

Introduction to Radiometry and Photometry

William Ross McCluney
Principal Research Scientist
Florida Solar Energy Center

Artech House
Boston • London

Library of Congress Cataloging-in-Publication Data

McCluney, William Ross.

Introduction to radiometry and photometry/William Ross McCluney.

Includes bibliographical references and index.

ISBN 0-89006-678-7

1. Radiation—Measurement 2. Photometry I. Title

QC795.42.M33 1994

539.2'028'7—dc20

94-7674

CIP

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

© 1994 ARTECH HOUSE, INC.

685 Canton Street

Norwood, MA 02062

All rights reserved. Printed and bound in the United States of America. No part of this book may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying, recording, or by any information storage and retrieval system, without permission in writing from the publisher.

International Standard Book Number: 0-89006-678-7

Library of Congress Catalog Card Number: 94-7674

10 9 8 7 6 5 4 3 2 1

*This book is dedicated to my mother, Luella Benjamin McCluney
and to the memory of my father, William Jones McCluney. Your
many years of hard work and sacrifice are not forgotten.*

Contents

Preface	xiii
Chapter 1 Fundamental Concepts of Radiometry	1
1.1 Electromagnetic Radiation	1
1.2 Terminology Conventions	3
1.3 Wavelength Notations and Solid Angle	4
1.4 Fundamental Definitions	7
1.5 Lambertian Radiators and Lambert's Cosine Law	13
1.6 Radiance, Irradiance, Intensity, and Flux Relationships	15
1.7 Connection with Electromagnetic Theory	20
1.8 Polarization	22
1.9 Photon Flux	25
Example Problem 1.1	28
Example Problem 1.2	30
References	32
Chapter 2 Fundamental Concepts of Photometry	33
2.1 Light	33
2.2 Photometric Definitions	37
2.2.1 Radiation Luminous Efficacy, K_r , and the V-lambda Function	41
2.2.2 Lighting System Luminous Efficacy, K_s	43
2.3 Luminance and Brightness	45
2.4 Luminance and Vision	47
2.5 Disability Glare	50
2.6 Discomfort Glare	52
2.7 Illumination	54
2.7.1 Illuminance Selection	55
Example Problem 2.1	57
Example Problem 2.2	59

Example Problem 2.3	61
References	61
Chapter 3 Blackbodies and Other Sources	63
3.1 Blackbody Radiation	63
3.2 Planck's Law	64
3.3 Wien Displacement Law	68
3.4 Luminous Efficacy of Blackbody Radiation	69
3.5 Color and Distribution Temperatures	71
3.6 Emission into an Imperfect Vacuum	72
3.7 Radiation Exchange	73
3.8 Experimental Approximation of a Blackbody	73
3.9 Other Real Sources	74
Example Problem 3.1	82
Example Problem 3.2	82
Example Problem 3.3	83
Example Problem 3.4	83
Example Problem 3.5	84
References	84
Chapter 4 Source/Receiver Flux Transfer Calculations	87
4.1 Introduction	87
4.2 Geometry and Definitions	87
4.2.1 Case 1	90
4.2.2 Case 2	91
4.2.3 Case 3	92
4.2.4 Case 4	93
4.2.5 Case 5	95
4.2.6 Case 6	96
4.2.7 Case 7	98
4.2.8 Monte Carlo Method	99
4.3 Configuration Factor	100
4.4 Net Exchange of Radiation	102
4.5 Summary	103
Example Problem 4.1	104
References	105
Chapter 5 The Invariance of Radiance and the Limits of Optical Concentration	107
5.1 Introduction	107
5.2 Radiance is a Field Quantity	107
5.3 Pencils of Rays	108
5.4 Elementary Beam of Radiation	109
5.5 Radiance Invariance	111

5.6	Radiance Invariance at an Interface	112
5.7	Radiance Through a Lens	114
5.8	Radiance in Absorbing and Scattering Media	115
5.9	Concentrating Radiance Meter	116
5.10	The Limits of Optical Concentration	120
	Example Problem 5.1	123
	Example Problem 5.2	124
	References	125
	Chapter 6 Optical Properties of Materials	127
6.1	Introduction	127
6.2	Terminology	128
6.3	Surface and Interface Optical Properties	130
6.3.1	Conductor Optical Properties	130
6.3.2	Nonconductor Optical Properties	131
6.3.3	Surface Emission Properties	132
6.3.4	Angular Dependence of Dielectric Optical Properties	136
6.3.5	Rough Surfaces	141
6.4	Bulk Medium Optical Properties	142
6.5	Properties of Plane Parallel Plates	148
6.5.1	Nonscattering Media	148
6.5.2	Scattering Media	154
6.6	Angular Dependence	156
6.7	Broadband Angle Properties	160
6.7.1	Transmittance and Reflectance Equations	160
6.7.2	Specular and Diffuse Optical Properties	162
6.8	Spectral Dependence	164
6.9	Broadband Spectral Properties	165
6.10	Spectral Selectivity	167
	Example Problem 6.1	174
	Example Problem 6.2	175
	References	175
	Chapter 7 The Detection of Radiation	179
7.1	Introduction	179
7.2	Basic Concepts	180
7.3	Classification of Detectors	187
7.3.1	Thermal Detectors	187
7.3.2	Photemissive Detectors	191
7.3.3	Semiconductor Devices	197
7.3.4	Multi-element Detectors, Charge Transfer Devices, and Imagers	205
7.4	Detector Noise	209

7.5	Signal Modulation and Radiation Chopping	211
7.6	Characterization of Detector Performance	215
7.6.1	Responsivity, R	216
7.6.2	Quantum Efficiency, η	216
7.6.3	Noise Equivalent Power, NEP	217
7.6.4	Detectivity, D	218
7.6.5	Photon Noise-Limited Performance	219
7.7	Flux Conditioning Prior to the Detector	220
7.7.1	Cosine Response Correction	221
7.7.2	Photopic Correction	224
7.7.3	Spectral Filtering	224
7.8	Signal Conditioning After the Detector	227
7.9	Detector Calibration	227
7.10	Example Detectors and Their Characteristics	229
	Example Problem 7.1	234
	References	237
	Appendix 7A	240
	Chapter 8 Optical Systems	243
8.1	Introduction	243
8.2	Optical Axis	244
8.3	Idealized (Thin) Lens Theory	245
8.4	Radiance and Irradiance of Images	250
8.5	Vignetting	253
8.6	Aberrations	254
8.6.1	Spherical Aberration	254
8.6.2	Chromatic Aberration	257
8.6.3	Distortion	258
8.6.4	Coma	258
8.6.5	Astigmatism	259
8.6.6	Field Curvature	261
8.6.7	Correcting Aberrations	261
8.6.8	The Diffraction Limit	261
8.7	Image Quality	263
8.8	Flux Distribution	265
8.9	Nonimaging Optical Systems	266
8.10	Throughput	269
8.11	Integrating Spheres	271
8.11.1	Cosine Correction	274
8.11.2	Transmissometers and Reflectometers	274
8.12	Monochromators	280
8.12.1	Spectral Filters	280

8.12.2 Scanning Monochromators	287
8.13 Windows	293
8.14 Sources	294
8.15 Goniometers	295
8.16 Transmissometers/Reflectometers	296
8.17 Scattering Meters, Nephelometers, Turbidimeters, and Haze Meters	296
Example Problem 8.1	297
References	300
 Chapter 9 Radiometers and Photometers	303
9.1 Introduction	303
9.2 General Design Factors	305
9.3 Broadband Irradiance and Radiance Meters	306
9.4 Restricted Spectral Band Irradiance Meters for the Ultraviolet Through the Infrared	310
9.5 Illuminance and Luminance Meters	311
9.6 Spectroradiometers	312
9.7 Calibration of Radiometers and Photometers	314
9.7.1 Transfer Standards	316
9.7.2 Broadband Irradiance Standard Sources	318
9.7.3 Standard Sources for Spectral Irradiance and Spectral Radiance	319
9.7.4 Absolute Radiometry	322
9.7.5 Standard Illuminance and Luminance Sources	326
9.7.6 Radiometer/Photometer Calibration Using Standard Sources	327
9.7.7 Spectroradiometer Calibration	328
9.7.8 National Standards Laboratories	329
Example Problem 9.1	329
Example Problem 9.2	330
References	330
 Chapter 10 Metric Primer and Additional Radiometric and Photometric Quantities and Units	333
10.1 Introduction	333
10.2 The SI System of Units	334
10.2.1 Basic Metric Principles	334
10.2.2 Metric Units for Radiometry and Photometry	336
10.3 The I-P System of Units	336
10.4 Photon Flux Units	336
10.5 Other Quantities and Units	338
References	340

Chapter 11 Basic Concepts of Color Science	343
11.1 Introduction	343
11.2 Basic Concepts and Definitions	344
11.3 Systems of Color Specification	349
11.3.1 Munsell Color System	349
11.3.2 CIE 1976 ($L^*a^*b^*$) Color Space	352
11.3.3 Tristimulus Colorimetry	352
11.4 CIE 1931 Color System	354
11.5 CIE 1964 Supplementary Observer Color System	359
11.6 CIE 1976 Uniform Color Space	359
11.7 Color Temperature	364
11.8 Standard Illuminants and Reflection Colorimetry	366
11.8.1 Blackbody Illuminants	367
11.8.2 Daylight Illuminants	369
11.8.3 Reflection Colorimetry	371
11.9 Color Rendering Index	372
References	375
Appendix Correspondence Between Finite Elements and the Calculus	377
A.1 Introduction	377
A.2 Definition of the Derivative	378
A.3 Definition of the Integral	380
A.4 Integrals as Sums	382
A.5 Sums Over Solid Angles	383
References	387
About the Author	389
Index	391

Preface

Radiometry is a system of language, mathematics, and instrumentation used to describe and measure the propagation of electromagnetic radiation, including the effects on that radiation of reflection, refraction, absorption, transmission, and scattering by material substances in their solid, liquid, and gaseous phases.

Photometry is a system used for the same purpose when the radiation is to be detected by the human eye.

Many introductory treatments of photometry tend to invite confusion rather than understanding. In some cases this accurately reflects a real confusion in the existing literature. The fundamental nature of the basic photometric quantities easily can be obscured by a seemingly inconsistent, or at least confusing, set of definitions and units.¹ This was well stated by Biberman in an article titled "Apples, Oranges, and Unlumens."²

Some of the confusion faced by students stems from presenting new concepts defined in terms of quantities that are themselves unfamiliar. It is hoped that the approach incorporated herein overcomes at least some of these problems, by defining first only the most basic of the radiometric quantities and then by showing one simple equation that represents the conversion of all relevant radiometric quantities to the corresponding photometric ones.

Students are often confronted by bewildering radiometric and photometric terminology appearing in the literature, including a large number of names for radiometric and photometric quantities and the use of different systems of units, many of which are unfamiliar in other disciplines. These names include: lumen, luminance, illuminance, lux, foot-candle, candela, nit, foot-Lambert, stilb, apostilb, and many more. Fortunately there is a trend toward increased uniformity in nomenclature and units, but valuable older references still contain outdated terminology. Chapter 10 is included to provide descriptions of some of the old usage and units

¹McCluney, W. R. "Radiometry and Photometry," *Amer. Jour. Phys.* Vol. 36, 977-979 (1968).

²Biberman, L. M., *Appl. Opt.*, Vol. 6, p. 1127 (1967).

and to offer some conversion factors and procedures for converting to the units in current use. It also offers a summary of the current international metric system, the *System Internationale* or SI system, used consistently in this book.

In recent years there has been a welcome move to standardize radiometric and photometric terminology, symbols, and nomenclature. The primary standard for this in the U.S. is "USA Standard Nomenclature and Definitions for Illuminating Engineering ANSI/IES RP-16-1986", published by the American National Standards Institute and the Illuminating Engineering Society of North America.³ This publication is in general agreement with the primary international standard for terminology in this field, the *International Lighting Vocabulary* published by the International Lighting Commission (CIE), and referenced repeatedly in this book.

The trend toward increased uniformity has been promoted by Fred Nicodemus in a multidecade devotion to increasing the understanding of radiometric concepts and the standardization of sensible terminology and units. Nicodemus' work culminated in a series of reports from the National Bureau of Standards (now the National Institute of Standards and Technology, NIST) making up the separate chapters of a work called "*Self-Study Manual on Optical Radiation Measurements*".⁴ Readers interested in exploring the fields of radiometry and photometry in more depth than is offered in this text are referred to this series of reports and to a number of additional treatments of the subject.⁵

Many publications, including textbooks on optical physics, dealing with electromagnetic radiation and its propagation in space speak of only one radiometric quantity (probably irradiance), calling it intensity and relating it to the square of the amplitude of an oscillating field quantity (the electric field or the magnetic field). These texts seldom explain in adequate detail how this quantity relates to

³IESNA, 345 East 47 Street, New York, NY 10017.

⁴Nicodemus, F. E., ed. *Self-Study Manual on Optical Radiation Measurements*, NBS Technical Notes 910-1 through 910-7, U. S. Department of Commerce, National Bureau of Standards (now National Institute of Standards and Technology), March 1976 through April 1984.

⁵Grum, F. *Optical Radiation Measurements, A Treatise*, Volumes 1 through 5, Academic Press, 1979–1083. Wyatt, C. L. *Radiometric Calibration: Theory and Methods*, Academic Press: New York, NY, 1978. Boyd, R. W. *Radiometry and the Detection of Optical Radiation*, John Wiley & Sons: New York, NY, 1983. Driscoll, W. G., ed. *Handbook of Optics*, McGraw-Hill Book Company: New York, NY, 1978, Section 1, Radiometry and Photometry. Nicodemus, F. E. "Radiance," *Am. Jour. Phy.* Vol. 31, 1963, p. 368. Meyer-Arendt, J. R. "Radiometry and Photometry: Units and Conversion Factors," *Applied Optics* Vol. 7, 1968, p. 2081. Tyler, J. E. "Applied Radiometry," in *Oceanography and Marine Biology, An Annual Review*, Vol. 11, Harold Barnes, ed., George Allen & Unwin Ltd.: London, 1973. Nicodemus, F. E., et. al. "Geometrical Considerations and Nomenclature for Reflectance," NBS Monograph 160, U. S. Department of Commerce, National Bureau of Standards (now National Institute for Standards and Technology), October 1977. Siegel, R., and J. R. Howell. *Thermal Radiation Heat Transfer*, 2nd ed., McGraw-Hill: New York, NY, 1981.

internationally standardized concepts in radiometry and photometry. A discussion of this subject can be found in Section 1.6.

It is hoped that this book will help readers overcome such problems by presenting in one relatively concise document the fundamental concepts of radiometry and photometry, the basic concepts of optical properties of materials, and some information on the uses of the radiometric and photometric quantities in various application areas. A number of worked examples are included at the end of most chapters to aid the student in mastering the subject quickly.

Here are some examples of general fields and application areas where radiometry and photometry have proven to be very useful, if not essential:

- Studies of radiative transfer in free space as well as in scattering and absorbing media, including natural bodies of water and planetary atmospheres.
- Studies of electromagnetic radiation produced in a variety of nuclear reactions and other sub-atomic processes, including that produced in particle accelerators.
- Studies of illumination, both electrically produced and from the sun and sky. These could include bioluminescence and light produced by chemical reactions.
- The design of electromagnetic radiation sources and measuring equipment or instruments for analytical or other purposes. Although specific treatment of radiation at microwave and radio frequencies is not included here, the concepts of radiometry can be used in these fields.
- The design of photographic, video, and other image capture and display equipment.

A few more specific applications can be mentioned, in order to show the variety of fields where radiometric and photometric concepts are used. These include:

- Water turbidity measurement instrumentation design and use.⁶
- Measurement of solar radiation for crop growth studies and for the design of solar energy collection and conversion systems.⁷
- Astronomical measurements of general electromagnetic radiation from stars, planets, and other luminous heavenly bodies.
- Environmental remote sensing, including aircraft and satellite measurements of upwelling radiation from the earth's atmosphere and surface⁸ and ground-

⁶McCluney, R., "Radiometry of Water Turbidity Measurements," *J. Water Poll. Contr. Fed.* Vol. 47, 1975, pp. 252–266.

⁷Duffie, J. A., and Beckman, W. A.. *Solar Engineering of Thermal Processes*, 2nd ed., John Wiley & Sons: New York, NY, p. 919, 1991.

⁸McCluney, W. R. "Remote measurement of water color," *Rem. Sens. of Environ.* Vol. 5, 1976, pp. 3–33.

based measurements of downwelling radiation emanating from the bottom of the atmosphere.

This book can be used as a reference text by people working in radiometry and photometry only intermittently and who need a concise introduction or refresher course periodically. It can also be used as a textbook in a short course on radiometry and photometry at the upper division college or graduate level, or for a short course offered by a university or professional society. As such, the book presumes that the reader has general knowledge of basic physics at the college level, including knowledge of electromagnetic wave propagation. The mathematical skills presumed on the part of the reader include algebra at the college level, high school geometry, integral calculus, and some familiarity with basic analytical geometry.

Readers possessing limited mathematical skills can skip the detailed derivations and focus mainly on the results obtained. For example, much of Chapter 5 on radiance invariance is devoted to a derivation of Equations (5.6), (5.14), and (5.16), which are the functional equations expressing the invariance of what is called "basic radiance." The sections of Chapter 5 subsequent to these equations are not as involved mathematically.

The author gratefully acknowledges Frances Rubenstein, whose interest led to the idea for writing this book, and to Paul Jindra for his painstaking review of portions of the manuscript and many helpful suggestions. Dr. Arthur Tarrant is kindly thanked for his many helpful suggestions for improving the manuscript. David Goebel and William E. Schneider are thanked for the material they provided about colorimetry and standards in radiometry and photometry. Dr. Laney R. Mills is thanked for his extensive and invaluable assistance in checking the worked examples and making numerous suggestions throughout the text. An associate professor of physics at the College of Charleston, Dr. Mills was on sabbatical leave at the National Center for Atmospheric Research when this assistance was provided.

Chapter 1

Fundamental Concepts of Radiometry

1.1 ELECTROMAGNETIC RADIATION

Radiation is the propagation of energy in a variety of forms through space. *Radiometry* is a system of concepts, terminology, mathematical relationships, measurement instruments, and units of measure devised to describe and measure radiation and its interaction with matter. *Photometry* is a similar system for radiation intended for detection by the human eye. Photometry is discussed in Chapter 2.

Although the concepts of radiometry can be used to describe nuclear radiation, including flux fields of energetic particles, they are generally used only for electromagnetic radiation. This book deals with the latter exclusively.

In radiometry and photometry, one is concerned with how radiation is distributed over the electromagnetic spectrum. Such a distribution is called a *spectral distribution*, or *spectrum*. Thus we begin our study of radiometry with the electromagnetic spectrum, illustrated in Figure 1.1, which has labels for different portions of the spectrum of interest in different fields of study. The International Commission on Illumination (*Commission Internationale de l'Eclairage*, or CIE) has standardized the terminology shown in Table 1.1 for the portions of the spectrum [1] containing and surrounding the visible portion dealt with most in the fields of radiometry and photometry. The spectral range of visible radiation does not have precise limits because these limits vary from person to person.

Monochromatic radiation is radiation having only one frequency and wavelength. Wavelength is given the Greek symbol λ and has units of meters, or the submultiples of the meter given in Figure 1.1. (More information on the metric system and metric terminology can be found in Chapter 10.) Frequency is given the Greek symbol ν and has units of cycles or periods per second. One cycle per second is called one hertz, abbreviated Hz.

The wavelength λ and frequency ν of a monochromatic beam are related by the equation

$$c = \lambda\nu \quad (1.1)$$

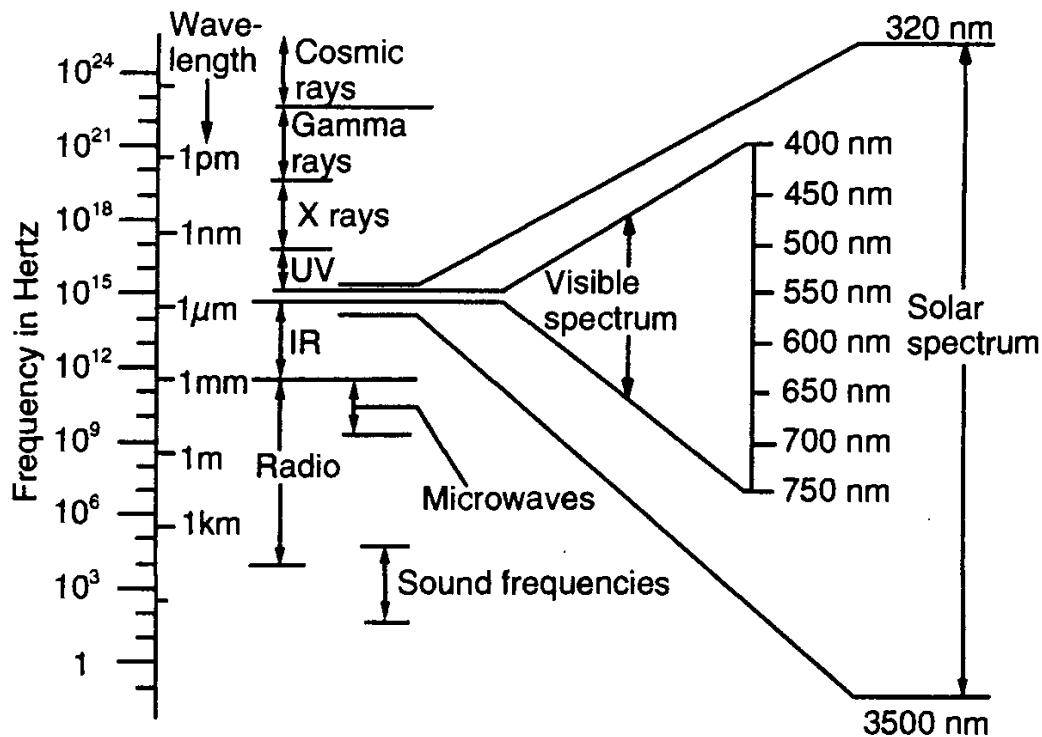


Figure 1.1 Portions of the electromagnetic spectrum of interest in many fields of study. “UV” means ultraviolet and “IR” means infrared.

Table 1.1
CIE Vocabulary for Spectral Regions

Name	Wavelength Range
UV-C	100 to 280 nm
UV-B	280 to 315 nm
UV-A	315 to 400 nm
VIS	Approximately 360–400 to 760–800 nm
IR-A*	780 to 1,400
IR-B	1.4 to 3 μ m
IR-C~	3 μ m to 1 mm

*Also called “near IR” or “NIR”

~Also called “far IR”

where c is the speed of light in the medium through which the radiation is propagating. The speed of light in a vacuum is $299,792,458 \text{ m} \cdot \text{s}^{-1}$. It is not much different from this value in air at sea level. One can select a wavelength (or frequency) of interest and use (1.1) to determine the corresponding frequency (or wavelength) in vacuum. The human eye is most responsive to radiation at about 555 nanometers

in wavelength, for example. The corresponding frequency is 5.4×10^{14} Hz. The eye is not at all responsive to radiation below 360 nm or above 830 nm.

Knowledge of the spectral distribution of radiation is crucial for many applications of radiometry and photometry. The optical properties of materials, such as transmittance, reflectance, and absorptance, are generally dependent upon the wavelength of radiation incident upon them or passing through them. One is therefore interested in spectral variations in the transmission, reflection, refraction, and absorption properties of materials.

There are four fundamental quantities in radiometry that are central to both radiometry and photometry. They are listed below along with their symbols and units in the metric system. The abbreviation sr refers to “steradian,” the unit of solid angle, described in Section 1.3.

- Radian flux Φ , in watts;
- Irradiance E , in watts/m²;
- Radian intensity I , in watts/sr;
- Radiance L , in watts/(sr · m²).

Radian flux is frequently also called *radiant power*, since dimensionally that is what it is. There is one additional quantity, radiant energy, which is needed for the definition of radiant flux and the remaining quantities listed above. Radiant energy is used alone in relatively few applications, but it is an important concept in radiometry that cannot be overlooked. However, in order to keep the fundamental definitions few in number and uncluttered by the large number of additional quantities that are used in the fields of radiometry and photometry, this chapter will concentrate on the four radiometric quantities listed above. The definitions of the four basic quantities provided in Section 1.4 are preceded by the definition of radiant energy at the beginning of that section.

The four photometric quantities that correspond to the above radiometric quantities are discussed in the next chapter, along with their symbols and units. They are:

- Luminous flux;
- Illuminance;
- Luminous intensity;
- Luminance.

1.2 TERMINOLOGY CONVENTIONS

Before beginning the definitions of the basic radiometric quantities, some preliminary information on notation conventions and on the concept of solid angle will be needed, and should make the definitions that follow more understandable. The symbols, units, and other nomenclature used herein conform to *USA Standard*

Nomenclature and Definitions for Illuminating Engineering, report RP-16, published in 1986 as a joint national standard of the American National Standards Institute (ANSI) and the Illuminating Engineering Society of North America (IESNA) [2]. The terminology used in this standard is entering common practice in a large number of fields dealing with optical radiation [3] and is in general agreement with the terminology standardized by the CIE [1]. According to Wyszecki and Stiles [4], "The CIE was founded in 1913 by taking over the functions of the earlier Commission Internationale de Photometrie, which was established in 1903. Since that time the CIE has grown into the international body that, by general consent in all countries, develops standards and standard procedures of measurement in virtually all fields of illuminating engineering and those fields that are directly relevant to it." The CIE is the primary international authority on terminology and basic concepts in radiometry and photometry, and its standards and technical reports are cited repeatedly in this book.

1.3 WAVELENGTH NOTATIONS AND SOLID ANGLE

When the wavelength symbol λ is used as a subscript on one of the radiometric quantities, say as Q_λ , the result denotes the concentration of the quantity at a specific wavelength as if one were dealing with a monochromatic beam of radiation at this wavelength and this wavelength only, meaning that the range of wavelengths $\Delta\lambda$ in the beam around the defining wavelength λ is infinitesimally small. Spectral radiant quantities can be thought of and defined as the mathematical derivative of the quantity Q with respect to wavelength λ , denoted as

$$Q_\lambda = \frac{dQ}{d\lambda} \quad (1.2)$$

which is equivalent to the quotient $\Delta Q/\Delta\lambda$ of the amount ΔQ of quantity Q contained in a small wavelength interval (say $[\lambda_1, \lambda_2]$) symbolized as $\Delta\lambda$, if the interval $\Delta\lambda$ between λ_1 and λ_2 is made to be infinitesimally small; that is to say, if one takes the limit of this ratio as $\Delta\lambda$ goes to zero:

$$Q_\lambda = \frac{dQ}{d\lambda} = \lim_{\Delta\lambda \rightarrow 0} \frac{\Delta Q}{\Delta\lambda} \quad (1.3)$$

This is a definition of the mathematical derivative Q_λ of the quantity Q with respect to wavelength λ , familiar to students of differential calculus. In order to denote the *distribution* of the radiometric quantity Q_λ over a *range* of wavelengths, it is therefore acceptable to show the quantity as a function of wavelength: $Q_\lambda(\lambda)$. Properties of materials, such as reflectance ρ or transmittance τ , also vary with and

may be shown as functions of the wavelength, as in $\tau(\lambda)$. However, since they are not distributions of flux (only weighting functions) and cannot be defined as a derivative with respect to wavelength, they are *not* to be subscripted with the wavelength symbol. This same convention applies to the directional dependencies of such properties. The directional emittance of a surface (introduced in Chapter 6), for example, would be shown as $\epsilon(\theta, \varphi)$, where θ and φ are suitably chosen plane angles specifying a direction of interest from a point on the surface.

Plane angle is defined by two straight lines intersecting at a point. The space between these lines in the plane defined by the lines, is the plane angle. It is measured in radians (2π radians to a circle) or degrees (360 degrees to a circle).

In preparation for defining solid angle, it can be pointed out that two lines meeting at a point can be thought of as linear projections of two points in space onto the point of intersection. If the two points lie on a circle centered at the point of intersection, their linear projections to this point can be thought of as *radial projections*, projections along radii to the center of the circle. Thus, we can give an alternative definition to the plane angle:

A *plane angle* is the arc length s of a radial projection of a segment of a curve C in a plane *onto a circle of unit radius* (called a unit circle) lying in that plane and centered at the vertex point P about which the angle is being defined.

The drawing in Figure 1.2 is used to define the plane angle more generally as follows. It is the quotient of the length s of the radial projection of curve C onto

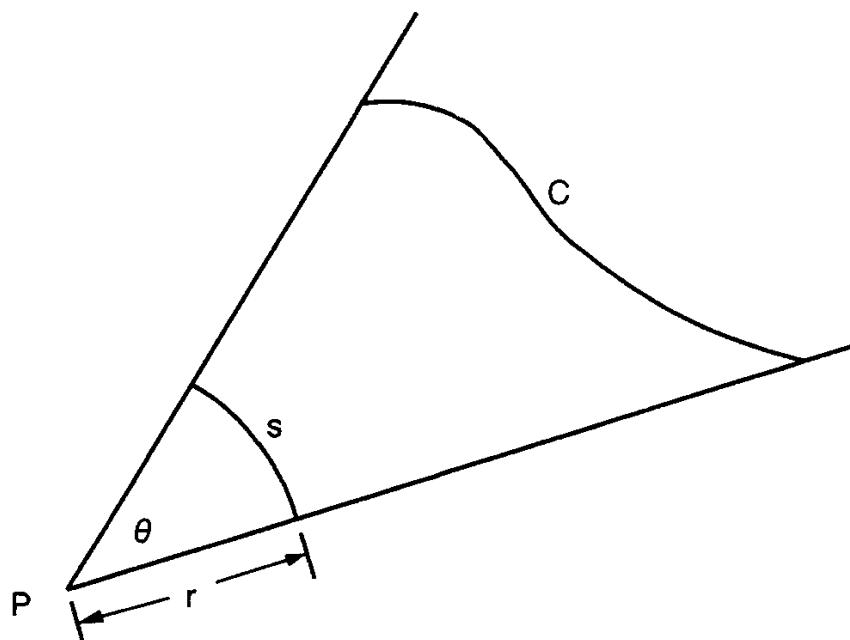


Figure 1.2 Illustration for defining plane angle.

a circle of radius r centered at the point P of definition of the plane angle and that radius. That is, it is the arc length s divided by the radius r :

$$\theta = \frac{s}{r} \quad (1.4)$$

According to this formula, the plane angle is dimensionless. However, to aid in communication, it has been given the unit: *radian*. The radian measure of an angle, as defined in (1.4), can be converted to degree measure by multiplying by a conversion constant ($180/\pi$). Since the circumference of a circle is 2π times its radius, a circle can be thought of as subtending a plane angle of 2π radians. A similar approach can be used to define solid angle:

Solid angle is defined by a closed curve and a point in space. Its magnitude is the area *on a sphere of unit radius* of the radial projection of the closed curve, the projection being from each point on the curve to the point from which the solid angle is to be measured.

Equivalently, the solid angle is defined by the quotient of the area on a sphere of radius R of the projection of the curve onto the sphere and R^2 , as illustrated in Figure 1.3. This quotient gives a dimensionless quantity. However, to aid in communication, the solid angle has been given a unit, called the *steradian*, abbreviated sr.

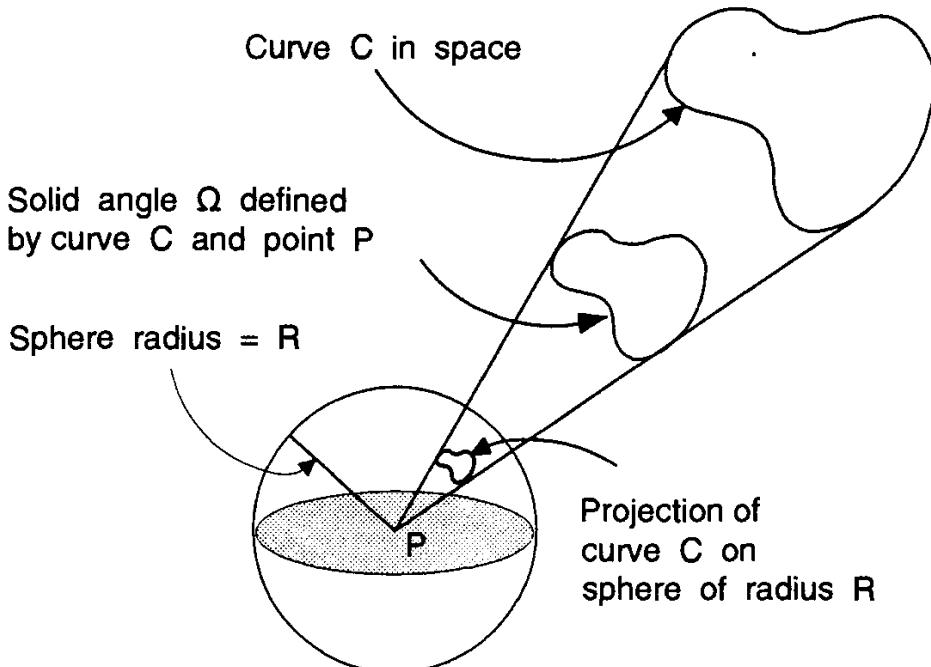


Figure 1.3 Drawing illustrating the definition of solid angle. The curve C in space subtends solid angle Ω at point P.

The area of a sphere is 4π times the square of its radius and the radius of a unit sphere is unity. Thus a sphere of any radius subtends a solid angle of 4π sr. A hemisphere of any radius subtends 2π sr. The language used in speaking of solid angles is important to convey the needed information about the solid angle properly. Since a solid angle is defined by a closed curve and a point in space, one must mention both the curve and the point whenever describing a solid angle in words. For example, one speaks of the solid angle formed, or subtended, by the aperture of a lens at its focus, the “curve” part of the definition clearly being the boundary of the aperture and the “point” part being the focal point of the lens.

Problem 1.1 at the end of this chapter illustrates the use of the solid angle definition to derive the following formula for the solid angle Ω within a circular cone of half angle α :

$$\Omega = 2\pi(1 - \cos \alpha) \quad (1.5)$$

For a radially symmetric lens with a centered circular aperture whose radius subtends the angle α with respect to the lens axis relative to the focus of the lens, (1.5) gives the solid angle subtended by that aperture at the focal point. For other geometries, the computation of the solid angle can become more complex and the reader is referred to the relevant literature on the subject [5–13].

1.4 FUNDAMENTAL DEFINITIONS

The definition of radiant energy is now introduced, followed by definitions of the four fundamental quantities of radiometry and, in the next chapter, a generic equation that can be used to convert each of these radiometric quantities to its corresponding photometric version. Additional radiometric and photometric quantities are introduced in Chapter 10. That chapter also contains a discussion of some discontinued photometric units and terminology found in older treatments.

Radiant energy, Q , is the quantity of energy propagating onto, through, or emerging from, a specified surface of given area in a given period of time. All wavelengths contained in the radiation are to be included. Alternatively, if a limited wavelength range is to be considered, this range must be stated (units: joule).¹ Radiant energy is of interest in applications involving pulses of electromagnetic radiation where not only the instantaneous rate of flow, or flux, of the radiation is important, but also the total quantity of energy delivered in the pulse over a

¹The international system of units (*System Internationale*, or SI) called the metric system is based on seven fundamental, or base, units of measurement: meter, kilogram, second, ampere, kelvin, mole, and candela. In this system, energy is a derived quantity and the primary unit of energy is the joule. In terms of the base units, a joule is a $\text{Kg} \cdot \text{m}^2 \cdot \text{sec}^{-1}$. The joule is equivalent to the Newton meter. Note that the photometric unit, candela, discussed in the next chapter, is one of the seven base units.

specific and restricted period of time. This is the case, for example, when absorbed radiation stimulates chemical or biological changes. In such cases, the integrated dose of radiation over a period of time can be as important as the rate of flow of the energy involved. A related term is “radiant exposure,” which is mentioned in Chapter 10.

Spectral radiant energy, Q_λ , is the radiant energy per unit wavelength interval at the specific wavelength λ (units: joule · nm⁻¹). Using the notation of differential calculus the defining equation is

$$Q_\lambda = \frac{dQ}{d\lambda} \quad (1.6)$$

Radiant flux (power), Φ , is the time rate of flow of radiant energy (unit: watt). (One watt is one joule per second: 1W = 1 J·s⁻¹).² The defining equation is

$$\Phi = \frac{dQ}{dt} \quad (1.7)$$

The flux or flow of radiation is described by the quantity of energy transferring through a surface or region of space per unit time. The flow of or use of energy is indicated by the quantity of energy flowing or used over a period of time, divided by the time interval. This is called *power* and it has units of energy per unit time. Whenever radiation is made to fall on a device that produces an electronic current, voltage, or other signal that is proportional to the magnitude of the radiation incident on the device (called a *detector* or *sensor*), it is generally the total quantity of flux over the whole device that is important, rather than the flux per unit area. When speaking of a certain quantity of radiant flux, the spatial extent of the radiation field containing the flux should be described.

Spectral radiant flux (power), Φ_λ , is the radiant flux per unit wavelength interval at the wavelength λ (units: watt · nm⁻¹ or J · nm⁻¹ · s⁻¹). It may also be written as $\Phi_\lambda(\lambda)$. The defining equations are

$$\Phi_\lambda = \frac{dQ_\lambda}{dt} = \frac{d\Phi}{d\lambda} \quad (1.8)$$

²Although the metric system is used throughout this book, in some fields and in older books the U.S. customary system of units, also called the *inch-pound system*, is used rather than the modern metric system. Some conversions are provided in Chapter 10. For the purpose of comparing flux values described here in metric units with their inch-pound counterparts, the British thermal unit, or Btu, is the inch-pound unit of energy and the Btu per hour is the unit of power or flux. There are 1054.35 joules in a BTU and 1 watt equals 3.414 Btu per hour.

The spectral distribution of radiant flux is important for most applications, including calculations of the response of a detector to radiation incident on its sensitive surface. Detectors are discussed in Chapter 7.

Irradiance, E , is the area density of radiant flux, the radiant flux per unit area in a specified surface that is incident on, passing through, or emerging from a point in the specified surface. All directions in the hemispherical solid angle above or below the point in the surface are to be included. See Figure 1.4 (units: $\text{watt} \cdot \text{m}^{-2}$). The defining equation is

$$E = \frac{d\Phi}{ds_o} \quad (1.9)$$

where $d\Phi$ is an element of radiant flux and ds_o is an element of area in the surface.

Irradiance is a function of position on the specified surface. The irradiance leaving a surface can be called *exitance*, M , and it has the same units and defining equation. (The exitance has in the past been called *emittance* but this term is now generally being reserved as a replacement term for *emissivity*, a property of a material surface, to be defined in Chapter 6.) The irradiance is generally thought to refer only to flux incident upon a surface, excluding flux leaving the surface, but since there is no mathematical distinction between these two cases, the term irradiance is used in this book for the flux per unit area incident upon and/or leaving a point in a described surface. One reason for this is the occasional case in which the surface in question is a purely geometrical one through which a beam of radiation is passing without alteration, save for the fact that the rays are approaching

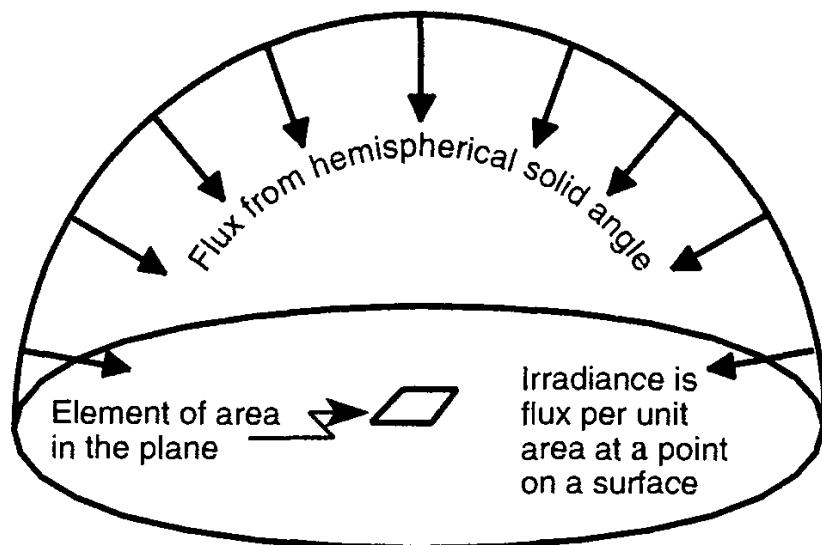


Figure 1.4 Irradiance refers to radiant flux per unit area incident at a point in a surface coming from a hemispherical solid angle.

the point of definition on one side of the surface and leaving this point on the other side.

The flux incident on a surface can come from any direction in the hemispherical solid angle above the surface at the point in question, or all of them, in this definition. The flux can also be that leaving the surface in any direction in the hemispherical solid angle, or all directions in that solid angle.

Irradiance has meaning only with respect to a point on a surface in space. When speaking of irradiance, one should be careful both to describe the surface and to indicate at which point on the surface the irradiance is being evaluated. Specification of the specific point on the surface can be omitted if it is known or if it can be presumed that the irradiance is constant over the portion of the surface of interest. For example, one could speak of that portion of the total irradiance on the outside wall of a building that comes from the solar disk. One could also speak of the irradiance on the wall from the sky, and that portion from the ground and other surrounding objects. In each case, we are describing the surface and either implying that there is one point or a small region on the wall that is of interest, or that we are considering these irradiances as uniform over the area of the wall.

Irradiance is the most important quantity for describing radiation incident on or leaving a surface when it is not essential to describe the angular or directional distribution of that radiation in detail (see Fig. 1.4).

Spectral irradiance, E_λ , is the radiant flux per unit area and per unit wavelength interval at the wavelength λ at a point in a specified surface (units: $\text{watt} \cdot \text{m}^{-2} \cdot \text{nm}^{-1}$ or $\text{J} \cdot \text{m}^{-2} \cdot \text{nm}^{-1} \cdot \text{s}^{-1}$). It may also be written as $E_\lambda(\lambda)$. The defining equations are

$$E_\lambda = \frac{dE}{d\lambda} = \frac{d^2Q_\lambda}{ds_o dt} = \frac{d^2\Phi}{ds_o d\lambda} \quad (1.10)$$

Radiant intensity, I , is the solid angle density of radiant flux, the radiant flux per unit solid angle incident on, passing through, or emerging from a point in space and propagating in a specified direction (units: $\text{watt} \cdot \text{sr}^{-1}$). The defining equation is

$$I = \frac{d\Phi}{d\omega} \quad (1.11)$$

where $d\Phi$ is the element of flux incident on or emerging from a point within element $d\omega$ of solid angle in the specified direction.

Radiant intensity is a function of direction from or toward the point for which it is defined. For most real sources, it is a strongly varying function of direction. Intensity is a useful concept for point sources (or sources that are very small

compared with the distance the observer or detector is from the source) such as stars in the sky and tiny filaments of electric light bulbs (called *lamps*), but it is generally inappropriate to apply this term to extended sources such as the daytime sky or large light fixtures (known as *luminaires* to the illuminating engineer). An exception would be in a discussion of the concept of extended sources, describing them as being made up of an infinite number of point sources. The term intensity is frequently used in certain subdisciplines of physics, such as optics. The quantity called intensity in most of these discussions is irradiance (or its photometric counterpart illuminance). Intensity is a very specific term in the field of radiometry that is seldom used for anything but point sources of radiation. Its use for irradiance outside this field can cause confusion for beginning students of radiometry and photometry. The connection between irradiance and electromagnetic field quantities is discussed in Section 1.7.

Since intensity is defined in terms of a point in space and a direction from that point, when speaking of intensity it is important to state what the point is and which direction or multiple directions are of interest. For example, the angular distribution of intensity emanating from the light fixtures in street lamps is of great importance in designing a road lighting system. When speaking of the intensity from such a lamp, we clearly imply that the lamp itself is the “point” of definition and the direction is specified with respect to some reference direction relative to the lamp. (Street lamps can be treated as “point sources” without significant error when one is mostly concerned with the flux at considerable distances from the lamps, distances great enough that the lateral extent of the lamp is very much smaller than the distance to it.)

Spectral radiant intensity, I_λ , is the radiant flux per unit solid angle and per unit wavelength interval at the wavelength λ incident on, passing through, or emerging from a point in space in a given direction (units: $\text{watt} \cdot \text{sr}^{-1} \cdot \text{nm}^{-1}$ or $\text{J} \cdot \text{sr}^{-1} \cdot \text{nm}^{-1} \cdot \text{s}^{-1}$). It may also be written as $I_\lambda(\lambda)$. The defining equations are

$$I_\lambda = \frac{dI}{d\lambda} = \frac{d^2Q_\lambda}{d\omega dt} = \frac{d^2\Phi}{d\omega d\lambda} \quad (1.12)$$

Radiance, L , is the area *and* solid angle density of radiant flux, the radiant flux per unit projected area and per unit solid angle incident on, passing through, or emerging in a specified direction from a specified point in a specified surface (units: $\text{watt} \cdot \text{m}^{-2} \cdot \text{sr}^{-1}$). The defining equation is

$$L = \frac{d^2\Phi}{d\omega ds} = \frac{d^2\Phi}{d\omega ds_o \cos \theta} \quad (1.13)$$

where $ds = ds_o \cos \theta$ is a quantity called the projected area, the area of the projection of elemental area ds_o in the surface containing the point at which the

radiance is being defined, projected in the direction of propagation onto a plane perpendicular to this direction. $d\omega$ is an element of solid angle in the specified direction and θ is the angle between this direction and the normal to the surface at the specified point (see Figure 1.5).

The projected area is the projection ds of the element ds_o of area in the specified surface onto a plane perpendicular to the given direction, as illustrated in Figure 1.5. Radiance can be thought of as the intensity per unit projected area or as irradiance per unit solid angle from the projected area.

Radiance is a function of both position and direction. For many real sources, it is a strongly varying function of direction. It is the most general quantity for describing the propagation of radiation through space and transparent or semi-transparent materials. The radiant flux and irradiance can be derived from the radiance by the mathematical process of integration over a finite surface area and/or over a finite solid angle, as will be demonstrated in Section 1.6. Chapter 5 is devoted to the subject of radiance invariance and Chapter 8 to the use of radiance in analyzing the propagation of flux through optical systems.

Since radiance is a function of position in a defined surface as well as the direction from it, it is important when speaking of radiance to specify the surface, the point in it, and the direction from it. All three pieces of information are important for the proper specification of radiance. For example, we may wish to speak of the radiance emanating from a point on the ground and traveling upward

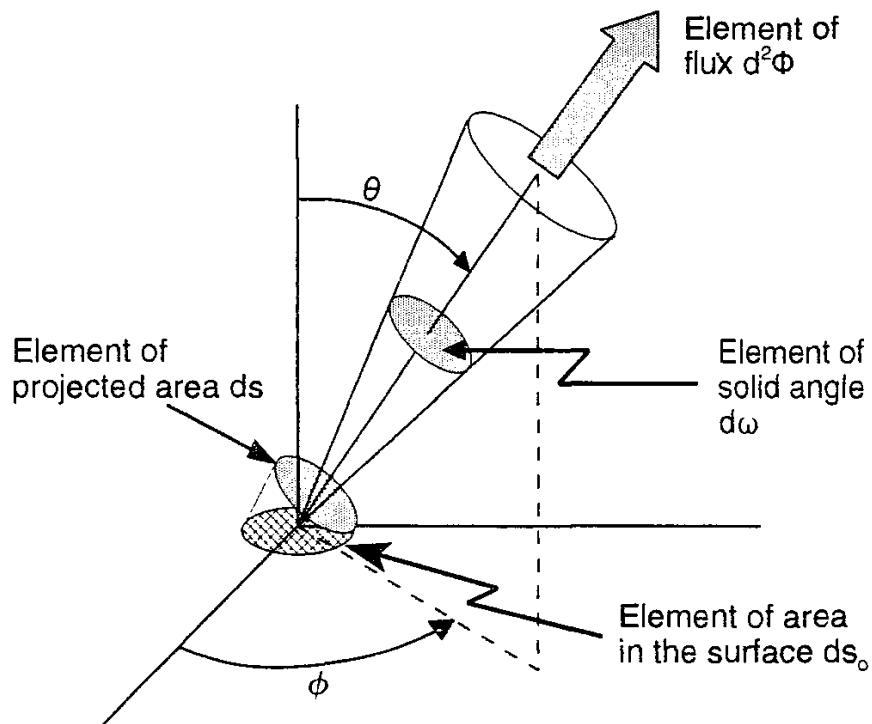


Figure 1.5 Geometry for the definition of radiance.

toward the lens of a camera in an airplane or satellite traveling overhead. We specify the location of the point, the surface from which the flux emanates, and the direction of its travel toward the center of the lens.

Since the words “radiance” and “irradiance” can sound very similar in rapidly spoken or slurred English, one can avoid confusion by getting in the habit of always speaking of the *point* and the *surface* that is common to both concepts, and then to clearly specify the *direction* when talking about radiance.

Spectral radiance, L_λ , is the spectral concentration of radiance, the radiant flux per unit projected area, per unit solid angle, and per unit wavelength interval incident on, passing through, or emerging in a specified direction from a specified point in a specified surface (units: $\text{watt} \cdot \text{m}^{-2} \cdot \text{sr}^{-1} \cdot \text{nm}^{-1}$). It may also be written as $L_\lambda(\lambda)$. The defining equations are

$$L_\lambda = \frac{dL}{d\lambda} = \frac{d^3\Phi}{d\omega ds_o \cos \theta d\lambda} \quad (1.14)$$

with θ being the angle between the specified direction and the normal to the surface at the specified point.

Some representative values of the radiometric quantities defined above are given in Table 1.2.

1.5 LAMBERTIAN RADIATORS AND LAMBERT'S COSINE LAW

Chapter 5 on the optical properties of materials deals with various definitions of surface optical properties such as transmittance and reflectance. However, one very

Table 1.2
Representative Values of Radiometric Quantities

Quantity	Value
Total radiant flux from a 100W tungsten incandescent light bulb, gas-filled, coiled coil filament	82W
Radiant flux output of typical, medium power, helium-neon laser, emission at 632.8 nm	5 mW
Radiant flux output from a 40W fluorescent lamp	23.2W
Extraterrestrial solar irradiance at mean earth orbit	1367 W/m ²
Terrestrial direct normal solar irradiance, clear sky, southeastern U.S., winter, solar noon	852 W/m ²
Terrestrial global (hemispherical) solar irradiance, on horizontal plane, clear sky, southeastern U.S., winter, solar noon	686 W/m ²
Radiance of the sun at its surface	$2.3 \times 10^7 \text{ W} \cdot \text{m}^{-2} \cdot \text{sr}^{-1}$
Apparent radiance of the sun from earth's surface	$1.4 \times 10^7 \text{ W} \cdot \text{m}^{-2} \cdot \text{sr}^{-1}$

important concept is needed here. Any surface, real or imaginary, whose radiance is independent of direction is said to be a *Lambertian radiator*, mainly because it obeys Lambert's cosine law, which is that the irradiance (or exitance) from an element of area in the surface varies as the cosine of the angle θ between that direction and the perpendicular to the surface element: $E(\theta) = E(0) \cos \theta$. The radiance of a Lambertian radiator is independent of direction. Because of this, a Lambertian radiator can be thought of as a window onto an isotropic radiant flux. This is Lambert's cosine law.

There is another version of the cosine law that is a kind of inverse of the previous one. It has to do not with the radiance *leaving* a surface but with how radiation from a uniform and collimated beam (a beam with all rays parallel to each other and equal in strength) *incident on* a plane surface is distributed over that surface as the angle of incidence changes. This is illustrated with the following problem: A horizontal rectangle of length L and width W receives flux from a homogeneous beam of collimated radiation of irradiance E_o , making an angle θ with the normal (perpendicular) to the plane of the rectangle, as shown in Figure 1.6. How does the average irradiance over the horizontal rectangle vary with the angle θ of incidence?

Here is the solution to this problem: Let Φ_o be the flux over the projected area A_o , given by E_o times A_o . This same flux Φ_o will be falling on the larger horizontal area $A = L W$, producing horizontal irradiance $E = \Phi_o/A$. The flux is the same on the two areas. Equating them gives

$$EA = E_o A_o. \quad (1.15)$$

From the geometry of Figure 1.6, $A_o = A \cos \theta$. Substituting this into (1.15) and solving for E yields

$$E = E_o \cos \theta \quad (1.16)$$

This function is plotted in Fig. 1.7. Although this problem speaks of the irradiance falling *on* a surface, if the surface is perfectly transparent, or even imaginary, it will also describe the irradiance (or exitance) *emerging from* the other side of the surface. If the collimated flux on the surface were coming from all directions simultaneously, then the emerging flux on the other side would be described as Lambertian and the surface could be called a Lambertian radiator. The fact that the surface itself is not really emitting the radiation—it is coming from a source (or infinite number of sources) on the other side—does not change the nature of the radiation emerging from the surface. An emitting surface is mathematically indistinguishable from an imaginary surface through which the same pattern of radiation is passing. This is how this latter cosine law is related to Lambert's Cosine Law. They are basically two manifestations of the same thing.

One consequence of the cosine effect is that the output of a perfect irradiance meter illuminated uniformly with collimated radiation that fully fills the sensing area will decrease with the cosine of the angle of incidence as that angle increases from zero, according to the curve plotted in Figure 1.7.

By the same reasoning, if a detector receives flux from a distant finite source of isotropic radiation (a Lambertian radiator), the amount of this flux will decrease with the cosine of the angle between the perpendicular to the source surface and the line from the source point to the detector, because the size of the projected area of this finite source decreases with angle. This effect is the origin of the $\cos \theta$ term in the denominator of (1.13). For example, suppose a telescope is pointed toward a flat Lambertian luminous plate in orbit around the earth. Suppose the telescope is not powerful enough to resolve the plate—it appears as a point source in the image—but that a detector in the focal plane of the telescope can measure the flux from the plate received by the telescope. Suppose also that the plate is spinning about an axis perpendicular to the line of sight. The flux measured by the detector will be sinusoidal with respect to time because of Lambert's Cosine Law.

1.6 RADIANCE, IRRADIANCE, INTENSITY, AND FLUX RELATIONSHIPS

Radiance and irradiance are quite different quantities. Radiance describes the angular distribution of radiation while irradiance adds up all this angular distribution

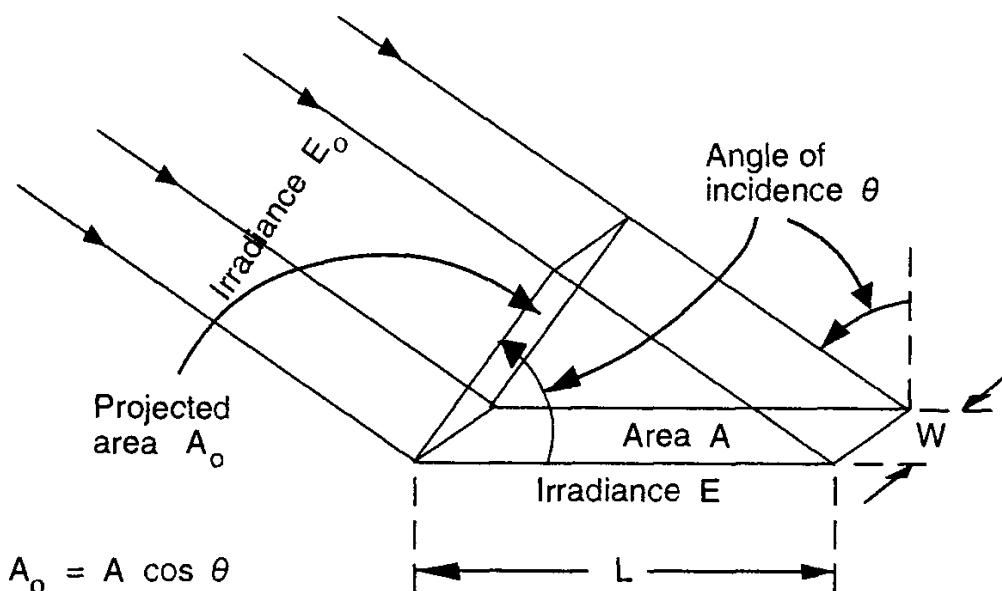


Figure 1.6 Geometry for deriving the cosine law.

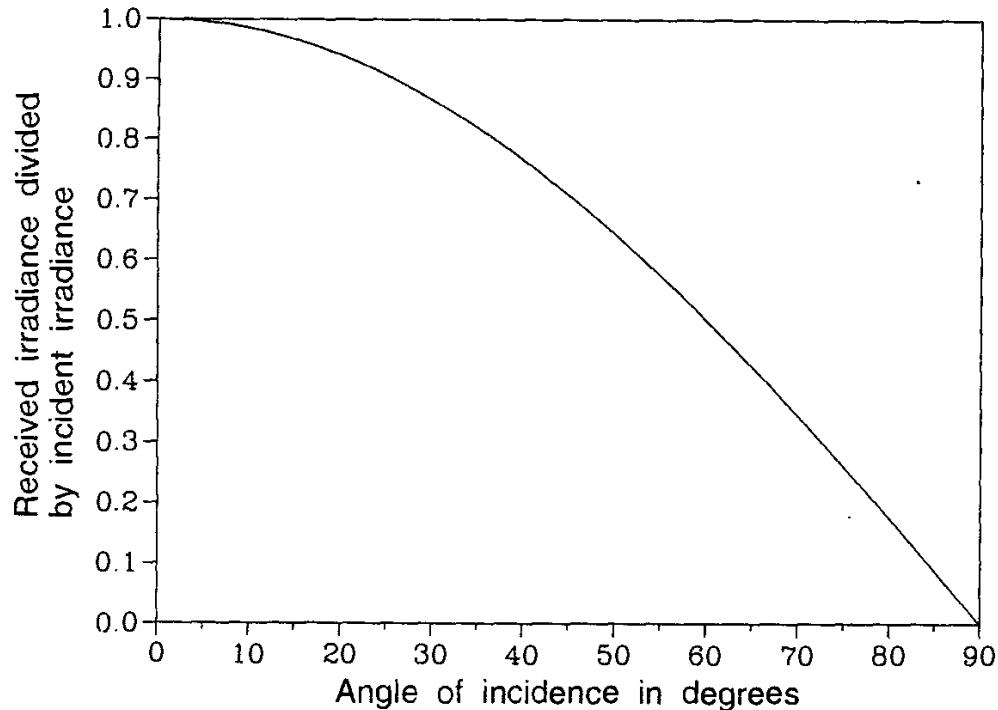


Figure 1.7 Irradiance ratio versus angle of incidence.

over a specified solid angle Ω and lumps it together. The fundamental relationship between them is embodied in the equation

$$E = \int_{\Omega} L(\theta, \phi) \cos \theta d\omega \quad (1.17)$$

for a point in the surface on which they are defined.

This equation involves an integral over a finite solid angle Ω . If $\Omega = 0$, there is no solid angle and there is no irradiance! When we speak of a collimated beam of some given irradiance, say E_o , we are talking about the irradiance contained in a beam of nearly parallel rays, but which necessarily have some small angular spread in them, filling a small but finite solid angle Ω , so that (1.17) can be finite. A perfectly collimated beam can contain no irradiance, because there is no directional spread to the radiation. Perfect collimation is a useful concept for theoretical discussions, however, and it is encountered frequently in optics.

A perfectly collimated beam can carry a nonzero quantity of flux only by a special mathematical trick attributed to physicist P. A. M. Dirac. The Dirac delta function is defined here to be the mathematical function $\delta(\theta, \phi)$ having the properties:

$$\delta(\theta, \phi) = \begin{cases} \infty & \text{for } \theta = 0 \text{ and } \phi = 0 \\ 0 & \text{for } \theta \neq 0 \text{ and } \phi \neq 0 \end{cases}$$

$$\int_0^{2\pi} \int_0^{\pi/2} \delta(\theta, \phi) d\theta d\phi = 1.0$$

If the radiance L_c of a collimated beam is expressed as $L_c = E_o \delta(\theta, \phi)$, then the beam's irradiance will be E_o .

Taking Ω to be 2π —the whole hemispherical solid angle—(1.17) indicates that the angular distribution of radiance at a point $L(\theta, \phi)$ is to be integrated over 2π sr. Using a spherical coordinate system with its origin at the point in question on a surface, as illustrated in Fig. 1.8, the element of solid angle is given by $\sin \theta d\theta d\phi$, so that the integral of (1.17) over the hemisphere becomes

$$E = \int_0^{2\pi} \int_0^{\pi/2} L(\theta, \phi) \cos \theta \sin \theta d\theta d\phi \quad (1.18)$$

If the radiance $L(\theta, \phi)$ is constant over the range of integration (over the hemispherical solid angle), then it can be removed from the integral and the final result is

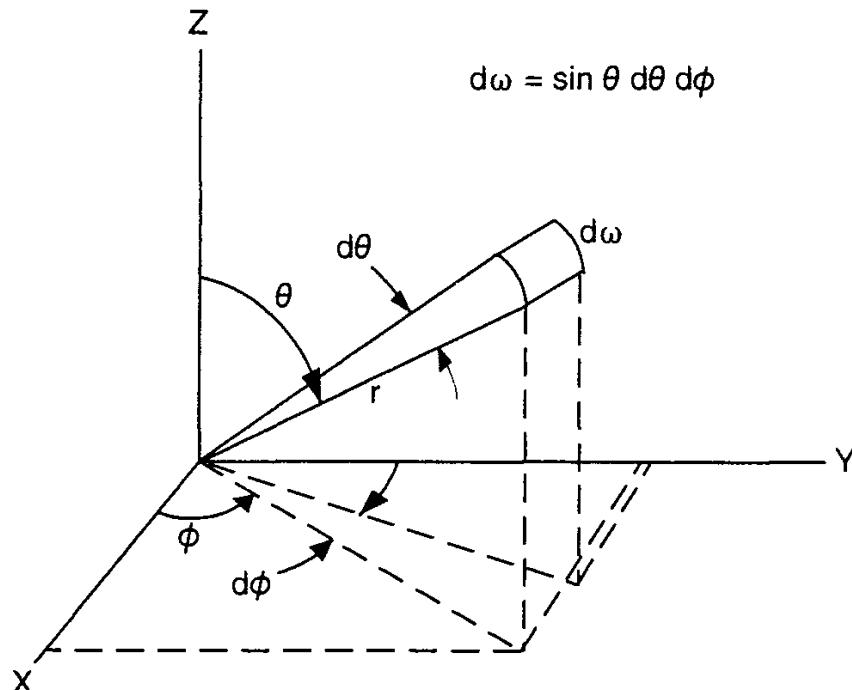


Figure 1.8 Geometry for defining element of solid angle in spherical coordinates.

$$E = \pi L \quad (1.19)$$

for a point on a surface whose radiance is constant over direction. As we have seen, such a surface is called a Lambertian surface and (1.19) applies only to such surfaces. The reader can verify analytically or by computation that

$$\int_0^{2\pi} \int_0^{\pi/2} \cos \theta \sin \theta d\theta d\phi = \pi \quad (1.20)$$

For readers curious as to the origin of the $\cos \theta$ factor in (1.17), it arises from the need for projected area in the definition of radiance. This equation can be derived from the definition of radiance as follows. Solving (1.13) for $d^2\Phi$, we have

$$d^2\Phi = L \cos \theta ds_o d\omega \quad (1.21)$$

Dividing this by ds_o , we can write

$$d\left(\frac{d\Phi}{ds_o}\right) = \frac{L \cos \theta ds_o d\omega}{ds_o} \quad (1.22)$$

which, from the definition of irradiance, is the element dE of irradiance produced by the element $d\omega$ of solid angle

$$dE = L \cos \theta d\omega \quad (1.23)$$

Integrating this yields (1.17). Equations (1.17) and (1.18) show how to convert from radiance L to irradiance E . The next step is to show relationships between intensity and radiance and irradiance.

It is noted that substituting the differential $d^2\Phi/d\omega = dI$ from (1.11) into (1.13) yields

$$L = \frac{dI}{ds_o \cos \theta} \quad (1.24)$$

Solving for dI and integrating both sides over the source area yields

$$I = \int_{S_o} L \cos \theta ds_o \quad (1.25)$$

However, intensity is normally only applied to point sources. Thus, it should be understood that S_o in (1.25) is very small in relationship to the distance to the point where the intensity I is observed.

Solving (1.11) for $d\Phi = Id\omega$, writing $d\omega$ as da_o/R^2 , and dividing both sides by da_o yields the expression

$$E_o = \frac{I d\omega}{da_o} = \frac{I da_o}{da_o R^2} = \frac{I}{R^2} \quad (1.26)$$

for the irradiance E_o a distance R from a point source of intensity I on a surface perpendicular to the line between the point source and the surface where E_o is measured. This is an explicit form for what is known as the *inverse square law* for the decrease in irradiance with distance from a point source.

Next comes the conversion from radiance L to flux Φ . Let the dependence of the radiance on position over the surface for which it is defined be indicated by generalized coordinates (u, v) in the plane of the surface of interest. Let the directional dependence be denoted by (θ, ϕ) , so that L may be written as a function $L(u, v, \theta, \phi)$ of position and direction. Integrating (1.21) over both the area S_o of the surface and the solid angle Ω of interest yields

$$\Phi = \int_{S_o} \int_{\Omega} L(u, v, \theta, \phi) \cos \theta d\omega ds_o \quad (1.27)$$

In this equation, the lower case ω is used to identify an *element* of solid angle $d\omega$ and the upper case Ω to identify a finite solid angle. As mentioned previously, in spherical coordinates $d\omega$ is given by $\sin \theta d\theta d\phi$, so that, if we let the solid angle Ω over which we integrate extend to the full 2π sr of the hemisphere, we have the total flux emitted by the surface in all directions

$$\Phi = \int_{u1}^{u2} \int_{v1}^{v2} \int_0^{2\pi} \int_0^{\pi/2} L(u, v, \theta, \phi) \cos \theta \sin \theta d\theta dv du \quad (1.28)$$

If the surface is a planar one, and we describe points in the surface by their rectangular coordinates (x, y) , then u and v would be replaced by x and y in (1.28) and $u1, u2, v1$, and $v2$ would be replaced by their x - y counterparts.

This completes the definitions of the four most fundamental quantities of radiometry. At the end of this chapter are some sample problems that illustrate the use of these concepts. In the next chapter, we introduce the corresponding photometric quantities and their units.

In many cases one cannot or does not always want to determine the *integrated* value of a spectrally or angularly distributed quantity. Or perhaps it is the angular or spectral distribution itself that is of interest. In such cases, one would like to be able to measure the full spectral or angular distribution, such as E_λ or $L(\theta, \phi)$. However, infinitesimally small quantities cannot be measured. In these situations it is preferred, *or required by the limitations of the measuring device*, to perform

measurements of some quantity B , symbolized by B_i , at discrete wavelengths or angles, represented generically by x_i . Then the integrated quantity B is approximated by calculating the sum of the discrete measurements as follows:

$$B = \int_{x_1}^{x_2} B(x) dx \approx \sum_{i=1}^N B_i \Delta x_i \quad (1.29)$$

If all the Δx 's are of the same width, Δx , then

$$B = \int_{x_1}^{x_2} B(x) dx \approx \Delta x \sum_{i=1}^N B_i \quad (1.30)$$

Using this approach, all the integrals shown above can be approximated with sums as indicated in (1.29). This is discussed further in the Appendix.

This discussion and the one in Appendix show readers not very familiar with integral calculus that the integrals given in the above equations (and subsequently throughout this book) are equivalent in concept to sums of the quantities being integrated, multiplied by the interval in the dependent quantity separating the discrete values of the quantities being summed. Another way of putting it is that the integral of Q_λ over λ is the “area” under a plot of Q_λ versus λ .

As we will see in Chapter 5, real problems of flux transfer in radiometry and photometry generally produce integrals of very complex functions that are very difficult if not impossible to integrate analytically in closed form. When we evaluate these integrals numerically on a fast computer, we convert the integrals into their finite difference counterparts, as illustrated in (1.29) and (1.30), and then perform the calculations. The reader is referred to textbooks on integral calculus for further discussions of these concepts, which are central to the solution of complex problems in radiometry and photometry.

1.7 CONNECTION WITH ELECTROMAGNETIC THEORY

Many textbooks on optical physics and on electricity and magnetism speak of the flux of energy associated with electromagnetic radiation in terms of a quantity they call *intensity*. The intensity is related to the square of the amplitude of an oscillating field quantity, usually the electric field or the magnetic field. From the derivations provided, this ends up being a flux density in W/m^2 and as such is what is called *irradiance* in radiometry. Looking at how several of these texts [14–16] treat the concept of flux density in a propagating electromagnetic field, we see a commonality of terminology and usage.

Electromagnetic radiation is described as oscillations in electric and magnetic fields that propagate together with the same speed and direction. The usual starting place is with Maxwell's theory of electromagnetism as embodied in the set of differential equations that have been given his name. For more information, the reader is referred to Chapter 1 of the text by Boyd [17]. It is stated that the quantity

$$\mathbf{S} = \mathbf{E} \times \mathbf{H}, \quad (1.31)$$

being the vector product (or cross product) of the electric field vector \mathbf{E} with the magnetic strength vector \mathbf{H} and called the Poynting vector, gives the instantaneous rate at which electromagnetic energy passes through a unit area whose normal is in the direction of \mathbf{S} . For infinitely extended plane waves and plane polarized waves of angular frequency $\omega = 2\pi\nu$, where ν is the frequency, Poynting's vector exhibits rapid oscillations at frequency ν . The time average of this over one complete period of the oscillation gives the rate at which energy is transported by the propagating oscillations in the fields.

The term plane polarization refers to the fact that the oscillations in the electric field vector \mathbf{E} for this solution to Maxwell's equations lie within a plane parallel to the direction of propagation. Other modes of oscillation are possible, including an elliptical rotation of this vector around the direction of propagation or, at the limit, a circular one, in which case the radiation is termed to be circularly polarized. Polarization is discussed at greater length in the next section.

The time average of Poynting's vector for plane polarized infinite waves is

$$\langle \mathbf{S} \rangle = \frac{1}{2} \sqrt{\frac{\epsilon}{\mu}} E_o^2 \mathbf{k} \quad (1.32)$$

where \mathbf{k} is a unit vector in the direction of propagation parallel to S , ϵ is the dielectric permittivity, and μ is the magnetic permeability, properties of the medium in which the wave propagates, and E_o is the amplitude of the electric field oscillations. In a vacuum, $\epsilon = \epsilon_o = 8.85 \times 10^{-12}$ farad/m and $\mu = \mu_o = 4\pi \times 10^{-7}$ henry/m. It can be shown that

$$\lambda\nu = \frac{1}{\sqrt{\epsilon\mu}} \quad (1.33)$$

and that

$$c = \frac{1}{\sqrt{\epsilon\mu}} \quad (1.34)$$

is the speed of propagation of the phase of the oscillations, commonly known as the speed of light. The refractive index of the medium is defined to be the ratio of the speed $c_o = 2.99792 \times 10^8$ m/s of propagation in a vacuum to the speed c of propagation in the medium so that the refractive index is given by

$$n = \frac{c_o}{c} = \sqrt{\frac{\epsilon\mu}{\epsilon_o\mu_o}} \quad (1.35)$$

and from (1.33) and (1.34) we have

$$c = \lambda\nu \quad (1.36)$$

as given in (1.1) above. The quantity $\langle S \rangle$ in (1.32) is the irradiance and it has units of W/m². Thus, we finally see that irradiance is proportional to the time average of the amplitude of the oscillations of the electric field vector in propagating electromagnetic waves.

1.8 POLARIZATION

The concept of polarization of electromagnetic radiation was introduced in the previous section. It is important in radiometry and photometry because the optical properties of materials can be different for different states of polarization of the radiation incident upon them. This will be discussed more in Chapter 6 on the optical properties of materials, but some of the formalism for dealing with polarization in radiometry is introduced here.

Lasers are made that produce light which is plane polarized—the oscillations of the electric field strength vector E are confined to a plane containing the direction of propagation of the laser beam. If one thinks of a point along this beam and examines the direction and magnitude of E at this point over time, it is found that the vector changes its length but not its direction, except when its length shrinks to zero and it reverses direction, growing in the opposite one along the same line. The vector is therefore confined to a line perpendicular to the laser beam and having one specific direction that does not vary. This light can be said to be *linearly polarized* because the electric vector at a point along the beam oscillates back and forth along a line. Choosing some direction perpendicular to the laser beam, the line of polarization can be parallel to this direction or make some angle α to it, still being perpendicular to the laser beam.

We can also think of the linearly polarized laser beam as consisting of a train of electric vectors, one for each point along the beam at an instant of time. Since all of these vectors are parallel to each other, perpendicular to the direction of propagation and extended along this direction and therefore lying in a plane containing the direction of propagation, the laser beam can be said to be *plane polarized*. This state of polarization is depicted in Figure 1.9. The term *plane polarization* can be ambiguous, as pointed out by Shurcliff [18], and probably should be avoided.

Other regular patterns of oscillation of the electric vector are possible. For example, instead of oscillating back and forth along a line at a point along the propagation direction, the electric vector can rotate around the direction of propagation. If the length of this vector remains constant, then its tip describes a circle around the point of observation and the radiation is said to be *circularly polarized*. Since the beam is propagating at the speed of light while the electric vector is rotating about the direction of propagation, an instantaneous snapshot of the tip of the electric vector would reveal a helical curve as depicted in Figure 1.10. The direction of rotation can be clockwise (*right circular polarization*) or counterclockwise (*left circular polarization*), viewing along the direction of propagation.

Polarization states intermediate between circular and linear polarization are possible and are called *elliptically polarized* radiation because this is the figure drawn by the tip of the electric vector at a point along the beam. If the minor axis of an ellipse has the same length as the major axis, then the ellipse becomes a circle. If the length of the minor axis is taken to zero, then the ellipse shrinks into a line. Because of this, linear and circular polarization states are limiting cases of elliptical polarization. To completely describe polarized light, one needs to specify the angle α of rotation of the major axis of the ellipse and its ellipticity β , which is the ratio of the lengths of the minor to the major axis of the ellipse. Ellipticity is zero for linear polarization and 1 for circular polarization.

Unpolarized light is light whose electric field strength vector oscillates in random directions. It thus exhibits no long-term orderliness in terms of the direction

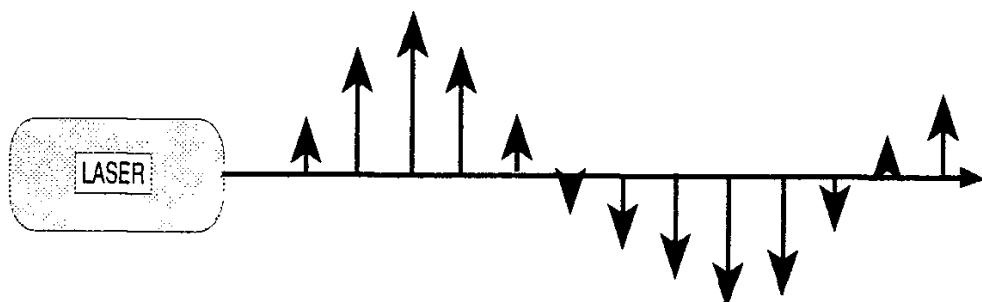


Figure 1.9 Depiction of the electric field strength vectors along the direction of propagation of a linearly polarized beam at an instant of time.

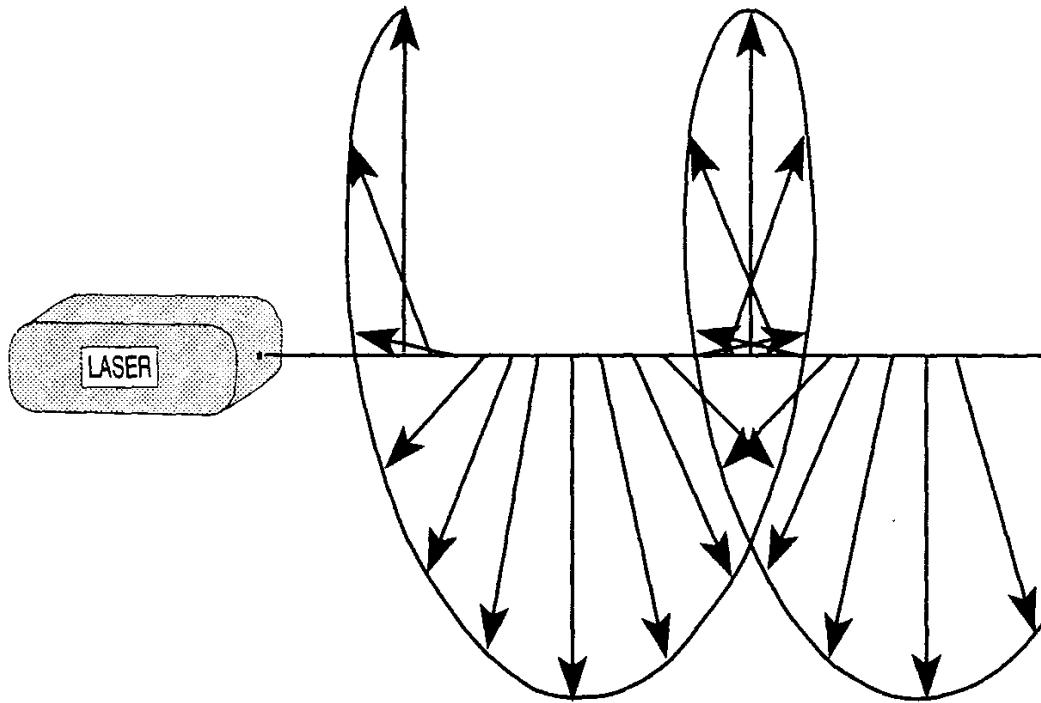


Figure 1.10 Depiction of the electric field strength vectors along the direction of propagation of a circularly polarized beam at an instant of time.

or motion of its electric field vector. Shurcliff [18] defines an unpolarized beam as “a beam that, when operated on by any elementary kind of energy-conserving device that divides the beam into two completely polarized subbeams, yields subbeams that have *equal power* (in a time interval long enough to permit the powers to be measured). Thus if a beam is to qualify as unpolarized, it must exhibit no long-term preferences as to lateral direction of vibration or as to handedness.”

Except in the case of single-wavelength laser light and a few other specialty sources, most visible light has an appreciable range of wavelengths and can contain a large number of forms of polarization simultaneously. If one cannot detect any preferred azimuth or circularity, then one must conclude that the light is unpolarized.

This is different from the term *randomly polarized* in that the latter refers to light that is polarized but the state of polarization at any one time is randomly selected from all possible states of elliptical polarization, including the limiting cases of circular and linear polarization. Of course, with randomly polarized light, if the time interval over which the polarization state is constant shrinks to a period shorter than we can make measurements of, then the result is really just unpolarized light, light whose polarization state changes rapidly and randomly during the measurement period.

A beam of light can contain both polarized and unpolarized components and a number of techniques are available for determining (*analyzing*) the polarization states contained in a beam of radiation. There are several schemes for describing these states, including both polarized and unpolarized components. The *degree of polarization* is an important concept that helps us understand the interaction of a partially polarized beam with optical systems which may respond differently to different states of polarization. For further information on this subject the reader is referred to the excellent text by Shurcliff [18] on polarization, or to any of a number of modern textbooks on optics.

Polarization is of importance to the discussion in Chapter 6 on the optical properties of materials. The reason is that many materials react differently to different forms of polarization incident upon them. The reflectance of most materials, for example, is different for what are called the parallel and perpendicular polarization components of light reflected from them (except at a zero angle of incidence). Parallel in this case means parallel to the plane of incidence, a plane containing the direction of incidence (and reflection) and a perpendicular (called a *normal*) to the surface.

Light from the blue sky, for example, is partially polarized. In studying the radiometry and photometry of this light, one must be aware that measuring instruments may react differently to the polarized component than to the unpolarized component.

1.9 PHOTON FLUX

At the beginning of the 20th century, Albert Einstein, Max Planck, and others developed a quantum theory of electromagnetic radiation, in which this radiation is thought of not as continuous waves moving through space but as discrete packets or quanta of energy propagating with the same speed and direction. Each quantum of radiation has a specific frequency ν associated with it and the energy e_ν contained in a single quantum of this radiation, called a *photon*, is given by

$$e_\nu = h\nu \quad (1.37)$$

where $h = 6.63 \times 10^{-34} \text{ J}\cdot\text{s}$ is Planck's constant. Since $c = \lambda\nu$, we can write this equation explicitly in terms of the wavelength as follows:

$$e_\lambda = \frac{hc}{\lambda} \quad (1.38)$$

This equation gives the correspondence between the wave nature of light and its

particle nature. These two natures are called the wave-particle duality of electromagnetic radiation. In many applications of radiometry, photon flux quantities are of more interest than radiant flux, radiant intensity, irradiance, or radiance, which are connected to the wave nature of radiation as described in Section 1.6. For such applications, we may replace the latter quantities with photon flux, photon intensity, photon irradiance, and photon radiance, respectively. Definitions for these photon quantities as well as their relationships to the radiant quantities defined above are presented next.

Photon number, N_p , is the number of photons emitted by a source or propagating onto, through, or emerging from, a specified surface of given area in a given period of time. Photons of all wavelengths emitted or contained in the beam are to be included. If a limited wavelength range is to be considered, this range must be stated (unit: 1). The photon number N_p is determined from the spectral radiant energy Q_λ , defined above, as follows. The element of photon number dN_p in the element of frequency interval $d\nu$ at frequency ν is given by

$$dN_p = \frac{Q_\nu}{h\nu} d\nu = \frac{\lambda Q_\lambda}{hc} d\lambda \quad (1.39)$$

Integrating this expression over all frequencies and wavelengths contained in the beam yields

$$N_p = \int dN_p = \int \frac{Q_\nu}{h\nu} d\nu = \int \frac{\lambda Q_\lambda}{hc} d\lambda \quad (1.40)$$

Photon flux, Φ_p , is the time rate of flow of photons (units: s^{-1}). The defining equation is

$$\Phi_p = \frac{dN_p}{dt} \quad (1.41)$$

Whenever radiation is made to fall on a device that produces an electronic current, voltage, or other signal that is proportional to the number of photons per second incident on the device (called a *photon counter*), it is generally the total number of photons over the whole device that is important rather than the photon flux per unit area. The relationship between photon flux and radiant flux is as follows:

$$\Phi_p = \frac{1}{hc} \int \lambda \Phi_\lambda d\lambda \quad (1.42)$$

Once we have defined photon flux, the definitions of photon irradiance, photon

intensity, and photon radiance are essentially the same as the previous definitions of irradiance, intensity, and radiance, respectively, as shown in the following paragraphs.

Photon irradiance, E_p , is the area density of photon flux, the photon flux per unit area in a specified surface that is incident on, passing through, or emerging from a point in the specified surface. All directions in the hemispherical solid angle above or below the point in the surface are to be included (see Fig. 1.4) (units: m^{-2}). The defining equation is

$$E_p = \frac{d\Phi_p}{ds_o} \quad (1.43)$$

where $d\Phi_p$ is an element of photon flux and ds_o is an element of area in the surface.

Photon irradiance is a function of position on the specified surface. The photon irradiance leaving a surface can be called photon exitance, M_p , and it has the same units and defining equation. If E_λ is the spectral irradiance, then the photon irradiance will be given by

$$E_p = \frac{1}{hc} \int \lambda E_\lambda d\lambda \quad (1.44)$$

Photon intensity, I_p , is the solid angle density of photon flux, the photon flux per unit solid angle incident on, passing through, or emerging from a point in space and propagating in a specified direction (units: sr^{-1}). The defining equation is

$$I_p = \frac{d\Phi_p}{d\omega} \quad (1.45)$$

where $d\Phi_p$ is the element of photon flux incident on or emerging from a point within element $d\omega$ of solid angle in the specified direction.

Photon radiance, L_p , is the area and solid angle density of photon flux, the photon flux per unit projected area and per unit solid angle incident on, passing through, or emerging in a specified direction from a specified point in a specified surface (units: $\text{m}^{-2} \cdot \text{sr}^{-1}$). The defining equation is

$$L_p = \frac{d^2\Phi_p}{d\omega ds} = \frac{d^2\Phi_p}{d\omega ds_o \cos \theta} \quad (1.46)$$

where $ds = ds_o \cos \theta$ is the projected area, the area of the projection of elemental area ds_o in the surface containing the point at which the photon radiance is being

defined, projected in the direction of propagation onto a plane perpendicular to this direction. $d\omega$ is an element of solid angle in the specified direction and θ is the angle between this direction and the normal to the surface at the specified point.

EXAMPLE PROBLEM 1.1

Problem: Given a circle of radius r whose center is a distance f from a point P lying on the perpendicular to the plane of the circle and passing through the center of the circle, as illustrated in Figure 1.11. Determine the relationship between the solid angle Ω subtended by the circle at point P and the half angle α .

Solution: First we must either project the circle onto a unit sphere or onto a sphere of known radius. We choose the latter, selecting a sphere of radius $f + \Delta f$ so that the circle lies on a sphere of this radius, which we call $R = f + \Delta f$, so that, from Figure 1.12, we have

$$\Delta f = R - f = \frac{f}{\cos \alpha} - f \quad (1.47)$$

The solid angle subtended by the circle at P will be the quotient of the area A of the portion of the sphere enclosed by the circle, divided by the square of the radius R of the sphere. To determine area A on the sphere, we construct an infinitesimal

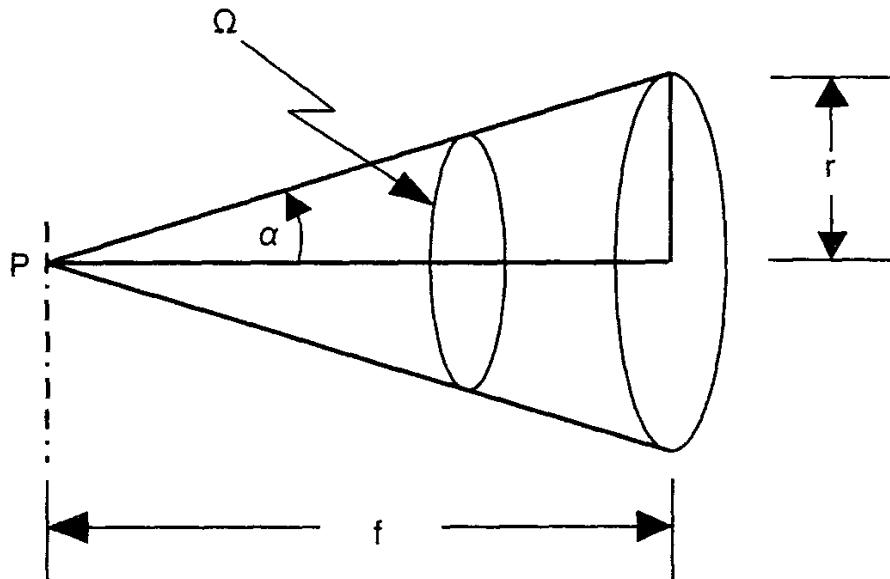


Figure 1.11 Geometry for Example Problem 1.1.

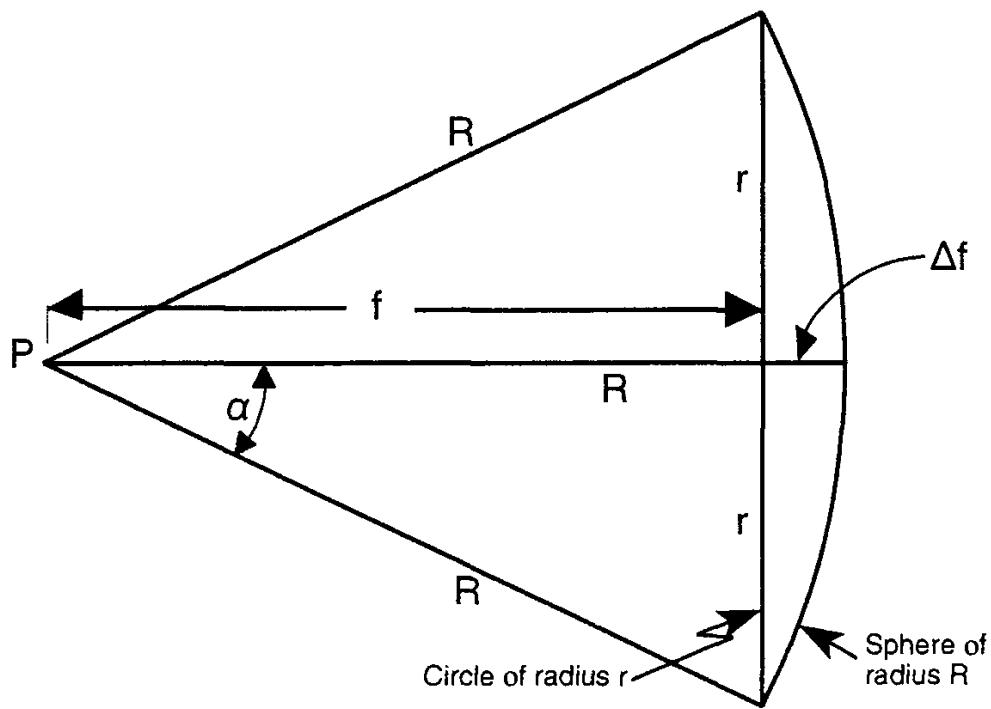


Figure 1.12 Geometry for determining Δf .

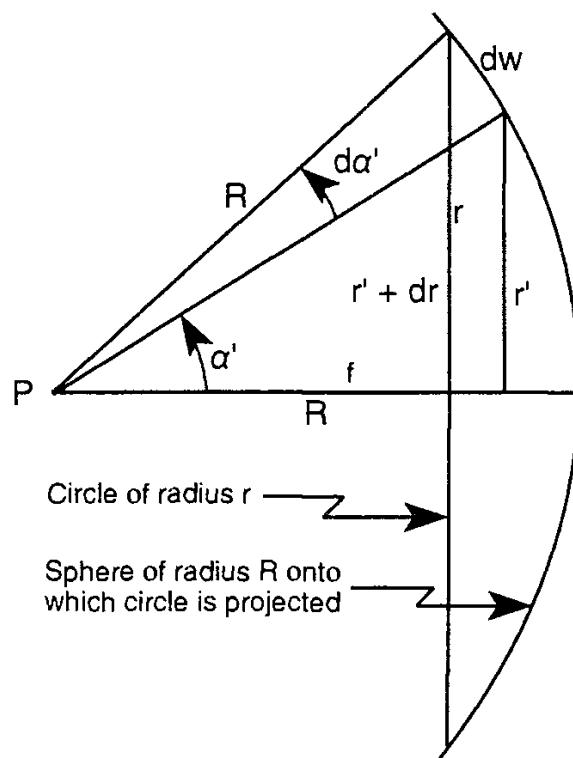


Figure 1.13 Geometry for determining dw .

area da on the sphere and integrate it over the interior of the circle on the sphere, as follows.

The element da of area forms an annular ring of inner radius r' , outer radius $r' + dr'$ and width dw , as illustrated in Fig. 1.13, with $r' \leq r$. We may write da as its length $2\pi r'$ (the circumference of the circle of radius r') times its width dw , which is $R d\alpha'$

$$da = 2\pi r' R d\alpha' \quad (1.48)$$

but

$$r' = R \sin \alpha' \quad (1.49)$$

so that

$$da = 2\pi R^2 \sin \alpha' d\alpha' \quad (1.50)$$

Integrating (1.50) gives

$$\begin{aligned} A &= \int_0^\alpha 2\pi R^2 \sin \alpha' d\alpha' \\ A &= 2\pi R^2 [1 - \cos \alpha] \end{aligned} \quad (1.51)$$

From the definition of solid angle, we see that the solid angle Ω subtended by the circle is the area A of its projection onto a sphere of radius R divided by the square of the radius R

$$\Omega = \frac{A}{R^2} = 2\pi [1 - \cos \alpha] \quad (1.52)$$

which is the desired result.

EXAMPLE PROBLEM 1.2

Problem: The radiant flux received by detector D of area A_d at the bottom of a black tube of length X and aperture area A_a , as shown in Figure 1.14, is Φ_d . Assuming that A_d is much less than (\ll) X^2 , $A_a \ll X^2$, and $A_a \ll R^2$, what is the average radiance of the area of the wall toward which the tube is pointed?

Solution: Dividing the flux Φ_d by the area A_d of the detector and the solid angle of the entrance aperture area A_a at the center of the detector, we have the average radiance in the aperture, the average being taken over the solid angle

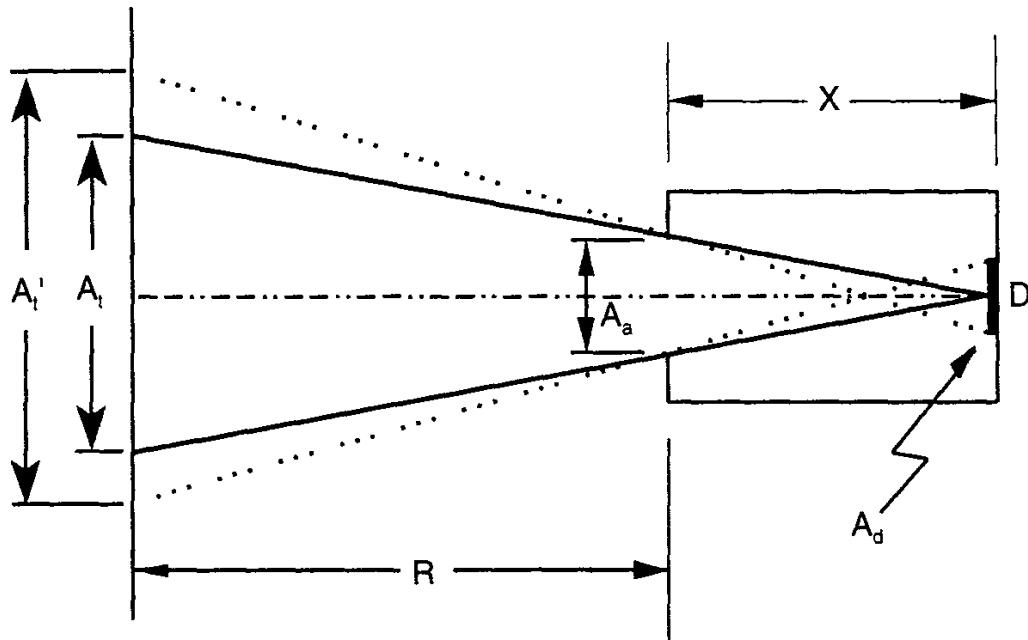


Figure 1.14 Geometry for Example Problem 1.2.

subtended by the center of the detector, called the field of view from the center of the detector.

$$L = \frac{\Phi_d}{A_d \frac{A_a}{X^2}} = \frac{\Phi_d X^2}{A_d A_a} \quad (1.53)$$

(In using A_a/X^2 for the solid angle of the aperture at the center of the detector, we have made use of the fact that the area on a sphere of radius X bounded by the circular aperture equals the plane area A_a of the aperture in the small-angle approximation that we are using here.) Equation (1.53) is also the apparent radiance of the portion of the wall in the field of view of the device. This tube and detector therefore can be used as a radiance meter. The equivalence of wall and aperture radiances is explained in Chapter 5.

Note that the area of the field of view from the point at the center of the detector, called the *target* area, is A_t , as shown in the illustration in Figure 1.14. However, around the edge of the detector, the field of view is shifted to the side somewhat. Thus, the total area “seen” by the whole detector, A'_t , is larger than that seen from the center of it. For the purposes of talking about the area of the field of view, for calibration purposes we use area A_t , since this is the magnitude of the area seen by each point in the detector if the detector is small. For the purposes of designating what was actually “seen” by the meter, we use A'_t .

REFERENCES

- [1] CIE, International Lighting Vocabulary, CIE Publ No. 17.4 IEC Publ. 50(845), Central Bureau of the *Commission Internationale de l'Eclairage*, Kegelgasse 27, A-1030 Vienna, P.O. Box 169, Austria, and *Bureau Central de la Commission Eletrotechnique Internationale*, 2, rue de Varembe, Geneva, Switzerland, 1987. Also available from the national committees of the CIE. In the U.S., this document can be obtained from TLA-Lighting Consultants, Inc., 7 Pond Street, Salem, MA 01970-4893.
- [2] Illuminating Engineering Society of North America, "USA Standard Nomenclature and Definitions for Illuminating Engineering," RP-16, IESNA, 345 E. 47 St., New York, NY, 10017, 1986.
- [3] Kaufman, J. E., ed. *IES Lighting Handbook—Reference Volume*, IESNA: New York, NY, 1984.
- [4] Wyszecki, G., and W. S. Stiles. *Color Science: Concepts and Methods, Quantitative Data and Formulae*, 2nd ed., John Wiley & Sons: New York, NY, 1982, p. 253.
- [5] Olivier, P., and D. Gagnon. "Mathematical Modeling of the Solid Angle Function, Part I: Approximation in Homogeneous Medium," *Optical Engineering*, Vol. 32, No. 9, Sept. 1993, pp. 2261–2265.
- [6] Olivier, P., and D. Gagnon. "Mathematical Modeling of the Solid Angle Function, Part II: Transmission Through Refractive Media," *Optical Engineering*, Vol. 32, No. 9, Sept. 1993, pp. 2266–2270.
- [7] Naito, M. "A Method of Calculating the Solid Angle Subtended by a Circular Aperture," *J. Phys. Soc. Japan*, Vol. 12, No. 10, 1957, pp. 1122–1129.
- [8] Gardner, R. P., and A. Carnesale. "The Solid Angle Subtended at a Point by a Circular Disk," *Nucl. Inst. Meth.*, Vol. 73, 1957, 228–230.
- [9] Garrett, M. W. "Solid Angle Subtended by a Circular Aperture," *Rev. Sci. Instr.*, Vol. 25, No. 12, 1954, pp. 1208–1211.
- [10] Jaffey, A. H. "Solid Angle Subtended by a Circular Aperture at Point and Spread Sources: Formulas and Some Tables," *Rev. Sci. Instr.*, Vol. 25, No. 4, 1954, pp. 349–354.
- [11] Khakjavi, A. "Calculation of Solid Angle Subtended by Rectangular Apertures," *J. Opt. Soc. Am.*, Vol. 58, No. 10, 1968, pp. 1417–1418.
- [12] Masket, A. V. "Solid Angle Contour Integrals, Series, and Tables," *Rev. Sci. Instr.*, Vol. 28, No. 3, 1957, pp. 191–197.
- [13] Masket, A. V. "Solid Angle Subtended by a Circular Disk," *Rev. Sci. Instr.*, Vol. 30, No. 10, 1959, p. 950.
- [14] Jackson, *Classical Electrodynamics*, John Wiley & Sons: New York, NY, 1962.
- [15] Jenkins and White. *Fundamentals of Optics*, 3rd ed., McGraw-Hill: New York, NY, 1957.
- [16] Born, M., and E. Wolf. *Principles of Optics*, 3rd ed., Pergamon Press: New York, NY, 1965.
- [17] Boyd, R. W. *Radiometry and the Detection of Optical Radiation*, John Wiley & Sons: New York, NY, 1983.
- [18] Shurcliff, W. A. *Polarized Light—Production and Use*, Harvard University Press: Cambridge, MA, 1966.

Chapter 2

Fundamental Concepts of Photometry

2.1 LIGHT

Photometry can be thought of as a subset or branch of radiometry, where all the radiant quantities defined in the previous chapter have been modified to indicate the human eye's response to them. Being familiar with the basic radiometric concepts introduced in that chapter makes much easier the study of the corresponding photometric concepts introduced in this chapter.

As was shown in the previous chapter, the human eye responds only to light having wavelengths between about 360 and 800 nm. Radiometry deals with electromagnetic radiation at all wavelengths and frequencies, while photometry deals only with visible light—that portion of the electromagnetic spectrum which stimulates vision in the human eye. Radiation having wavelengths below 360 nm, down to about 100 nm, is called *ultraviolet*, or UV, meaning “beyond the violet.” Radiation having wavelengths greater than 830 nm, up to about 1 mm, is called *infrared*, or IR meaning “below the red.” The “below” in this case refers to the frequency of the radiation, not to its wavelength. Solving (1.1) for frequency yields the equation $\nu = c/\lambda$, showing the inverse relationship between frequency and wavelength. The infrared portion of the spectrum lies beyond the red, having frequencies below and wavelengths above those of red light. Since the eye is very insensitive to light at wavelengths between 360 and about 410 nm and between about 720 and 830 nm, and many people cannot see at all radiation in portions of these ranges, the visible edges of the UV and IR spectra are as uncertain as the edges of the visible spectrum.

The term “light” should *only* be applied to electromagnetic radiation in the visible portion of the spectrum, lying between 360 and 830 nm. With this terminology, there is no such thing as “ultraviolet light,” nor does the term “infrared light” make any sense either. Radiation outside these limits is radiation—not light—and should not be referred to as light.

To begin our study of photometry, we first look at how the human eye responds to different wavelengths of light. After passing through the cornea, the

aqueous humor, the iris and lens, and the vitreous humor, light entering the eye is received by the retina, which contains two general classes of receptors: rods and cones. Photopigments in the outer segments of the rods and cones absorb radiation and the absorbed energy is converted within the receptors, into neural electrochemical signals which are then transmitted to subsequent neurons, the optic nerve, and on to the brain.

The cones are primarily responsible for day vision and the seeing of color. Cone vision is called *photopic vision*. The rods come into play mostly for night vision, when illumination levels entering the eye are very low. Rod vision is called *scotopic vision*. An individual's relative sensitivity to various wavelengths is strongly influenced by the absorption spectra of the photoreceptors, combined with the spectral transmittance of the preretinal optics of the eye. The relative spectral sensitivity depends on light level and this sensitivity shifts toward the blue (shorter wavelength) portion of the spectrum as light levels drop due to the shift in spectral sensitivity when going from cones to rods.

The spectral response of a human observer under photopic (cone vision) conditions was standardized by the International Lighting Commission, the *Internationale de l'Eclairage* (CIE), in 1924 [1]. Although the actual spectral response of humans varies somewhat from person to person, an agreed standard response curve has been adopted, as shown in Figure 2.1 and listed numerically in Table 2.1. The values in Table 2.1 are taken from the *IES Lighting Handbook* of the

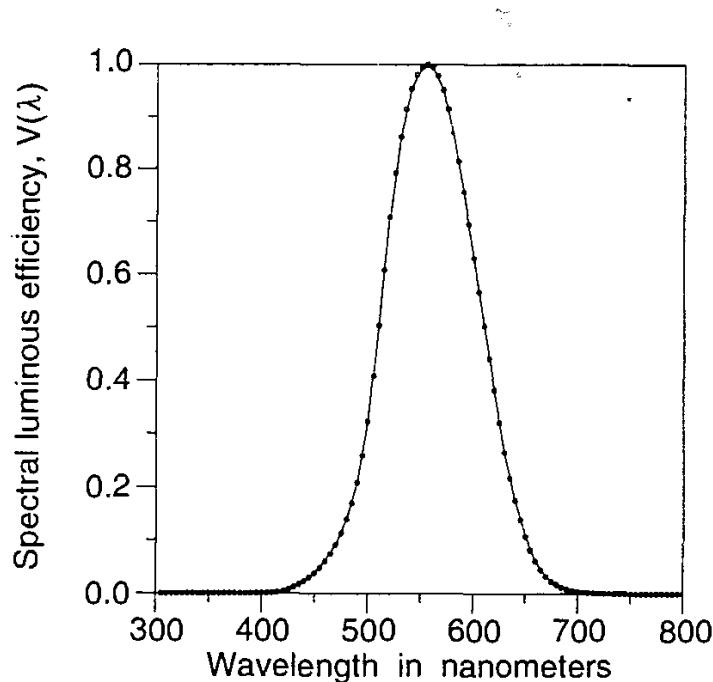


Figure 2.1 Photopic spectral luminous efficiency function, $V(\lambda)$, for CIE 1924 standard photometric observer.

Table 2.1
Photopic Spectral Luminous Efficiency $V(\lambda)$

<i>Wavelength Standard</i>	<i>λ, nm</i>	<i>Values</i>	<i>Values Interpolated at Intervals of 1 Nanometer</i>								
			1	2	3	4	5	6	7	8	9
	380	.00004	.000045	.000049	.000054	.000058	.000064	.000071	.000080	.000090	.000104
	390	.00012	.000138	.000155	.000173	.000193	.000215	.000241	.000272	.000308	.000350
	400	.0004	.00045	.00049	.00054	.00059	.00064	.00071	.00080	.00090	.00104
	410	.0012	.00138	.00156	.00174	.00195	.00218	.00244	.00274	.00310	.00352
	420	.0040	.00455	.00515	.00581	.00651	.00726	.00806	.00889	.00976	.01066
	430	.0116	.01257	.01358	.01463	.01571	.01684	.01800	.01920	.02043	.02170
	440	.023	.0243	.0257	.0270	.0284	.0298	.0313	.0329	.0345	.0362
	450	.038	.0399	.0418	.0438	.0459	.0480	.0502	.0525	.0549	.0574
	460	.060	.0627	.0654	.0681	.0709	.0739	.0769	.0802	.0836	.0872
	470	.091	.0950	.0992	.1035	.1080	.1126	.1175	.1225	.1278	.1333
	480	.139	.1448	.1507	.1567	.1629	.1693	.1761	.1833	.1909	.1991
	490	.208	.2173	.2270	.2371	.2476	.2586	.2701	.2823	.2951	.3087
	500	.323	.3382	.3544	.3714	.3890	.4073	.4259	.4450	.4642	.4836
	510	.503	.5229	.5436	.5648	.5865	.6082	.6299	.6511	.6717	.6914
	520	.710	.7277	.7449	.7615	.7776	.7932	.8082	.8225	.8363	.8495
	530	.862	.8739	.8851	.8956	.9056	.9149	.9238	.9320	.9398	.9471
	540	.954	.9604	.9961	.9713	.9760	.9083	.9480	.9873	.9902	.9928
	550	.995	.9969	.9983	.9994	1.0000	1.0002	1.0001	.9995	.9984	.9969
	560	.995	.9926	.9898	.9865	.9828	.9786	.9741	.9691	.9638	.9581
	570	.952	.9455	.9386	.9312	.9235	.9154	.9069	.8981	.8890	.8796
	580	.870	.8600	.8496	.8388	.8277	.8163	.8046	.7928	.7809	.7690
	590	.757	.7449	.7327	.7202	.7076	.6949	.6822	.6694	.6565	.6437
	600	.631	.6182	.6054	.5926	.5797	.5668	.5539	.5410	.5282	.5156
	610	.503	.4905	.4781	.4568	.4535	.4412	.4291	.4170	.4049	.3929
	620	.381	.3690	.3575	.3449	.3329	.3210	.3092	.2977	.2864	.2755
	630	.265	.2548	.2450	.2354	.2261	.2170	.2082	.1996	.1912	.1830
	640	.175	.1672	.1596	.1523	.1452	.1382	.1316	.1251	.1188	.1128
	650	.107	.1014	.0961	.0910	.0862	.0816	.0771	.0729	.0688	.0648
	660	.061	.0574	.0539	.0506	.0475	.0446	.0418	.0391	.0366	.0343
	670	.032	.0299	.0280	.0263	.0247	.0232	.0219	.0206	.0194	.0182
	680	.017	.01585	.01477	.01376	.01281	.01192	.01108	.01030	.00956	.00886
	690	.0082	.00759	.00705	.00656	.00612	.00572	.00536	.00503	.00471	.00440
	700	.0041	.00381	.00355	.00332	.00310	.00291	.00273	.00256	.00241	.00225
	710	.0021	.001954	.001821	.001699	.001587	.001483	.001387	.001297	.001212	.001130
	720	.00105	.000975	.000907	.000845	.000788	.000736	.000668	.000644	.000601	.000560
	730	.00052	.000482	.000447	.000415	.000387	.000360	.000335	.000313	.000291	.000270
	740	.00025	.000231	.000214	.000198	.000185	.000172	.000160	.000149	.000139	.000130
	750	.00012	.000111	.000103	.000096	.000090	.000084	.000078	.000074	.000069	.000064
	760	.00006	.000056	.000052	.000048	.000045	.000042	.000039	.000037	.000035	.000032

Illuminating Engineering Society of North America (IES) [2]. Since the symbol $V(\lambda)$ is normally used to represent this spectral response, the curve in Figure 2.1 is often called the “V-lambda curve.”

The data shown in Table 2.1 give the *1924 CIE spectral luminous efficiency function for photopic vision*. These data define what is called [3] “the CIE 1924 Standard Photopic Photometric Observer.” The official values were originally given for the wavelength range from 380 to 780 nm at 10-nm intervals [4] but were then [3,5] “completed by interpolation, extrapolation, and smoothing from earlier values adopted by the CIE in 1924 and 1931” to the wavelength range from 360 to 830 nm on 1-nm intervals [6] and these were then recommended [7] by the International Committee of Weights and Measures (CIPM) in 1976 [5]. The values below 380 and above 769 are so small that they are of little value for most photometric calculations and are therefore not included in Table 2.1.

Any individual’s eye may depart somewhat from the response shown in Figure 2.1, and when light levels are moderately low, the other set of retinal receptors (rods) comes into use. This regime is called “scotopic vision” and is characterized by a different relative spectral response.

The relative spectral response curve for scotopic vision is similar in shape to the one shown in Figure 2.1, but the peak is shifted from about 555 nm to about 510 nm. The lower wavelength cutoff in sensitivity remains at about 380 nm, however, while the upper limit drops to about 640 nm. More information about scotopic vision can be found in various books on vision as well as in the *IES Lighting Handbook-Reference Volume* [2]. The latter contains both plotted and tabulated values for the scotopic spectral luminous efficiency function, $V'(\lambda)$.

There is a wide range at medium to low light levels, where *both* rod and cone receptors are involved. This is called *mesopic vision*. There is some recent evidence [8] that at the moderate light levels encountered inside typical office buildings, our vision is some combination of photopic and scotopic, but for the purposes of this discussion we are concerned only with illumination at moderately high light levels where the standard response curve is a reasonable approximation for photopic vision.

The colors of the visible spectrum all fall in the wavelength range from 360 to 830 nm where the $V(\lambda)$ function is nonzero. As described in Chapter 11, any given color perception can result from a variety of different spectral distributions over the visible range, a phenomenon known as *metamerism*. Monochromatic radiation in the visible portion of the spectrum also produces the sensation of color. Wavelengths stimulating the principal colors of the spectrum are given below [9]:

Purple: 360–450 nm
Yellow: 570–591 nm

Blue: 450–500 nm
Orange: 591–610 nm

Green: 500–570 nm
Red: 610–830 nm

2.2 PHOTOMETRIC DEFINITIONS

Four fundamental quantities in radiometry were defined in the previous chapter. The corresponding photometric ones are easily defined in terms of the radiometric ones as follows. Let Q_λ be one of the following: spectral radiant flux Φ_λ , spectral irradiance E_λ , spectral intensity I_λ , or spectral radiance L_λ . The corresponding photometric quantity, Q_v , is defined as follows:

$$Q_v = 683 \int_{380}^{770} Q_\lambda V(\lambda) d\lambda \quad (2.1)$$

with wavelength λ having units of nanometers.

The subscript v (standing for visible or visual) is placed on photometric quantities to distinguish them from radiometric quantities, which are given the subscript e (standing for energy). These subscripts may be dropped, as they were in Chapter 1, when the meaning is clear and no ambiguity results. The four fundamental radiometric quantities, and the corresponding photometric quantities are listed in Table 2.2 along with the units for each. The reason for the presence of the factor 683 in (2.1) is given in Section 2.2.1.

The beauty of representing the conversion from radiometry to photometry in one simple equation, (2.1), must be counterbalanced by the need which newcomers to photometry may have to understand the individual meanings and uses of the four basic photometric quantities. These are now discussed.

The basic unit of *luminous flux*, the lumen, is like a “light-watt.” It is the luminous equivalent of the radiant flux or power. Similarly, *luminous intensity* is the photometric equivalent of radiant intensity. It gives the luminous flux in lumens emanating from a point, per unit solid angle in a specified direction, and therefore has the units of lumens per steradian or lm/sr. This unit is one of the seven base

Table 2.2
Basic Quantities of Radiometry and Photometry

Radiometric Quantity	Symbol	Units	Photometric Quantity	Symbol	Units
Radiant flux	Φ_e	watt	Luminous flux	Φ_v	lumen (lm)
Radiant intensity	I_e	W/sr	Luminous intensity	I_v	lumen/sr = candela, abbrev. cd
Irradiance	E_e	W/m ²	Illuminance	E_v	lumen/m ² = lux, abbrev. lx
Radiance	L_e	W·m ⁻² ·sr ⁻¹	Luminance	L_v	lm·m ⁻² ·sr ⁻¹ = cd/m ²

units of the metric system and is also called the *candela*. More information about the metric system is provided in Chapter 10. Luminous intensity is a function of direction from its point of specification, and may be written as $I_v(\theta, \phi)$ to indicate its dependence upon the spherical coordinates (θ, ϕ) specifying a direction in space.

Illuminance is the photometric equivalent of irradiance and is like a “light-watt per unit area.” Illuminance is a function of position (x, y) in the surface on which it is defined and may therefore be written as $E_v(x, y)$. Most light meters measure illuminance and are calibrated to read in lux. The lux is an equivalent term for the lumen per square meter. In the widely used inch-pound (I-P) system of units, the unit for illuminance is the lumen/ ft^2 , which also has the odd name “foot-candle,” abbreviated fc, even though it has nothing to do with candles or the candela. This terminology is slowly being discontinued in favor of the metric system that is used throughout this book. For more information on the connections between modern metric units and terminology in photometry and the antiquated units from other systems found in older writings in the field, the reader is directed to Chapter 10. As with irradiance and radiant exitance, illuminance leaving a surface can be called *luminous exitance*.

The inch-pound system has been the predominant system in use for engineering work in the U.S. for some time. Consequently, most U.S. electrical and lighting engineers work with the foot candle as the unit of illuminance. The foot-candle is a misnomer, being defined to be one lumen of light flux per square foot of area. Since there are 10.76 ft^2 in 1 m^2 , there are 10.76 lux in a fc. Rather than round this number up to 11, most people round it down (truncate it) to 10, making it quite easy to convert between lux and foot-candles. Most modern light meters have metric calibration options. There is a move toward metric standards within the illumination engineering community, and the archaic and frequently confusing nonmetric units are slowly dying out. Chapter 10 is provided as a bridge to the nonmetric literature of the past.

Luminance can be thought of as “photometric brightness,” meaning that it comes relatively close to describing physically the subjective perception of “brightness.” This is discussed more in Section 2.3. Luminance is the quantity of light flux passing through a point in a specified surface in a specified direction, per unit projected area at the point in the surface and per unit solid angle in the given direction. The units for luminance are therefore $\text{lm} \cdot \text{m}^{-2} \cdot \text{sr}^{-1}$. A more common unit for luminance is the cd/m^2 , which is the same as the lumen per steradian and per square meter. See Section 1.4 for a discussion of projected area.

Other units have been used for photometric quantities, and many of them are described in Chapter 10. Except for the mention of the foot-candle above, in order to avoid confusion, further discussion of nonmetric and deprecated units is confined to that chapter.

Equation (2.1) shows that each photometric quantity has a radiometric counterpart. The names of photometric quantities are distinguished from their radiometric cousins primarily by having the adjective “luminous” precede them,

with the exception of the photometric counterparts of radiance (called *luminance*) and irradiance (called *illuminance*). All the spectrally integrated quantities, such as Φ , E , I , and L in Chapter 1, have their photometric equivalents. Thus, the equations in Chapter 1 not involving spectral quantities also have their photometric equivalents. Specifically, Equations (1.7), (1.9), (1.11), (1.13), (1.15–1.19), (1.21–1.28) and (1.53) can be converted to photometric equations by the simple expedient of adding the subscript v to the radiometric quantities contained in them. For example, the important radiometric equation (1.17), relating irradiance to radiance and repeated here as

$$E_e = \int_{2\pi} L_e(\theta, \phi) \cos \theta d\omega \quad [W \cdot m^{-2}] \quad (2.2)$$

becomes the following equation relating illuminance to luminance:

$$E_v = \int_{2\pi} L_v(\theta, \phi) \cos \theta d\omega \quad [lm \cdot m^{-2}] \quad (2.3)$$

As was the case with radiant intensity, when speaking of luminous intensity, it is important to specify both the point in space from which the luminous flux is emanating and the direction of evaluation from that point. When talking about illuminance, it is important to specify both the surface and the point in it where the illuminance is to be evaluated. Luminance discussions likewise should make it clear in what surface, where in that surface, and in which direction relative to the point in the surface the luminance is being evaluated.

As one becomes more familiar with the concepts of radiance and luminance, it may become burdensome to state explicitly the point, surface, and direction every time the luminance is discussed. This information is still important, however, and must be implied in the context of the discussion.

Some representative values of three of the four fundamental photometric quantities are listed in Table 2.3. To aid the reader, the individual defining equations for the four basic photometric quantities are presented separately below. Each of these is one specific embodiment of (2.1).

$$\Phi_v = 683 \int_{380}^{770} \Phi_\lambda V(\lambda) d\lambda \quad [lm] \quad (2.4)$$

$$I_v = 683 \int_{380}^{770} I_\lambda V(\lambda) d\lambda \quad [lm \cdot sr^{-1}] \quad (2.5)$$

$$E_v = 683 \int_{380}^{770} E_\lambda V(\lambda) d\lambda \quad [lm \cdot m^{-2}] \quad (2.6)$$

Table 2.3
Representative Values of Photometric Quantities

Quantity	Value
Total luminous flux from a 100W tungsten incandescent light bulb, gas-filled, coiled coil filament, outputting 82W total radiant flux, radiation luminous efficacy, $K_r = 21.2 \text{ lm/W}$, lighting system luminous efficacy, $K_s = 17.4 \text{ lm/W}$	1,740 lumens
Luminous flux output of typical, medium power, helium-neon laser, emission at 632.8 nm, 5.0 mW radiant flux, $K_r = 159 \text{ lm/W}$	796 lumens
Radiant flux output from a 40W cold cathode fluorescent lamp, 23.2W radiant flux, $K_r = 122 \text{ lm/W}$, $K_s = 71 \text{ lm/W}$	2,830 lumens
Airport beacon lamp, 1,200W	27,500 lumens
Extraterrestrial solar illuminance at mean earth orbit, solar irradiance = 1,367 W/m ² , $K_r = 99.3 \text{ lm/W}$	133 klx
Terrestrial direct normal solar illuminance, clear sky, southeastern U.S., winter. Irradiance = 852 W/m ² , $K_r = 111 \text{ lm/W}$	94.6 klx
Terrestrial global (hemispherical) solar illuminance, on horizontal plane. Irradiance = 686 W/m ² , $K_r = 115 \text{ lm/W}$	78.9 klx
Average solar luminance, at its surface. Radiance = $2.3 \times 10^7 \text{ W/m}^{-2}\text{sr}^{-1}$	$2.3 \times 10^9 \text{ cd/m}^2$
Average solar luminance, apparent, viewed from earth surface. Radiance = $1.4 \times 10^7 \text{ W}\cdot\text{m}^{-2}\cdot\text{sr}^{-1}$	$1.6 \times 10^9 \text{ cd/m}^2$
Lunar luminance, from earth surface, at bright spot on the moon	$2.5 \times 10^3 \text{ cd/m}^2$
Clear sky luminance, typical	8,000 cd/m ²
Overcast sky luminance, typical	2,000 cd/m ²
Candle flame luminance	10,000 cd/m ²
Blackbody luminance at 6,500K	$3 \times 10^9 \text{ cd/m}^2$
Tungsten filament luminance inside 100W light bulb	$1.2 \times 10^7 \text{ cd/m}^2$
Fluorescent lamp	$8.2 \times 10^3 \text{ cd/m}^2$
Lambertian surface luminance, reflectance 0.7, illuminated with 500 lux	111 cd/m ²

Note: radiation and lighting system luminous efficacies are defined in Sections 2.2.1 and 2.2.2, respectively.

$$L_v = 683 \int_{380}^{770} L_\lambda V(\lambda) d\lambda \quad [\text{lm}\cdot\text{m}^2\cdot\text{sr}^{-1}] \quad (2.7)$$

For completeness, we show similar equations for converting the four spectral radiometric quantities used in the above integrals into their integrated photometric counterparts:

$$\Phi_e = \int_0^\infty \Phi_\lambda d\lambda \quad [W] \quad (2.8)$$

$$I_e = \int_0^\infty I_\lambda d\lambda \quad [W\cdot\text{sr}^{-1}] \quad (2.9)$$

$$E_e = \int_0^{\infty} E_{\lambda} d\lambda \quad [W \cdot m^{-2}] \quad (2.10)$$

$$L_e = \int_0^{\infty} L_{\lambda} d\lambda \quad [W \cdot m^{-2} \cdot sr^{-1}] \quad (2.11)$$

Now that the basic quantities of photometry have been defined, some additional quantities of interest can be introduced.

2.2.1 Radiation Luminous Efficacy, K_r , and the V -lambda Function

Radiation luminous efficacy, K_r , is the ratio of luminous flux (light) in lumens to radiant flux (total radiation) in watts in a beam of radiation. It is an important concept for converting between radiometric and photometric quantities (units: lumens per watt).¹

Luminous efficacy is not an efficiency since it is not a dimensionless ratio of energy input to energy output—it is a measure of the effectiveness of a beam of radiation in stimulating the perception of light in the human eye. If Q_v is any of the four photometric quantities defined previously and Q_e is the corresponding radiometric quantity, then the luminous efficacy associated with these quantities is given by the following defining equation:

$$K_r = \frac{Q_v}{Q_e} \quad [lm \cdot W^{-1}] \quad (2.12)$$

Q_e is an integral over all wavelengths for which Q_{λ} is nonzero, while Q_v is an integral (2.1) only over the visible portion of the spectrum where $V(\lambda)$ is nonzero. The luminous efficacy of a beam of infrared-only radiation is zero since none of the flux in the beam is in the visible portion of the spectrum. The same can be said of ultraviolet-only radiation.

¹The International Committee for Weights and Measures (CPIM), meeting at the International Bureau of Weights and Measures near Paris, France in 1977 set the value 683 lumens per watt for the spectral luminous efficacy (K_m) of monochromatic radiation having a wavelength of 555 nm in standard air. In 1979, the candela was redefined to be the luminous intensity, in a given direction, of a source that emits monochromatic radiation of frequency 540×10^{12} hertz and that has a radiant intensity in that direction of 1/683 W/sr. The candela is one of the seven fundamental units of the metric system (*System Internationale*, or SI). As a result of the redefinition of the candela, 683 is now not a recommended good value for K_m but follows from the definition of the candela in SI units. Prior to 1979, the candela was realized by a platinum approximation to a blackbody. After the 1979 redefinition of the candela, it can be realized from the absolute radiometric scale using any of a variety of absolute detection methods.

We can use luminous efficacy to explain the factor 683 in (2.1) and (2.4–2.7). Since $V(\lambda)$ is only the *relative* spectral response function for the eye, normalized to 1.0 at its peak value, an additional constant is needed to account for the *absolute* magnitude of the conversion of watts of radiation into lumens of light. This is what 683 is. It is the luminous efficacy of monochromatic radiation at the wavelength (555 nm) of maximum relative response of the human eye and it has units of lumens per watt.

It is worth noting that the 1924 $V(\lambda)$ function is based on observations by young, dark-adapted subjects using small fields of view, subtending angles of only 2 deg and 3 deg at the eye, with central fixation [3]. According to the CIE [3], “The field luminances were often less than $10 \text{ cd} \cdot \text{m}^{-2}$ and barely high enough to ensure photopic vision, particularly towards the ends of the visible spectrum. . . . Many later investigators have questioned the accuracy with which the $V(\lambda)$ function represents foveal photopic vision, and it is almost certainly too low at wavelengths shorter than 460 nm, too low by a factor of about ten at 400 nm. However, the inconvenience of changing the function has been regarded as outweighing any potential advantage.”

It is pointed out in Chapter 11 that the V -lambda function is the same, by definition, as a function $\bar{y}(\lambda)$ used in colorimetry for color-matching according to the CIE 1931 system of colorimetry. However, the values provided in Table 2.1 below 460 nm “do not properly represent the spectral sensitivity of most color normal human observers” [10,11]. Because of this, in 1990 the V -lambda function was changed slightly from the values recommended by the CIPM in 1976. The changes are quite small, but noticeable. For example, at the 555-nm wavelength of the peak value of $V(\lambda)$, the value is reduced from the 1.0002 value found in Table 2.1 to 1.0000. The values below 460 nm are also altered slightly. The resulting new function is called the *CIE 1988 modified 2-deg spectral luminous efficiency function for photopic vision* and has been given the symbol $V_M(\lambda)$. It is a *supplement* to, not a replacement of, the 1924 function [12] and is not used in this book.

With the above discussion of K , in mind, it is clear that 683 lumens per watt is the theoretical maximum luminous efficacy that a beam of radiation can have, and this value is valid *only* for monochromatic light at the specific wavelength 555 nm. (To be strictly correct, the maximum theoretical luminous efficacy according to Table 2.1 is actually $1.0002 \times 683 = 683.14$ lumens per watt, since the interpolated peak in the $V(\lambda)$ table at 555 nm just happens to be 1.0002 rather than 1.0. Although it may have been intended for $V(\lambda)$ to have a peak value of 1.0 at 555 nm, the original 1924 CIE definition of $V(\lambda)$ [13] gives values only at 550 and 560 nm and does not therefore require $V(\lambda)$ to have a maximum value of 1.0. Nor does it specify the value at 555 nm. The value of 1.0002 at 555 nm resulted from interpolation based on a curve fit to the original 1924 CIE data at 10-nm intervals. The more recent modification of the interpolated $V(\lambda)$ values to the new

$V_M(\lambda)$ function addresses this problem and changes the value at 555 nm to 1.0. (See Section 11.4 for more information.)

Other sources of light have other wavelengths present and many have radiation outside the visible spectrum. Invisible ultraviolet and infrared radiation have watts of power but no lumens of light in them. Their luminous efficacies are zero and they cannot be called light. The more watts of invisible radiant power one has in a beam, the lower the luminous efficacy has to be.

Table 2.4 gives the luminous efficacies of several different types of radiation. All radiation is eventually absorbed and converted into either heat or chemical (or biochemical) energy. (In some cases, as with radiation-pumped lasers and fluorescent materials, some of the radiation absorbed by a material can be reemitted, usually at different wavelengths and in different directions.) Luminous efficacy can therefore be thought of as the ratio of light “energy” to heat energy contained in a beam of radiation.

2.2.2 Lighting System Luminous Efficacy, K_s

Lighting system luminous efficacy, K_s is the ratio of the light output in lumens from some (electric or other) lamp or lighting system to the total (electrical or other) power in watts supplied to the system (units: lm/W). This is a different kind of luminous efficacy than radiation luminous efficacy, but it has the same units. K_s is used mostly in the field of illuminating engineering, but it has other applications. The defining equation is

$$K_s = \frac{\Phi_v}{P_e} \quad (2.13)$$

Table 2.4
Luminous Efficacies of Radiation from Some Light Sources

Source	Luminous Efficacy
Monochromatic light, 555 nm	683 lm/W
White light, constant over visible spectrum	220 lm/W
Blue sky light	125–140 lm/W
Direct beam sunlight, midday	90–120 lm/W
Direct beam sunlight, extraterrestrial	99.3 lm/W
Direct beam sunlight, sunrise/sunset	50–90 lm/W
Overcast sky light	103–115 lm/W
Uncoiled tungsten wire at its melting point	53 lm/W
Radiation from typical tungsten filament lamp	15 lm/W
Phosphorescence of cool white fluorescent lamp	348 lm/W

where Φ_v is the total luminous flux in lumens produced by the system and P_e is the (electrical or other) power input to the system in watts. A laser beam producing radiation at 555 nm will have $K_r = 683 \text{ lm/W}$, while K_s might be only 1 lm/W.

Sometimes Φ_v can be defined as the *useful* luminous flux, the portion of the total emitted flux delivered to where it is needed. Some example electric lighting system luminous efficacies are shown in Table 2.5. Note that these do not include ballast losses, energy losses in the auxiliary electrical circuits that condition the power supplied to the lamps. When the latter are included, the system efficacies can drop considerably. New electronic ballasts, however, are being developed which promise substantially increased *electrical efficiencies*, and correspondingly higher *luminous efficacies* for modern lighting systems.

Again it is important to note that luminous efficacy K_s of an electric lighting source or system is not an *efficiency* of that source or system. The CIE uses the symbol η or η_e for the *radiant efficiency* of a source of radiation, the ratio of the radiant flux emitted by the source to the electrical power consumed by it. This comes close to being a true efficiency, being defined as the ratio of energy output to energy input. The symbol η is widely used for efficiency in other fields. However, the CIE also uses the symbol η or η_v for the *luminous efficacy of a source* of radiation, defined as the ratio of the luminous flux emitted to the electrical power consumed by the source. This is the same as the definition of K_s given above. I prefer to keep the symbol η for efficiency and K for efficacy, using subscripts to distinguish between radiation and source efficacy.

This ends the introduction of the basic concepts of photometry. The treatment is relatively short, since the subject of radiometry has already been introduced, so little additional effort is needed to introduce photometry. All of the concepts and relationships introduced in Sections 1.5 and 1.6 are applicable to photometry as well. Next is presented some information about the connection between the purely physical concepts introduced above and the *perception* of light by the human eye/brain combination. There is still more to be said about photometry in later chapters, most notably Chapters 7 through 11 forming the second half of the book, but also in the next chapter dealing with a special kind of radiation source, called a *black-body*, and concepts related to its properties.

Table 2.5
Electric Lighting System Luminous Efficacies

<i>Lighting System</i>	<i>Luminous Efficacy</i>
Tungsten filament lamp	12–17 lm/W
Hot cathode cool white fluorescent lamps, not including ballast	30–70 lm/W
Cold cathode cool white fluorescent lamps, not including ballast	38–52 lm/W
High pressure sodium vapor lamps, not including ballast	60–140 lm/W
Metal halide lamps, not including ballast	75–125 lm/W

2.3 LUMINANCE AND BRIGHTNESS

When an idealized lens forms a perfect image of an object, as indicated in Figure 2.2, rays from a point in the object (the base of the arrow in Fig. 2.2) are focused to a corresponding point in an (inverted) image of the object. If the limiting aperture in the lens is circular and we are looking at an image point on the axis (the geometrically symmetrical center line) of the lens, then the rays diverging from the point in the object to this aperture form what is called a *pencil* of rays, since they fill a right-circular cone solid angle, defined by the point in the object and the circular aperture of the lens. This pencil of rays can be seen on the left side of Figure 2.2.

Similarly, the rays converging to the corresponding image point to the right of the lens also form a solid-angle pencil of rays. The luminous flux contained in these solid angles can best be described by the luminance. Of first interest is the luminance of the surface of the object, at point P, in the directions contained in the *diverging* pencil of rays on the left side of the lens in Figure 2.2. Second is the luminance of the image surface at point P' in the directions contained in the *converging* pencil of rays on the right side of the lens in the figure.

Let $L_{vo}(x, y, \theta, \phi)$ be designated the luminance of the source at object point P having coordinates (x, y) in the source plane and in direction (θ, ϕ) inside the diverging solid angle. Designate $L_{vi}(x', y', \theta', \phi')$ as the luminance of the image at point P' having coordinates (x', y') in the image plane and in direction (θ', ϕ') . In Chapter 4 is described a way of relating $L_{vi}(x', y', \theta', \phi')$ to $L_{vo}(x, y, \theta, \phi)$. Basically, L_{vi} is proportional to L_{vo} for corresponding points in the object and its image. For now, let us restrict the discussion to the image alone and an attempt to describe it photometrically.

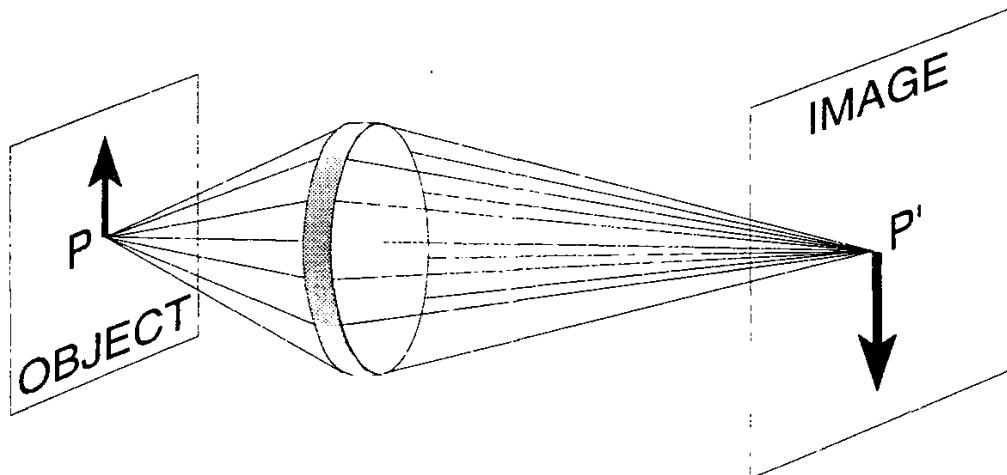


Figure 2.2 Formation of the (magnified) image at P' of an object at P by an idealized perfect lens.

It seems obvious that the image is made up of a spatial variation in luminance over the image plane of the ideal lens. This is true, but most image receiving surfaces, whether they be a white sheet of paper, the retina of the eye, or photographic film in a camera, gather luminance from all directions, covering the whole 2π -sr hemispherical solid angle, essentially converting it to luminous or radiant flux before expressing this flux as a chemical change or as reflected light. Thus we see that an image can also be thought of as a spatial distribution of irradiance or illuminance over an imaging surface, in this case the image plane shown on the right in Figure 2.2.

The action of most image-receiving surfaces is to redirect the total flux incident upon them. Other image-receiving surfaces convert the total spatial distribution of incident flux into physical or chemical patterns (as in a photographic emulsion) or electrical signals (as with two-dimensional detector arrays). In the latter case, the spatial distribution of *luminance* incident on the image-capturing surface is converted into a spatial distribution of response that is related to the *illuminance* incident on the surface. Thus, in describing the response of an imaging system to the incident flux, it is the image plane illuminance distribution in which one is interested.

If Ω_i is the solid angle containing the pencil of rays converging to point P' in Figure 2.2, then the photometric equivalent of (1.23) can be integrated over this solid angle to obtain the following expression for the distribution of illuminance $E_{vi}(x', y')$ in the image, in terms of the incident luminance distribution $L_{vi}(x', y', \theta', \phi')$:

$$E_{vi}(x', y') = \int_{\Omega_i} dE_{vi}(x', y') = \int_{\Omega_i} L_{vi}(x', y', \theta', \phi') \cos \theta' d\omega \quad (2.14)$$

If the lens in Figure 2.2 is the human eye, then it is clear that the illuminance distribution in (2.14) comes originally from the luminance distribution $L_{vo}(x, y, \theta, \phi)$ emerging from the object. From this strictly physical perspective, it appears that it is the luminance distribution of the source that we see with our eyes.

To be more correct, one should speak of the spectral radiance $L_{\lambda o}$ of the source, because most eyes are capable of distinguishing the different colors in this spectral distribution. To be even more correct, one should be still more careful and follow the principles stated in Chapter 11, since the perception of color involves complex, spectrally selective, biochemical and psychophysical processes that operate on the spectral distribution of irradiance received by the retina. However, for the purposes of this section, we are interested only in the more physical photometric aspects of the radiation received by the retina, before further processing takes place. Color vision will be discussed in Chapter 11 and the process of perception of objects is presented in the next section.

Here we speak of the eye as “seeing” luminance variations in the objects that we observe. Thus, when we perceive one object as “brighter” than another, the physical difference that most contributes to this perception is the difference in luminance between the two objects. The perception of “brightness,” which has not been defined, is linked to the luminance of the object being viewed. At this point, we are mainly interested in emphasizing that it is the emitted luminance of an object that the human eye responds to rather than its emitted illuminance. Only a portion of the flux incident on an object (and characterized by the illuminance distribution incident on it) is reflected by it and only a portion of the reflected flux enters the eye. Most objects reflect the flux received by them into the whole $2\pi\text{-sr}$ hemispherical solid angle. The lens and iris of the eye pass only a small portion of this flux on to the retina. Problem 2.3 at the end of this chapter provides an estimate of the magnitude of the solid angle of flux from a point in a distant object passed by the iris to the retina. It is shown that only a very small portion of the total flux reflected from a point in the object enters the eye.

The reader may recall the earlier instruction (Section 2.2) to always specify in discussions of luminance the surface, point in that surface, and direction. In the previous two paragraphs, it may seem that this dictum has been violated. However, it is implied in the discussion that the point is actually several points on the surface of an object and that the direction is from these points into the eye of a person looking at them. This implied geometrical arrangement is continued into the next section.

2.4 LUMINANCE AND VISION

Photometry is human centered. It deals exclusively with the human eye as a radiation detector and evaluates the flux of radiation in terms of how the human eye responds to it. When humans look at something, an image of what they are looking at is placed (inverted) on the retina at the back of the eye and the image so displayed is converted into electro-chemical signals that are sent to the brain, with interpretation and modification of those signals taking place along the way. It once was thought that most of the processing takes place only in the brain, but research indicates that processing takes place at many stages, from retinal reception to the perception of colors and shapes in the brain. From a strictly physical perspective, when we look at an object what we are seeing is a modified distribution of the spectral radiance leaving that object and converted to spectral irradiance incident on the retina. In reality, the eye/brain perception of brightness and color is a psychophysical activity. Color vision is discussed in Chapter 11. Here we are interested more in the perception of shapes and how brightness and contrast contribute to that perception. The reader may wish to read the excellent study of this subject provided in a mathematical model of color vision devised by Prof. R.W.G. Hunt [14].

According to Shapely [15], “Although it seems that the perception of brightness of objects is effortless, visual scientists have known for years that complex neuronal computations are required to perform this task. The perception of brightness is not simply a matter of counting photons. The primary determinant of brightness perception is local contrast—the local difference between luminances on either side of a boundary normalized by the (local) average luminance.” The *IES Lighting Handbook* puts it this way [2]: “Contrast detection is the basic task from which all other visual behaviors are derived. The visual system gives virtually no useful information when the retina is uniformly illuminated, but is highly specialized to inform about luminous discontinuities and gradients in the visual field.” Further information on this topic can be found in references [16] through [18].

Thus we see that luminance differences are as important as (or more so than) their absolute levels in our ability to see and perceive objects. Scientists attribute this to the evolutionary selection process in animals, which made them capable of seeing reflectances in nature rather than overall luminance levels. It is reflectance *differences* in nature that remain the same regardless of illumination levels. Thus, it is thought that we evolved very powerful abilities to see differences in luminances when viewed adjacent to each other, while at the same time being able to accommodate very wide variations in the overall light level.

According to Williams [19], “The ratio of the luminance of a white piece of paper that is just detectable to its luminance in bright light is $10^{[-11]}$.” This ability allowed us during our hunting/gathering period to see the luminance differences required for survival over the wide range of light levels from starlight to bright daylight. On the other hand, if we walk into a room whose walls are brighter than those in the room we just left, we are scarcely aware of the difference. Our eyes adjust easily to changes in overall illumination levels and these are difficult to perceive unless they are presented simultaneously, side by side, or sequentially, with little time for the transition.

Our ability to adjust to different light levels is called adaptation. It is accomplished by changes in the diameter of the pupil (from about 2 mm up to 8 mm, a 16-fold change in area), by neuronal adaptations, and by photochemical adjustments in the retina. Although these processes enable us to see over a wide range of illumination levels, it is also true that we generally tend to see better with more light. That is, our ability to detect luminance contrasts increases with the overall luminance of the viewed target.

In order to understand this phenomenon better, we introduce the concept of luminance contrast C , which can be defined in several ways. Three are shown below:

$$C = \frac{L_d - L_b}{L_b} \quad (2.15)$$

where L_d is the luminance of the detail and L_b is the luminance of the background,

$$C = \frac{L_g - L_l}{L_g} \quad (2.16)$$

where L_g is the greater of the luminances and L_l is the lesser of them, and

$$C = \frac{L_g - L_l}{L_g} \quad (2.17)$$

These different definitions are useful in different applications. An additional measure of contrast, called the *modulation*, M , is defined as

$$M = \frac{L_{\max} - L_{\min}}{L_{\max} + L_{\min}} \quad (2.18)$$

When the objects being viewed are opaque and not self-luminous, such as black text on white paper, the luminances of the object and background depend on their reflectances. If the viewing direction remains the same while the magnitude only of the illumination of the target is varied, the above contrasts do not vary. Reflectance contrast is independent of illumination level, all other things being equal. However, the reflectances of most objects vary with direction, so that the apparent contrast of such objects can also vary with positions of both the source and the observer.

According to the *IES Lighting Handbook* [1]:

The simplest visual function is one in which a small change in luminance must be detected in an otherwise uniform surround. This function has been studied in great detail and has related the probability of detecting a small round test object on a uniform background to the contrast and the luminance of the background. . . As the contrast is raised, the probability of seeing a test object increases until, at a certain contrast, the object can be detected 100 percent of the time. The contrast at which the object can be detected 50 percent of the time is called the *threshold contrast* and will vary among individuals, with the duration of exposure of the object, with the size of the object, and with the luminance of the background.

If a group of observers is tested for this effect, one can obtain a relationship giving the reduction in threshold contrast with increasing background luminance. Some approximate values of this relationship are given in Table 2.6.

Table 2.6
Human Visibility Reference Function

<i>Luminance Contrast Using (2.15)</i>	<i>Background Luminance, cd/m²</i>
0.08	1,000
0.2	10
2	0.1
100	0.001

The curve on which the data in Table 2.6 are based has been used by the Illuminating Engineering Society to illustrate the fundamental relationship between object detection and the luminance of the background. The data make clear the relationship between seeing ability and illumination level. For more complex seeing tasks, however, more explicit and detailed measures of visual performance are needed. For more information on this important subject the reader is referred to Section 3 of the *IES Lighting Handbook* or textbooks in the fields of vision and illumination.

2.5 DISABILITY GLARE

The way a visual target is illuminated and viewed can have a strong effect on its luminance contrast. The reflectance of most smooth or slick materials is greatest when the angle of reflection equals the angle of incidence, as illustrated in Figure 2.3. *Specularly reflecting* surfaces are surfaces whose reflectance is essentially zero at angles of reflectance not equal to the angle of incidence. *Diffusely reflecting* surfaces reflect radiation from an incident ray into all directions. Many diffuse surfaces still have a strong specular component in their reflectances. Reflecting surfaces can be categorized as follows:

Matt or matte surfaces—all diffuse and no specular reflection.

Glossy surfaces—mostly diffuse but with some specular reflection, as depicted in Figure 2.3.

Mirrors and polished glass surfaces—all specular and no diffuse reflection.

The subject of reflection is dealt with more completely in Chapter 6.

The reflectance of most slick or glossy surfaces, such as the lacquered paper found in most magazines, increases with angle of incidence. This is true of both the white uninked and the black-inked portions of the pages of such magazines. At some angles, the reflectance of the ink, which normally is much lower than that of the uninked paper, can equal or exceed that of the paper. This means that at some angle of incidence, the contrast of the text on the paper can disappear completely, with both the inked and uninked portions of the page being uniformly

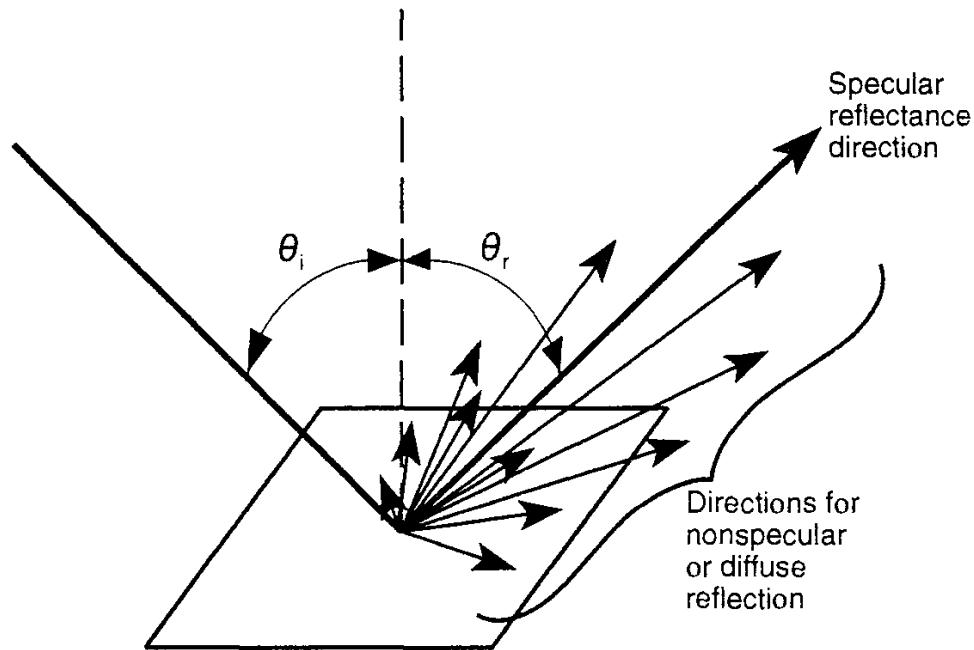


Figure 2.3 Illustration of the way surfaces reflect light. Generally the reflectance is maximum in the specular direction, where the angle θ_r of reflection equals the angle θ_i of incidence. Reflected rays in nonspecular (diffuse) directions are also shown.

luminous, causing the luminous contrast to go to zero, and rendering the text invisible. When this happens, the seeing task becomes impossible and the ability to read the text or see the text becomes disabled, resulting in a condition called *disability glare*. Even if the task contrast is not eliminated completely, it can still impair the performance of the visual task and produce disability glare. When the illumination of a visual target comes from a direction so as to reduce the luminance contrast compared with uniformly diffuse illumination, we speak of this illumination as being a source of unwanted disability glare. A common example of this condition is illustrated in Figure 2.4. Solutions to this problem include redesign of the light source and changing the relative positions of the source, the visual task, and the observer.

Another common example is that of a computer display screen reflecting light into the eye from a small or extended source, such as a window behind the computer operator. The reflected light from this source can produce a veiling luminance that is as bright or brighter than the screen, reducing contrast and producing disability glare. Solutions to this problem include reducing or removing the specular component of the reflectance from the screen, removing or repositioning the glare source, and reorienting the screen and operator so that the specular reflection does not enter the operator's eye.

Measurement of the luminances and reflectances leading to disability glare is an important aspect of the field of applied photometry or illumination engineering.

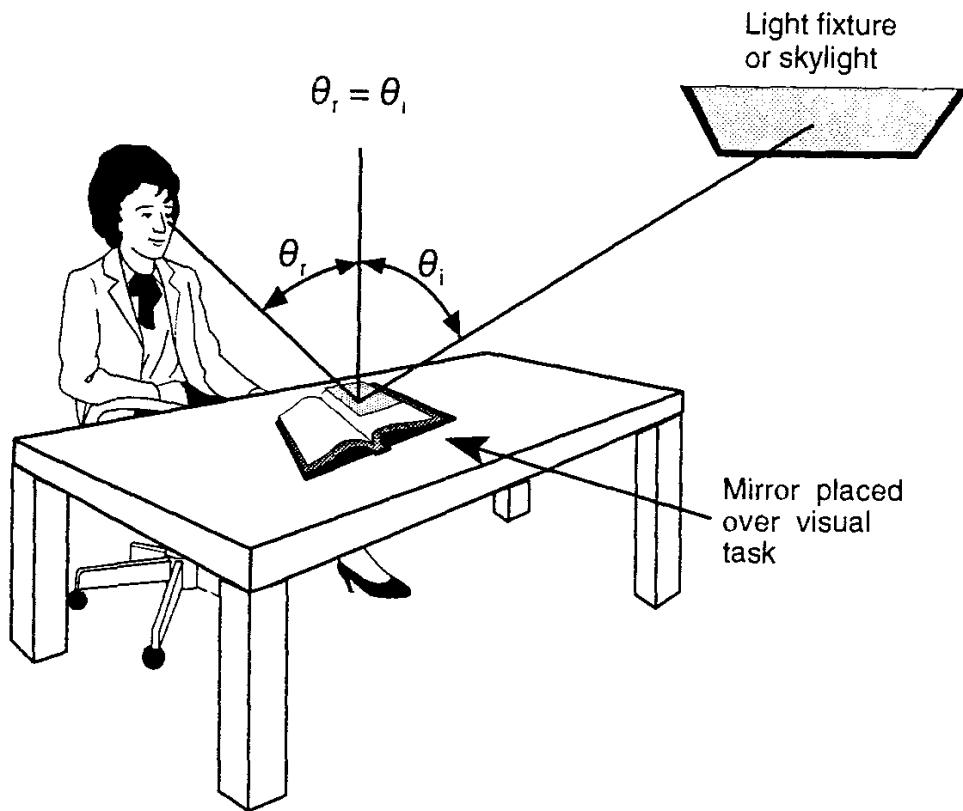


Figure 2.4 If a bright light source is visible in a mirror placed over a visual task, then the source is a potential source of disability glare, depending upon the nature of the target reflectance.

2.6 DISCOMFORT GLARE

There is another important type of glare that is very prevalent but largely unidentified. When bright light enters the eye from the side, away from the direction of view (away from the optical axis of the eye), it confuses the eye's natural adaptation response and produces some veiling flux on the retina resulting from scattering of light from inhomogeneities in the lens and vitreous humor of the eye. The condition is called *discomfort glare*. In extreme cases it can be very uncomfortable or even painful and can lead to noticeably reduced visual performance. It is caused by high or nonuniform distributions of brightness within the field of view and can be accompanied by disability glare, but it is a distinctly different phenomenon. An example of a condition producing discomfort glare is illustrated in Figure 2.5. Another example is the light from the high-beam headlights of an oncoming car. This extreme case of discomfort glare can actually make the road surface invisible to the driver and can lead to serious automobile accidents.

There is a fairly easy way to detect the presence of discomfort glare. An opaque object, such as a piece of cardboard or a large book, is held next to the



Figure 2.5 Bright light from the side can cause discomfort glare.

eyes out of the line of sight but to the side, above or below the direct line of view, shading the eyes from the glare source. Then it is removed from this position. If the presence of the shading object in one or more of these locations results in a sensation of improved comfort, a slightly greater sense of well being, and/or an improved ability to see straight ahead, then discomfort glare likely is present. This phenomenon is most noticeable when one sits in an under-illuminated, relatively dark room, with a bright window to one side. The problem is exacerbated if the interior room surfaces have relatively low reflectances, such as with dark wood paneling.

Most attempts at a quantitative assessment of discomfort glare are based on the size, luminance, location in the field of view, and number of glare sources, as well as the background luminance. The determination of visual discomfort is a psychological or psychophysical problem and is usually based on determining the conditions that cause the subject to just barely perceive discomfort. This threshold is termed the “borderline between comfort and discomfort,” or BCD. For more information, see the CIE’s publication on discomfort glare [20], the *IES Lighting Handbook*, or Murdoch’s textbook on illumination engineering [21].

We are almost continually confronted by discomfort glare. It occurs in nature when direct beam sunlight enters our eyes, either directly or after reflection from shiny surfaces. It is often produced outdoors by man-made, shiny objects. Sunlight from automobile surfaces can be a particularly bad source of glare. When this sunlight enters a building through a window, the glare effect is even more pronounced, due to the generally lower ambient interior light levels.

2.7 ILLUMINATION

The provision of adequate light for various human activities has been a matter of some importance since the emergence of civilization. The basic question that has faced the lighting designer is "how much light do we need to see." The IES has dealt with this central problem since its beginning in 1906. From the *IES Lighting Handbook* [2]:

It has been known since the mid-1800s that, in general, it is easier to perform simple tasks as the luminance level of a visual environment becomes higher. However, it was not until the 1930s that the effect of lighting level on specific industrial and office tasks was investigated. In England, the problem was approached by measuring the performance on actual industrial tasks under varying lighting conditions; from such field data it was possible to assess the relative increase in task performance (in most cases restricted to their visual components) for a given increase in luminance. For years the British lighting recommendations were based on that work.

In the United States, that approach was used to assess the relative importance of lighting levels on the speed of reading. Data from that study, too, were limited to specific tasks and could be generalized to other situations only with great caution.

In 1937, and again in 1959, a different approach was taken, based on improved understanding of the basic processes involved in vision, seeing, and performing visual tasks. The resulting work formed the basis of the illuminance recommendations of the CIE and for many years was the foundation of the IES method of prescribing illumination. This work was subsequently extended [2,3] and now includes the possibility of incorporating variations in the specified illuminance levels based on the contrast of the visual target, the criticality of the visual task being performed, and the age of the person performing the task. A summary of the IES illuminance selection procedure is presented below.

Determinations of adequate *quantities* of illumination cannot be made without some consideration being given to the *quality* of the illumination. Quality in illumination refers mainly to the absence of glare. It also refers to other visual effects that illumination can produce. It is the absence of glare, however, that is the most important quality of illumination to be considered when selecting appropriate illuminance levels for the performance of specific visual tasks.

We have seen that the lower the contrast of a visual target, the higher must be its overall luminance for us to see it. For constant reflectance tasks, illuminated from the same direction, if one needs more *luminance from* the target, the higher must be the *illuminance on* the target. Thus, the need for illuminance is directly

tied to the presence or absence of contrast reducing glare. The greater the glare, the greater the need for illuminance on the target.

We can avoid disability glare by presenting the target with light mostly from the side, so that the specularly reflected and hence brightest light is reflected off to the side and away from the field of view. If the light source itself, however, remains in the field of view and is bright, then it can become a source of discomfort glare, so it is best to position the light source behind the observer, either above or to the side. We can reduce the effects of discomfort glare by increasing task luminances relative to the luminance of the surroundings. It is important, however, not to make task luminances much greater than surround luminances. The balancing of illumination levels with lighting quality is central to the field of illumination engineering. For further information on this subject, the reader is referred to a variety of texts in the field.

2.7.1 Illuminance Selection

The IES illuminance selection guidelines are summarized in Table 2.7. The range from 20 to 20,000 lux of illuminance on the task is divided into nine ranges, lettered A through I. In each of the nine categories, a low, mid, and high illuminance value is given. The number 1, 2, or 3 is assigned to whether the average age of the occupants is under 40, between 40 and 65, or over 65, respectively. Another 1, 2, or 3 is assigned according to whether the background reflectance of the task is over 70%, between 30% and 70%, or below 30%, respectively. Finally, another 1, 2, or 3 is assigned according to whether the visual task demand for speed and/or accuracy is not important, is important, or is critical. The three numbers obtained from these selections are then added together and the result is divided by 3. If the result is 1, 2 or 3, then the low, medium, or high illuminance level in the selected category, respectively, is chosen.

To assist in the use of Table 2.7, the IES offers an extensive table of visual tasks and designates the illuminance selection category that goes with each. The reader is referred to the *IES Lighting Handbook* [2] or the shorter *IES Lighting Ready Reference* [22] for specific entries in this table. A couple of examples can be given: In the section on "Reading," one finds under "Handwritten Tasks" an entry for "#4 pencil and harder leads" that calls for illuminance category F. Under "Printed Tasks," the subheading for 8- and 10-point type calls for category D.

Under "Residences, general lighting, conversation, relaxed and entertainment," one finds category B, which happens to be the same as the one listed for "Dance halls and discotheques."

One can see in the very way in which the IES recommends illuminance levels a recognition of the importance of target contrast, criticality of task, and age of

Table 2.7
IES Illuminance Categories and Illuminance Values for Generic Types of Activities Indoors

<i>Type of Activity</i>	<i>Illuminance Category</i>	<i>Range of Illuminances</i>		<i>Reference Work Plane</i>
		<i>Lux</i>	<i>Footcandles</i>	
Public spaces with dark surroundings	A	20–30–50	2–3–5	General lighting throughout space
Simple orientation for short temporary visits	B	50–75–100	5–7.5–10	
Working spaces where visual tasks are only occasionally performed	C	100–150–200	10–15–20	
Performance of visual tasks of high contrast or large size	D	200–300–500	20–30–50	Illuminance on task
Performance of visual tasks of medium contrast or small size	E	500–750–1000	50–75–100	
Performance of visual tasks of low contrast or very small size	F	1k–1.5k–2k	100–150–200	
Performance of visual tasks of low contrast or very small size over a prolonged period	G	2k–3k–5k	200–300–500	Illuminance on task, obtained by combination of general and local lighting
Performance of very prolonged and exacting visual tasks	H	5k–7.5k–10k	500–750–1,000	
Performance of very special visual tasks of extremely low contrast and small size	I	10k–15k–20k	1k–1.5k–2k	

observer. To give another example, for the case of a 50-year old casual reader of a textbook having 10-point type in an office setting with a 75% reflectance, we first choose category D and then add the numbers 2 for age, 1 for reflectance, and 1 for criticality, to obtain the number 4. Dividing this by 3 yields 1.33 so that we would choose an illuminance about 1/3 of the way from 200 to 300 lux, or approximately 235 lux as our desired illuminance level. This 235 lux illuminance level is equivalent to about 23.5 fc in the inch-pound system of measurements. This value is close enough to the strictly correct converted value of 21.8 fc to be acceptable. This is why the IES uses the factor of 10 to convert between I-P and metric units in Table 2.7.

EXAMPLE PROBLEM 2.1

Problem: A collimated beam of radiation has the (incident) spectral irradiance distribution E_λ given in Table 2.8. This beam is incident normally (perpendicularly) on a parallel plate of clear-but-tinted glass whose transmittance spectrum T_λ is also given in Table 2.8 along with the relevant V_λ values.

Table 2.8
Worksheet for Example Problem 2.1

λ in nm	E_λ in $W m^{-2} nm^{-1}$	T_λ	V_λ	$E_\lambda \times T_\lambda$	$E_\lambda \times V_\lambda$	$E_\lambda \times T_\lambda \times V_\lambda$
500	0	0.8	0.323	0	0	0
510	1	0.81	0.503	0.81	0.503	0.40743
520	1.4	0.81	0.71	1.134	0.994	0.80514
530	2	0.81	0.862	1.62	1.724	1.39644
540	2.4	0.81	0.954	1.944	2.2896	1.854576
550	3.1	0.8	0.995	2.48	3.0845	2.4676
560	3.6	0.78	0.995	2.808	3.582	2.79396
570	4.5	0.76	0.952	3.42	4.284	3.25584
580	5.3	0.73	0.87	3.869	4.611	3.36603
590	6.3	0.7	0.757	4.41	4.7691	3.33837
600	7.1	0.68	0.631	4.828	4.4801	3.046468
610	8	0.64	0.503	5.12	4.024	2.57536
620	8.9	0.61	0.381	5.429	3.3909	2.068449
630	9.5	0.57	0.265	5.415	2.5175	1.434975
640	9.9	0.52	0.175	5.148	1.7325	0.9009
650	10.3	0.49	0.107	5.047	1.1021	0.540029
660	10.7	0.45	0.061	4.815	0.6527	0.293715
670	11.1	0.41	0.032	4.551	0.3552	0.145632
680	11.5	0.38	0.017	4.37	0.1955	0.07429
690	11.9	0.31	0.0082	3.689	0.09758	0.0302498
700	12.2	0.25	0.0041	3.05	0.05002	0.012505
710	12.3	0.2	0.0021	2.46	0.02583	0.005166
720	12.4	0.14	0.00105	1.736	0.01302	0.0018228
730	12.4	0.09	0.00052	1.116	0.006448	0.00058032
740	12.3	0.05	0.00025	0.615	0.003075	0.00015375
750	12.1	0	0.00012	0	0.001452	0
760	11.8	0	0.00006	0	0.000708	0
770	11	0	0	0	0	0
780	8	0	0	0	0	0
790	3.5	0	0	0	0	0
800	0	0	0	0	0	0
Sum $\times \Delta\lambda$	2365			798.84	444.90	308.16
Units	$W \cdot m^{-2}$			$W \cdot m^{-2}$		

Calculate, approximately (without interpolation), the total incident irradiance $E_{e,i}$, the incident illuminance $E_{v,i}$, the total transmitted irradiance $E_{e,t}$, the transmitted illuminance $E_{v,t}$, the integrated radiant transmittance $T_e = E_{e,t}/E_{e,i}$, the integrated luminous transmittance $T_v = E_{v,t}/E_{v,i}$, and the luminous efficacies $K_{r,i}$ and $K_{r,t}$ of the incident and transmitted beams, respectively.

Solution: First we prepare the worksheet shown in Table 2.8 and calculate the entries in it. Since we are asked only for an approximate result, we do not need to use an interpolation scheme to approximate the shape of the spectral distributions between the tabulated wavelengths. The incident irradiance spectrum, as well as the V -lambda and glass transmittance spectra, are drawn in Figure 2.6. We sum the entries in the incident spectral irradiance column $E_{\lambda i}$, the transmitted spectral irradiance column $T_\lambda E_{\lambda i}$, and the columns for the products of the incident and transmitted spectral irradiances with the V -lambda values, $T_\lambda E_{\lambda i}$ and $T_\lambda E_{\lambda i} V_\lambda$, respectively. The results are given at the bottom of Table 2.8, multiplied by the 10-nm wavelength interval. The final answers are as follows:

$$E_{ei} = 2365 \text{ W/m}^2$$

$$E_{vi} = 683 \text{ lm/W} \times 444.9 \text{ W/m}^2 \\ = 304 \text{ klm/m}^2 = 304 \text{ klx}$$

$$E_{et} = 798.8 \text{ W/m}^2$$

$$T_e = 798.84/2,365 = 0.338$$

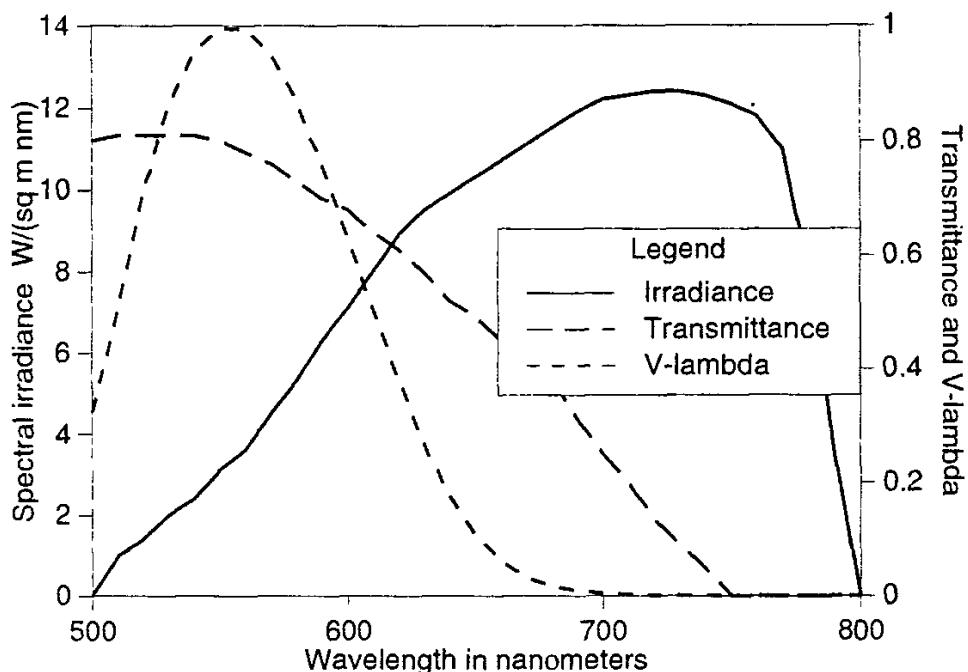


Figure 2.6 Spectral distributions of incidence irradiance, of glass transmittance, and of the V -lambda curve used in Example Problem 2.1.

$$E_{v,i} = 683 \text{ lm/W} \times 308.16 \text{ W/m}^2 \\ = 210 \text{ k lux}$$

$$T_v = 210/304 = 0.691$$

$$K_{ri} = 304,000/2365 = 128.5 \text{ lm/W} \text{ (incident radiation luminous efficacy)} \\ K_{rt} = 210,000/798.8 = 262.9 \text{ lm/W} \text{ (transmitted radiation luminous efficacy)}$$

It is seen from these numbers that the glass is acting as a spectral filter, transmitting only visible radiation. One result is an increase in the luminous efficacy of the radiation passing through it and a decrease in the heat-producing radiant flux.

It would be interesting to define a couple of figures of merit for this transmitting plate. One would be the ratio of the luminous to radiant transmittances, T_v/T_e , which in this case is 2.0, and another would be a kind of luminous efficacy transfer function, the ratio of the transmitted to incident luminous efficacies, K_{rt}/K_{ri} , which in this case is also 2.0. Writing these ratios in terms of $E_{e,i}$, $E_{v,i}$, $E_{e,t}$, $E_{v,t}$ and rearranging terms shows the ratios to be identical.

EXAMPLE PROBLEM 2.2

Problem: How much of the flux in the solar spectrum is contained in the visible part?

Solution: Since the range of the visible portion of the spectrum is variable depending upon the observer (see Table 1.1), we must integrate the solar spectrum between variable limits containing the visible portion of the spectrum. So, we begin by taking the solar spectrum shown in Figure 2.7 and integrating it both over its whole range and between a set of different wavelength limits on either side of and including portions of the V-lambda curve. Although data for $V(\lambda)$ are provided in Table 1.1, data for the solar spectrum are not provided. The reader is therefore not expected to work this problem in detail, only to follow the logic of the solution.

We then take the integrated solar irradiance between these limits and divide the result by the total integrated irradiance over the whole solar spectrum, calculated by the summation version of (2.10). (See Appendix for instructions on how to convert integrals to sums.) The illuminance contained within the indicated wavelength band is also calculated, (using the illuminance version of (2.1)), and divided by the total illuminance (2.6). The results are shown in Table 2.9.

These calculations are best performed using a spreadsheet computer program or a programming language such as Basic or Fortran. The solution to the problem is shown in the last two columns of Table 2.9. In looking at Table 2.9, it is important to note that all the colors of the spectrum are distributed over the wavelength range from 370 to 770 nm. Thus, if we cut out some of the light at the edges of the spectrum, as indicated by the lower entries in Table 2.1, we will be cutting out some colors, violet at the lower end and red at the upper end of the range. The effect of this on the apparent color of the resulting light is a question of some

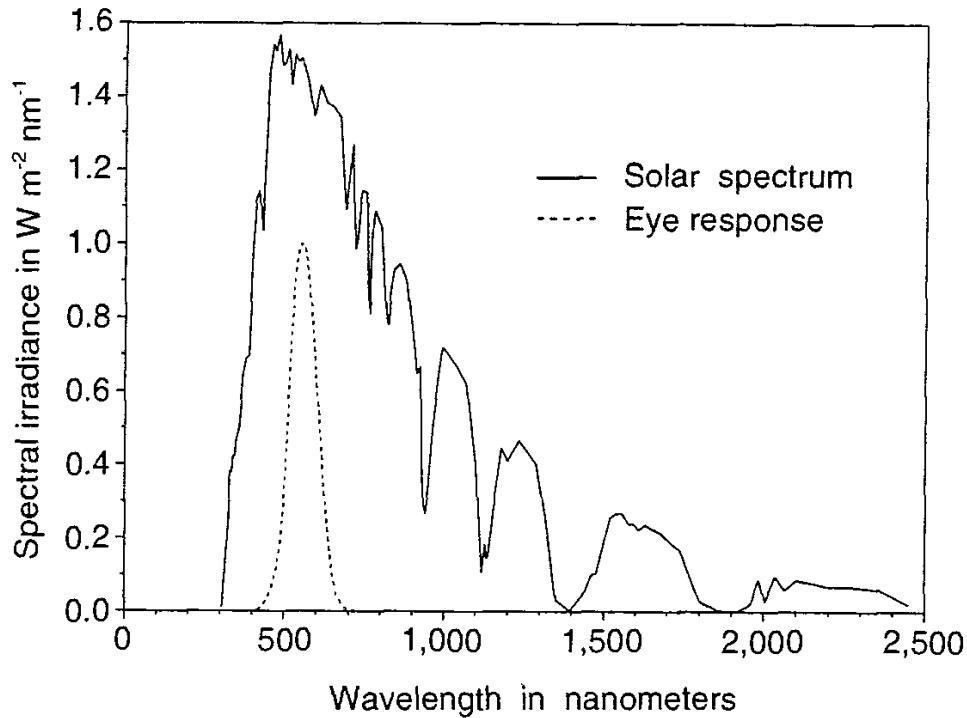


Figure 2.7 Spectra of the sun and the human photopic luminous efficiency curve, showing the shapes of the two spectra. The human visibility curve, which is dimensionless, peaks at a value of 1.0.

Table 2.9
Table of Values for Example Problem 2.2

Portions of total solar spectral irradiance contained in various portions of the visible spectrum.
Integrated total irradiance and illuminance, from Figure 2.7: 950 W/m² and 100 k·lux

Start Wavelength in nm	End Wavelength in nm	Irradiance Fraction in Percent	Illuminance Fraction in Percent
370	770	54.4	100.0
380	760	52.2	100.0
390	750	50.2	99.9
400	740	47.4	99.9
410	730	44.9	99.8
420	720	41.9	99.8
430	710	39.5	99.8
440	700	36.7	99.8
450	690	35.3	99.5
460	680	31.1	99.1

importance in the design of windows for buildings that are good at blocking solar radiant heat while admitting visible light.

EXAMPLE PROBLEM 2.3

Problem: The iris of the eye has a diameter of about 5 mm (0.005 m) when partially dilated. If one looks at an object on a wall 2 meters away, what fraction of the hemispherical solid angle does the iris subtend at a point in the middle of the object on the wall?

Solution: The radius of the iris is 0.0025m, giving a half-angle for the cone from the point on the wall to the iris of $\tan^{-1} (.0025/2) = \tan^{-1} (0.00125) = 0.0716$ deg. Using (1.5) with $\alpha = 0.0176$ deg, we have a solid angle Ω of 4.91×10^{-6} sr or $4.91 \mu\text{sr}$. Dividing this by 2π sr yields the fraction 7.81×10^{-7} of a hemispherical solid angle as our answer to the problem. If the source object is Lambertian at the point of observation, then this is the fraction of total emitted flux from the object point that falls within the iris of the eye.

REFERENCES

- [1] CIE Proceedings 1924, Cambridge University Press, Cambridge, 1926, pp. 67 and 232.
- [2] Kaufman, J. E., ed. *IES Lighting Handbook-Reference Volume*, Illuminating Engineering Society of North America: New York, NY, 1984.
- [3] CIE, "The Basis of Physical Photometry," CIE Publ. No. 18.2 (TC-1.2), Central Bureau of the CIE, Kegelgasse 27, A-1030 Vienna, Austria, 1983.
- [4] Wyszecki, G., and W. S. Stiles. *Color Science: Concepts and Methods, Quantitative Data and Formulae*, 2nd ed., John Wiley & Sons: New York, NY, 1982, p. 256.
- [5] CIE, "International Lighting Vocabulary," CIE Publ. No. 17.4, Central Bureau of the CIE, Kegelgasse 27, A-1030 Vienna, Austria, 1987, p. 13.
- [6] CIE, "Principles of Light Measurements," 1st ed., CIE Publ. No. 18 (E-1.2), 1970.
- [7] CIPM Proces-Verbaux 40, 145 (1972).
- [8] Berman, S., "Energy Efficiency Consequences of Scotopic Sensitivity," *J. IES*, Winter 1992, pp. 3–14.
- [9] IES Color Committee, "Color and Illumination," IES DG-1-1990, Illuminating Engineering Society of North America, 345 East 47 St., New York, NY 10017.
- [10] Judd, D. B., "Report of U. S. Secretariat Committee on Colorimetry Artificial Daylight," CIE Proceedings, Stockholm, Bureau Central de la CIE, 1951, Vol. 1, Part 7, p. 11.
- [11] Kaiser, P. K., "Photopic and Mesopic Photometry: Yesterday, Today, and Tomorrow," in *Golden Jubilee of Colour in the CIE 1931–1981*, The Society of Dyers and Colourists, Bradford, 1981, pp. 28–52.
- [12] CIE, "CIE 1988 2° Spectral Luminous Efficiency Function for Photopic vision," CIE Technical Report Publ. No. CIE 86, 1st ed., CIE Central Bureau, Kegelgasse 27, A-1030 Vienna, Austria, 1990.
- [13] International Lighting Commission, "International Lighting Vocabulary," 3rd ed., Publ. CIE No.

- 17 (E-1.1.) 1970, Table 1, p. 51 and 4th Edition, Publ. CIE No. 17.4, 1987, Table on p. 13 referred to in definition 845-01-22.
- [14] Hunt, R.W.G. *Measuring Colour*, 2nd ed., Ellis Horwood, Chichester, Sussex, UK, 1991, Chapter 12. Available in the U.S. from Simon & Schuster, Prentice Hall Building, Englewood Cliff, NJ 07632.
 - [15] Shapely, Robert, "Retinal Regulation of Visual Contrast Sensitivity," *Optics and Photonics News*, Aug. 1991, p. 17.
 - [16] Heinemann, E. G. "Simultaneous Brightness Induction as a Function of Inducing and Test Field Luminances," *J. Exp. Psychol.*, Vol. 50, 1955, pp. 89–96.
 - [17] E. G. Heinemann. "Simultaneous Brightness Induction," in *Handbook of Sensory Physiology*, Vol. VII/4, D. Jameson and L. M. Hurvich, eds., Springer Verlag: New York, NY, 1972, pp. 146–169.
 - [18] R. Shapely and C. Enroth-Cugell. "Visual Adaptation and Retinal Gain Controls," in *Progress in Retinal Research*, Vol. 3, N. Osborne and G. Chader, eds., Pergamon, 1984, pp. 263–346.
 - [19] Williams, D. "Progress in Vision Research," *Optics and Photonics News*, Aug. 1991, pp. 8–9.
 - [20] CIE, "Discomfort Glare in the Interior Working Environment," Publ. CIE No. 55 (TC-3.4), Central Bureau of the CIE, Kegelgasse 27, A-1030 Vienna, Austria, 1983.
 - [21] Murdoch, J. B. *Illumination Engineering—From Edison's Lamp to the Laser*, MacMillan Publishing Co.: New York, NY, 1985.
 - [22] Kaufman, J. E., and J. F. Christensen, eds. *IES Lighting Ready Reference*, Illuminating Engineering Society of North America: New York, NY, 1989.

Chapter 3

Blackbodies and Other Sources

3.1 Blackbody Radiation

All bodies above a temperature of absolute zero *emit* radiation. The hotter they are, the more they emit. The constant agitation of the atoms and molecules that make up all objects involves accelerated motion of electrical charges (the electrons and protons of the constituent atoms). The fundamental laws of electricity and magnetism, as embodied in Maxwell's equations, predict that any accelerated motion of charges will produce radiation. The constant jostling of atoms and molecules in material substances above a temperature of absolute zero produces electromagnetic radiation over a broad range of wavelengths and frequencies. The total radiant flux emitted from the surface of an object at temperature T is expressed by the Stefan-Boltzmann law, in the form

$$M_{bb} = \sigma T^4 \quad (3.1)$$

where M_{bb} is the exitance of (irradiance leaving) the surface in a vacuum, σ is the Stefan-Boltzmann constant, and T is the temperature in degrees kelvin (number of degrees above absolute zero on the kelvin scale). The currently recommended [1] value for the Stefan-Boltzmann constant σ is $5.67032 \times 10^{-8} \text{ W} \cdot \text{m}^{-2} \cdot \text{K}^{-4}$. Using this value, the units for M_{bb} in (3.1) are $\text{W} \cdot \text{m}^{-2}$. From (3.1), we can easily calculate that a blackbody at 27°C ($27 + 273 = 300\text{K}$) emits at the rate of 460 W/m^2 . At 100°C, the rate jumps up to 1,097 W/m^2 .

Equation (3.1) is for a perfect or *full* emitter, which is called a *blackbody*. Imperfect emitters, called *greybodies* if their spectral shape matches that of a true blackbody and *nonblackbodies* if the shape does not match, are defined and discussed further in Chapter 6 on the optical properties of materials. The first part of the current chapter deals only with perfect or full emitters, blackbodies. Then other sources of radiation commonly used in optical systems of a variety of kinds are discussed.

A blackbody is defined as an ideal body that allows all incident radiation to pass into it (zero reflectance) and that absorbs internally all the incident radiation (zero transmittance). This must be true for all wavelengths and all angles of incidence. According to the definition of a blackbody, it is a perfect absorber, with absorptance of 1.0 at all wavelengths and directions. Due to the law of the conservation of energy, the sum of the reflectance R and absorptance A of an opaque surface must be unity, $A + R = 1.0$. Thus, if a blackbody has an absorptance of 1.0, its reflectance must be zero. Accordingly, a perfect blackbody at room temperature would appear totally black to the eye, hence the origin of the name.

Only a few surfaces, such as carbon black, carborundum, and gold black, approach a blackbody in these optical properties. Most surfaces are not perfect absorbers and many have reflectances that are different for different wavelengths. They are therefore called spectrally selective surfaces, because their optical properties are selectively different for different portions of the spectrum.

The radiation emitted by a surface is in general distributed over a range of angles filling the hemisphere and over a range of wavelengths. As shown by Grum [2], the angular distribution of radiance from a blackbody is constant; that is, the radiance is independent of direction; it is a Lambertian surface. Specifically, this means that $L_\lambda(\theta, \phi) = L_\lambda(0, 0) = L_\lambda$. Thus, the relationship between the spectral radiance $L_{\text{bb}\lambda}$ and spectral exitance $M_{\text{bb}\lambda}$ of a blackbody is given by (1.19), repeated here as

$$M_{\text{bb}\lambda} = \pi L_{\text{bb}\lambda} \quad (3.2)$$

If $L_{\text{bb}\lambda}$ is in $\text{W}\cdot\text{m}^{-2}\cdot\text{nm}^{-1}\cdot\text{sr}^{-1}$, then the units of $M_{\text{bb}\lambda}$ will be $\text{W}\cdot\text{m}^{-2}\cdot\text{nm}^{-1}$.

3.2 PLANCK'S LAW

As the temperature changes, the spectral distribution of the radiation emitted by a blackbody shifts. In 1901, Max Planck made a radical new assumption—that radiant energy is quantized—and used this assumption to derive an equation for the spectral radiant energy density in a cavity at thermal equilibrium (a good theoretical approximation of a blackbody) [3]. By assuming a small opening in the side of the cavity and examining the spectral distribution of the emerging radiation, one can derive an equation for the spectral distribution of radiation emitted by a blackbody [4]. The equation, now called Planck's blackbody spectral radiation law, accurately predicts the spectral radiance of blackbodies in a vacuum at any temperature [5–8]. Using the notation of this text the equation is

$$L_{\text{bb}\lambda} = \frac{2hc^2}{\lambda^2(e^{hc/\lambda kT} - 1)} \quad (3.3)$$

From [1], $h = 6.626176 \times 10^{-34} \text{ J}\cdot\text{s}$ is Planck's constant, $c = 2.9979246 \times 10^8 \text{ m/s}$ is the speed of light in a vacuum, and $k = 1.380662 \times 10^{-23} \text{ J}\cdot\text{K}^{-1}$ is Boltzmann's constant. Using these values of the constants, the units of $L_{bb\lambda}$ will be $\text{W}\cdot\text{m}^{-2}\cdot\mu\text{m}^{-1}\cdot\text{sr}^{-1}$.

From (3.2), the spectral exitance $M_{bb\lambda}$ of a blackbody at temperature T is just the spectral radiance $L_{bb\lambda}$ given in (3.3) multiplied by π

$$M_{bb\lambda} = \frac{2\pi hc^2}{\lambda^5(e^{hc/\lambda kT} - 1)} \quad (3.4)$$

It is common to express (3.4) in the form

$$M_{bb\lambda} = \frac{C_1}{\lambda^5(e^{C_2/\lambda T} - 1)} \quad (3.5)$$

with constants C_1 and C_2 tabulated for different combinations of units [9]. Using common metric units, these constants have the following values [1]:

$$\begin{aligned} C_1 &= 2\pi hc^2 = 3.741832 \times 10^8 \text{ W}\cdot\mu\text{m}^4\cdot\text{m}^{-2} \\ &= 3.741832 \times 10^4 \text{ W}\cdot\mu\text{m}^4\cdot\text{cm}^{-2} \\ &= 3.741832 \times 10^{-4} \text{ W}\cdot\text{nm}^4\cdot\text{m}^{-2} \\ &= 3.741832 \times 10^{-16} \text{ W}\cdot\text{m}^2 \end{aligned}$$

$$\begin{aligned} C_2 &= hc/k = 14387.86 \mu\text{m}\cdot\text{K} \\ &= 0.01438786 \text{ m}\cdot\text{K} \end{aligned}$$

(Note: Some authors use c^2 for C_1 and then give $2\pi hc^2$ another name, such as C_{1p} . The constant C_{1L} is sometimes used for C_1/π , so that (3.5) can be used for spectral radiance with C_1 replaced by C_{1L} .)

Substituting the values for C_1 and C_2 into (3.5) and rearranging terms yields the following form of this equation, when the wavelength is given in μm , and the temperature in degrees kelvin:

$$M_{bb\lambda} = \frac{3.741 \times 10^8}{\lambda^5[\exp(14388/(\lambda T)) - 1]} \quad (3.6)$$

It is easy to determine blackbody spectral exitance values from this equation by substituting in values for wavelength and temperature and then performing the indicated operations. This is easily done on a scientific calculator or computer. To get the spectral radiance $L_{bb\lambda}$, just divide (3.6) by π .

Many textbooks on radiation, especially those on infrared radiation, contain tables of values for the spectral exitance of blackbodies at different temperatures. See, for example, the texts by Wyszecki and Stiles [10], Conn and Avery [11],

Pivovonsky and Nagel [12], Hahn and others [13], and Levi [14]. Because of the presence of λT in the denominator of (3.6), some authors calculate and publish $M_{bb\lambda}/T^5$ for various values of λT to save space [15]. One can easily recover the original function for any temperature T by multiplying the tabulated values by T^5 and dividing the λT product by T to recover λ . In the modern era, with inexpensive programmable calculators and desktop computers widely available, it seems easier to calculate (3.6) directly rather than go through the tedium of copying (and possibly mistakenly modifying) numbers from published tables. However, the tabulated values are available for those who need them and as a check on computer calculations.

Figure 3.1 shows a family of blackbody radiation spectra calculated with (3.6) for different temperatures. A log-log scale is used to cover the wide range of values produced with (3.6). Figure 3.2 shows a plot on a linear scale of the shape of the 6,000K blackbody spectrum. (Equation (3.6) gives spectral exitance in $\text{W} \cdot \text{m}^{-2} \cdot \mu\text{m}^{-1}$. The plots in Figures 3.1 through 3.4 were obtained by scaling the output from (3.6) to give exitance in $\text{W} \cdot \text{m}^{-2} \cdot \text{nm}^{-1}$.)

The 6,000K blackbody curve is compared with the extraterrestrial solar spectrum in Figure 3.3, where the blackbody curve has been scaled vertically to peak near the height of the solar spectrum. The sun is not a true blackbody radiator, so the curves do not quite match in shape. One reason is that the sun is surrounded

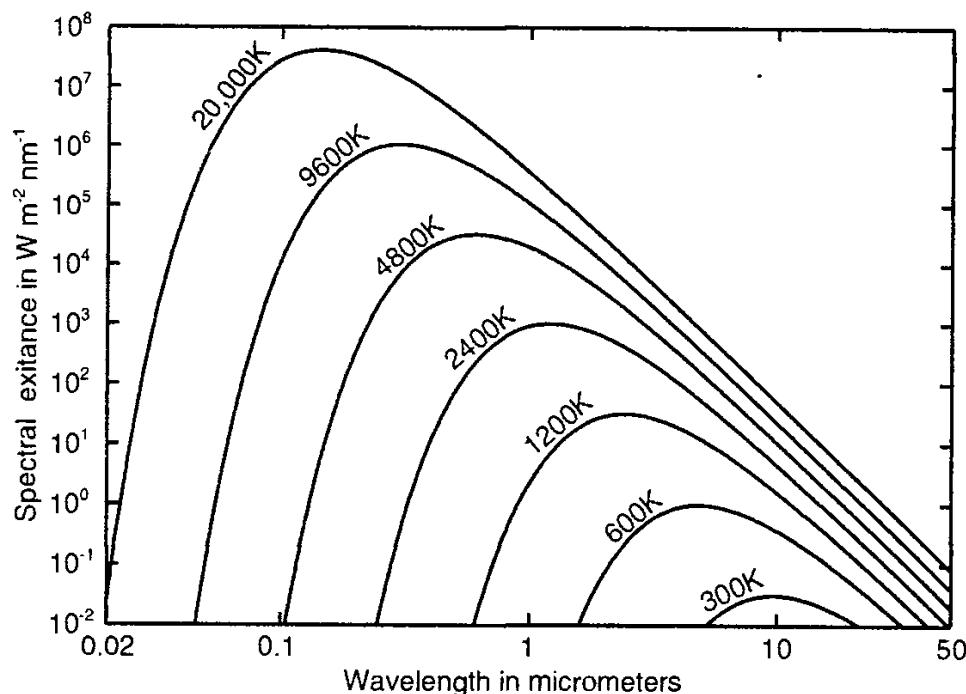


Figure 3.1 Exitance spectra for blackbodies at various temperatures from 300 to 20,000K, calculated with (3.6).

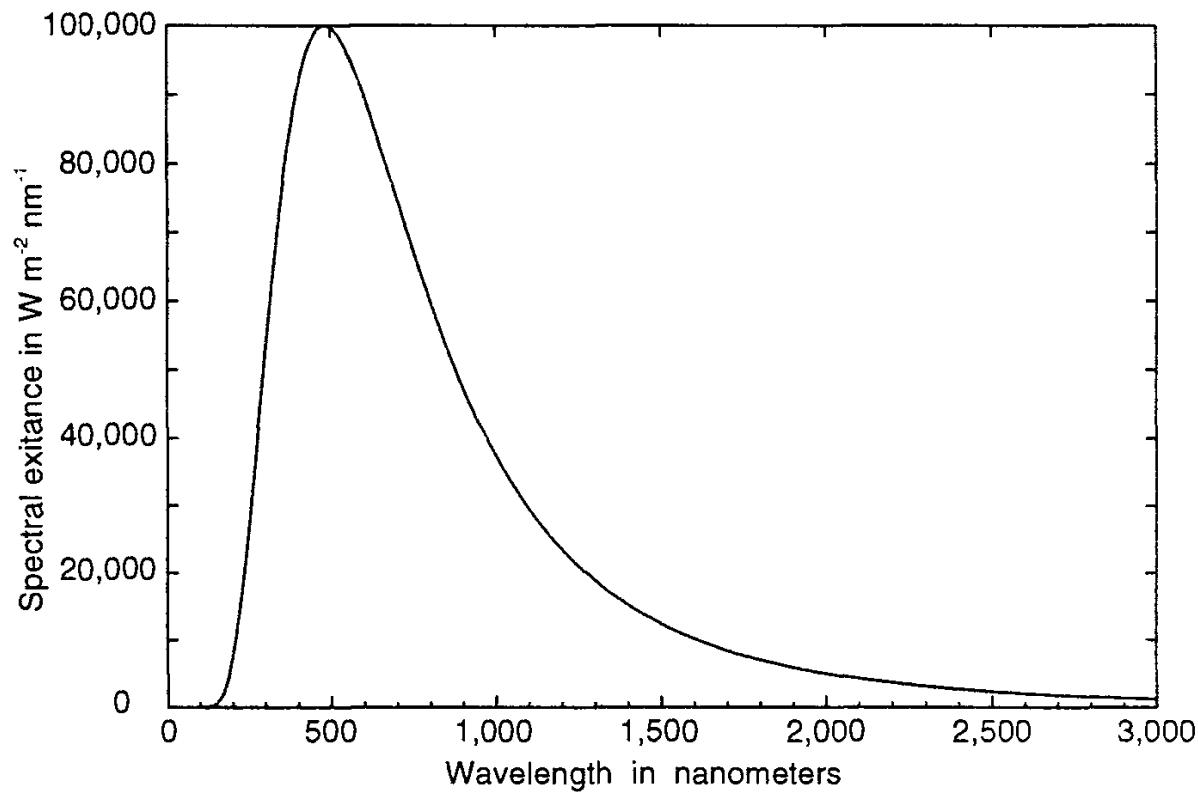


Figure 3.2 The exitance spectrum of a blackbody at 6,000K, calculated with (3.6).

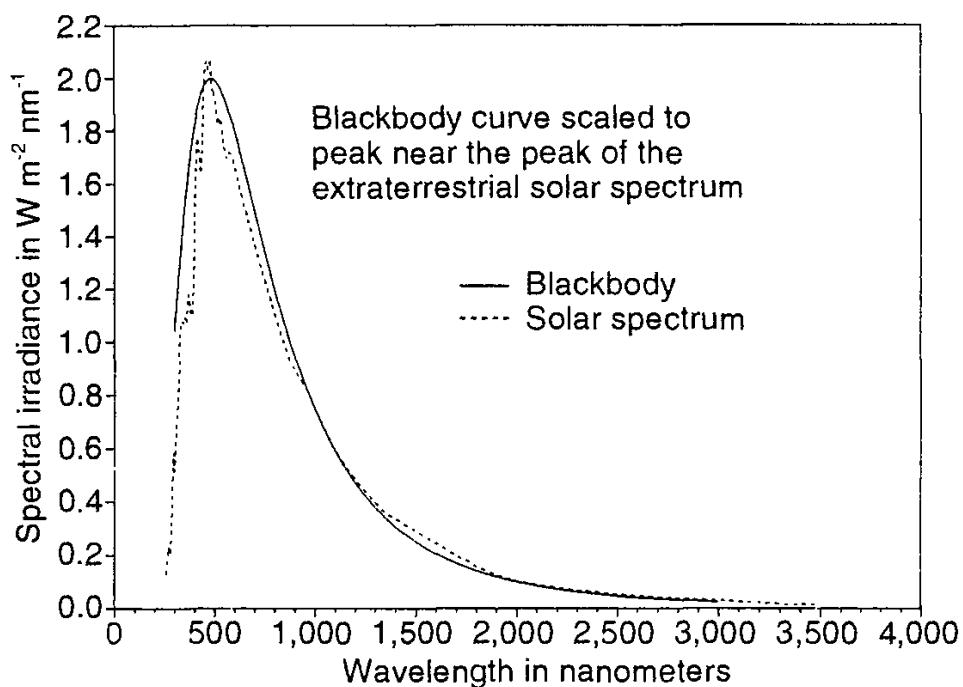


Figure 3.3 Extraterrestrial solar spectrum and scaled 6,000K blackbody spectrum.

by a large gaseous region or atmosphere whose atoms and molecules are cooler than those of the sun's surface and these atoms and molecules absorb radiation selectively in different wavelength regions.

It is clear to the human senses that objects as hot as the sun emit radiation, because we are able to see the visible portion of their spectra with our eyes and feel the radiant heat on our skins.

Other bodies also emit radiation. Figure 3.4 shows the emission spectrum of a blackbody at only 23.9°C or 297K (75°F), room temperature. You can see that the spectrum is shifted considerably toward the higher wavelength. There is no visible radiation produced. Also the total magnitude of the emission is much lower and not easy to feel on our skins, which are at a higher temperature.

3.3 WIEN DISPLACEMENT LAW

If one differentiates Planck's formula, (3.4), and sets the result equal to zero, a solution of the resulting equation for wavelength yields a simple relationship between the wavelength λ_m where the Planck radiation formula has its maximum

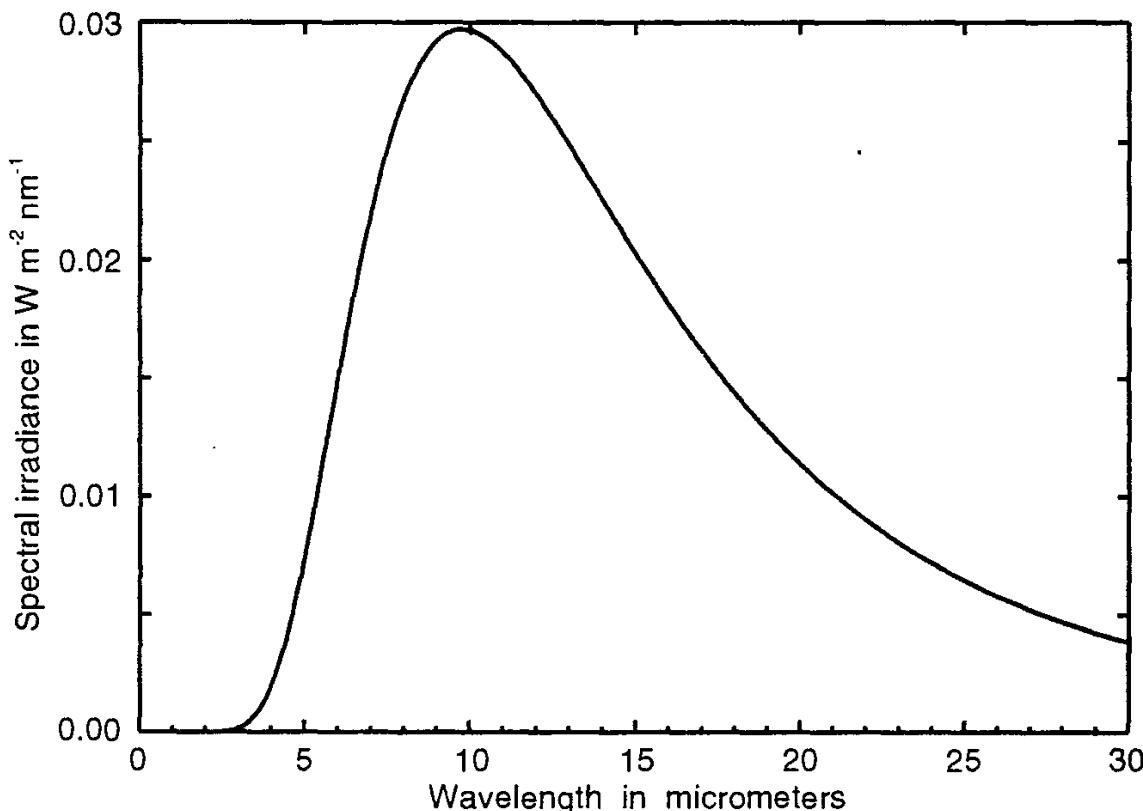


Figure 3.4 Exitance spectrum of a 23.9°C (75°F) blackbody.

value M_{λ_m} (or L_{λ_m}) and the temperature T of the blackbody. The resulting relationship is called Wien's displacement law and is given by

$$\lambda_m T = 2,897.8 \text{ } \mu\text{m} \cdot \text{K} \quad (3.7)$$

It can also be shown by integrating (3.4) from $\lambda = 0$ to $\lambda = \lambda_m$, where λ_m is the wavelength of maximum emission given by (3.7), that exactly one-fourth of the total exitance of a blackbody lies in the wavelength range from 0 to λ_m , below the emission peak at this wavelength.

3.4 LUMINOUS EFFICACY OF BLACKBODY RADIATION

Substituting (3.4) for the hemispherical spectral exitance of a blackbody into (2.6) and (2.10), and using the results in the irradiance version of (2.12) (replacing each Q with E), for each of several different temperatures T , one can calculate the radiation luminous efficacy K_{rb} of blackbody radiation as a function of temperature. Some numerical results are given in Table 3.1 and others are plotted in Figure 3.5,

Table 3.1
Luminous Efficacy Values

Temperature in Degrees Kelvin	Luminous Efficacy in lm/W
500	7.6×10^{-13}
1,000	2.0×10^{-4}
1,500	0.103
2,000	1.83
2,500	8.71
3,000	21.97
4,000	56.125
5,000	81.75
6,000	92.9
7,000	92.8
8,000	87.3
9,000	79.2
10,000	70.6
15,000	37.1
20,000	20.4
30,000	7.8
40,000	3.7
50,000	2.0

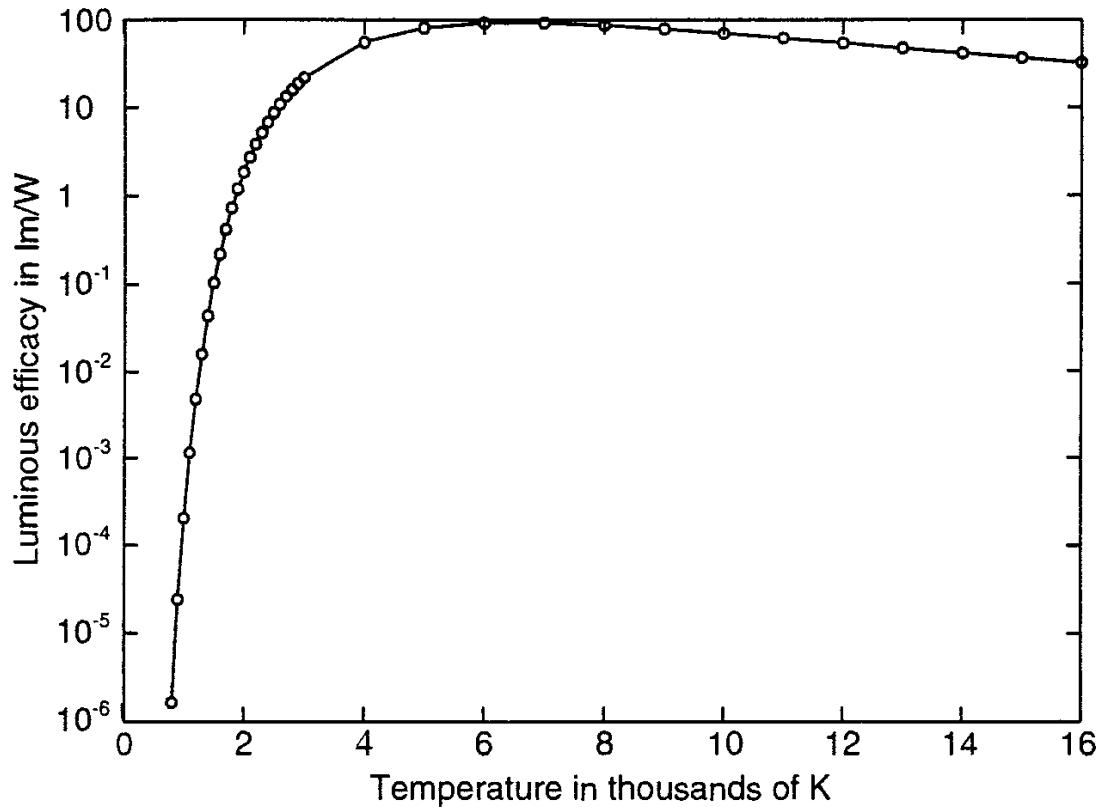


Figure 3.5 The luminous efficacy of blackbody radiation on a logarithmic scale, as a function of temperature.

where it can be seen that, as expected, the luminous efficacy drops very rapidly as the body cools down from white hot temperatures.

It would be a useful exercise for the reader to calculate the data presented in Table 3.1 following the procedure outlined in the previous paragraph. In preparing Table 3.1, (3.4) was evaluated at 10-nm intervals from 200 nm to 40,000 nm for each of the kelvin temperatures shown in the table up to 6,000 and from 10 nm to 40,000 nm for the remaining temperatures. The resulting spectra were then each integrated according to (2.4) and (2.8) and the results divided as indicated by (2.12) to yield the luminous efficacies shown in Table 3.1. For Figure 3.5, the calculations were performed from 10 to 40,000 nm in 10-nm steps over the wavelength range, and over the temperature range from 800 to 3,000K in 100K steps and from 4,000 to 50,000K in 1,000K steps. These data were only plotted to 16K in Figure 3.5. Figure 3.6 was plotted using this same data to determine the shape of the luminous efficacy curve on a linear scale on either side of the peak. According to Pivovonsky and Nagel (as reported by Driscoll [5]), the peak value is close to 95 lm/W at a temperature near 6,600K, in agreement with the curves shown in Figures 3.5 and 3.6.

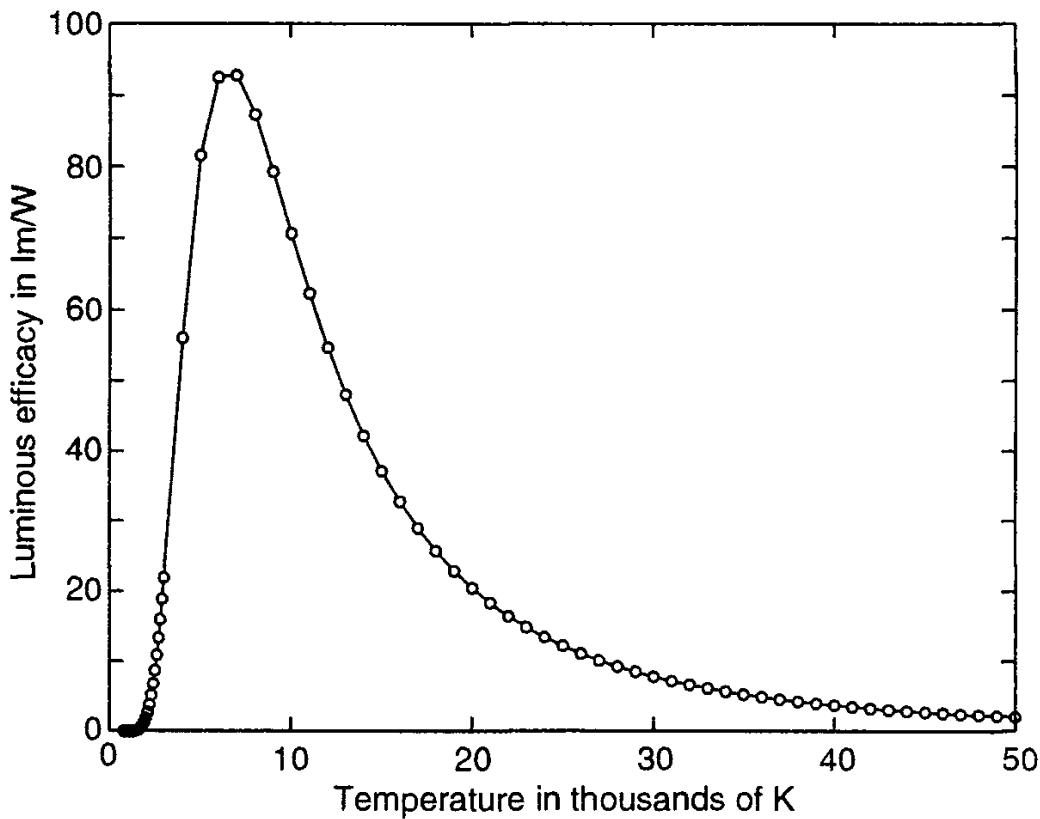


Figure 3.6 The luminous efficacy of blackbody radiation on a linear scale, as a function of temperature.

3.5 COLOR AND DISTRIBUTION TEMPERATURES

The color of the radiation emitted by a blackbody in the visible portion of the spectrum varies as the temperature is changed (over the range of temperatures where there is significant emission in the visible). It can be useful to think of matching the color of some general radiation source to one of these blackbody colors. This idea leads to the concept of *color temperature*, the temperature at which the color of the general source of radiation appears to be the same as that of a blackbody. If the color of the source does not quite match that of any blackbody, then a modification of the color temperature, called the *correlated color temperature* is available. That temperature which produces light from a blackbody having a color closest to the color of the source of interest is called the *correlated color temperature of the source*. The spectral distribution can be very unlike that of a blackbody, but the resulting color must be near to that of a blackbody at some temperature. Methods of obtaining this color temperature are described in Chapter 11.

The concept of color temperature is closely related to another concept called *distribution temperature*, which relates non-blackbody emitters to the spectral emis-

sion of a blackbody. The concept is applied only to radiators whose spectral shape is close to that of a blackbody at some temperature.

The radiator does not have to be at the same temperature as the blackbody, but the shape of its emission spectrum must be close (typically within 5% at all wavelengths) to that of a blackbody at some temperature. The distribution temperature T_d for such a source is that value of T which minimizes the following integral over variations in the arbitrary constant a and the temperature T :

$$\int_0^\infty \left(1 - \frac{M'_{e,\lambda}}{a M_{bb\lambda}(\lambda, T)} \right)^2 d\lambda \quad (3.8)$$

where $M'_{e,\lambda}$ is the spectral exitance of the test source and $M_{bb\lambda}(\lambda, T)$ is obtained from Planck's formula, given in (3.4). When the spectral distribution $M'_{e,\lambda}$ from the source differs a lot from that of a blackbody, the color temperature or the correlated color temperature, described above and detailed in Chapter 11, is generally used instead. A procedure for calculating the distribution temperature of measured spectra is described by Andreic [7].

3.6 EMISSION INTO AN IMPERFECT VACUUM

The refractive index of a perfect vacuum is 1.0, because the refractive index is defined as the ratio of the speed of light in a perfect vacuum to its speed in the medium. If a blackbody is emitting radiation into a gas, liquid, or solid having refractive index n , then the above equations for the Stefan-Boltzmann law, Planck's radiation Law, and the Wien displacement law have to be modified. We must replace the wavelength and speed of light in a vacuum with their values in the medium. The Wien displacement law becomes

$$n\lambda_m T = 2,897.8 \mu\text{m}\cdot\text{K} \quad (3.9)$$

where λ_m is the wavelength in the medium. The Stefan-Boltzmann law becomes

$$M_{bb} = n^2 \sigma T^4 \quad (3.10)$$

Planck's radiation law can be modified by representing wavelength in a vacuum as λ_o and the speed of light in a vacuum as c_o and replacing these with $n\lambda_m$ and nc_m , respectively, where λ_m is the wavelength in the medium and c_m is the speed of light in the medium.

The emission from a blackbody in contact with glass of refractive index $n = 1.5$ according to (3.10) is 2.25 times its value in a vacuum. Since the refractive index of air at atmospheric pressure and room temperature is about 1.00028 and

the square of this is 1.00056, the exitance of a blackbody into air is about the same as it is for emission into a vacuum.

3.6 RADIATION EXCHANGE

The emission of blackbody radiation into a medium, discussed in the previous section, brings up a host of important additional topics that will be dealt with in subsequent chapters. For example, the optical properties of the medium into which the radiation is emitted are very important to the subsequent fate of the radiation. Optical properties are discussed in Chapter 6. Furthermore, the presence of another medium adjacent to or in contact with the emitting source can affect the properties of the source and can therefore alter the amount and spectral distribution of the emission. For example, if an ideal blackbody is in contact with another blackbody (or nonblackbody) at a temperature that is different from that of the first blackbody, heat will tend to flow by conduction from the hotter to the colder body, possibly altering the temperature of the original blackbody and hence its emission characteristics. Even if the bodies are not in contact, radiant (and possibly convective) heat can flow between them, again altering their temperatures under transient conditions and changing the nature of the emissions.

This latter case is one of great interest in real applications of radiometry that involve the transfer of substantial quantities of radiant energy between different objects separated by a transparent medium. The importance of this consideration is illustrated in Example Problem 3.4. It is frequently the *net* radiant energy exchange between two objects that is of most importance. As is shown in Chapter 4, if one knows the temperatures of two blackbodies in radiant “view” of each other, as well as the geometrical relationships between them, it is possible to calculate the net radiation exchange between them if their temperatures are known. This problem is of very great importance in applications involving long-wavelength radiometry, where every room-temperature object in view of the detector is a source of radiant flux and can introduce errors into attempts to measure other sources having similar temperatures. Such issues are discussed in Chapter 4.

3.8 EXPERIMENTAL APPROXIMATION OF A BLACKBODY

The angular and spectral characteristics of a blackbody can be approximated with an arrangement similar to the one shown in Figure 3.7. A metal cylinder is hollowed out to form a cavity with a small opening in one end. At the opposite end is placed a conical shaped “light trap,” whose purpose is to multiply reflect incoming rays, with maximum absorption at each reflection, in such a manner that a very large number of reflections must take place before any incident ray can emerge back out the opening. With the absorption high on each reflection, a vanishingly small

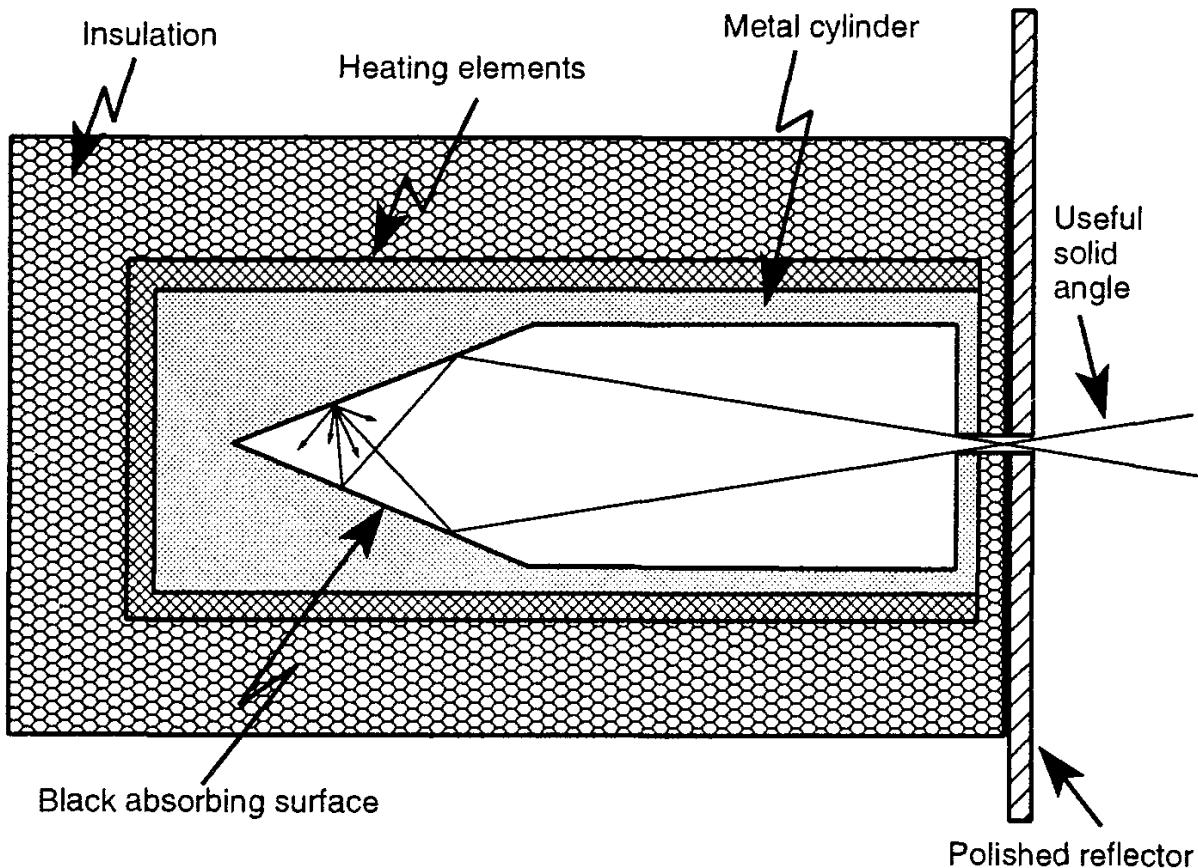


Figure 3.7 Schematic diagram of an experimental realization of a blackbody radiator.

fraction of incident flux, after being multiply reflected and scattered, would emerge from the opening. In consequence, only a very tiny portion of the radiation passing into the cavity through the opening can be reflected back out of the cavity.

The temperature of the entire cavity is controlled by heating elements and thick outside insulation so that all surfaces of the interior are at precisely the same (known) temperature and any radiation escaping from the cavity will be that emitted from the surfaces within the cavity. The emerging radiation will be rendered very nearly isotropic by the multiple reflections taking place inside (at least over the useful solid angle indicated in Figure 3.7, for which the apparatus is designed).

3.9 OTHER REAL SOURCES

Since it is not practical to use close approximations of blackbodies in real radiometric and photometric systems, many of the latter are designed to measure the radiation from a variety of different sources with varying spectral distributions. It will be helpful, therefore, at this point to describe the spectral distributions of a few common additional sources of radiation that radiometers and photometers are designed to measure.

Sources such as sunlight, incandescent lamps, and fluorescent lamps are mentioned in Chapters 1 and 2, but the spectral distributions from these sources are not described explicitly. Spectral distributions are probably the most important characteristics of sources that must be considered in the design of radiometric systems to measure all or portions of those distributions. The matching of a proper detector/filter combination to a given radiation source to be measured is one of the most important tasks facing the designer of a radiometric system. Chapter 7 deals with detectors of radiation.

Following its passage through the atmosphere, direct beam solar radiation, whose extraterrestrial spectrum is shown in Figure 3.3, exhibits the altered spectral distribution shown in Figure 3.8. The alterations are the result of absorption by various gaseous constituents of the atmosphere, the most noticeable of which are water vapor and CO₂.

The spectral distribution of blue sky radiation is similar to that shown in Figure 3.8, but the shape is skewed by what is called *Rayleigh scattering*, the scattering of radiation by molecular-sized particles in the atmosphere. Rayleigh scattering is proportional to the inverse fourth power of the wavelength. Thus, blue light is scattered more prominently than red light and this phenomenon is responsible for the blue appearance of sky light. (Conversely, this same removal of some blue light shifts the apparent color of beam sunlight toward the red end of the spectrum, a phenomenon primarily responsible for the orange-red appearance of

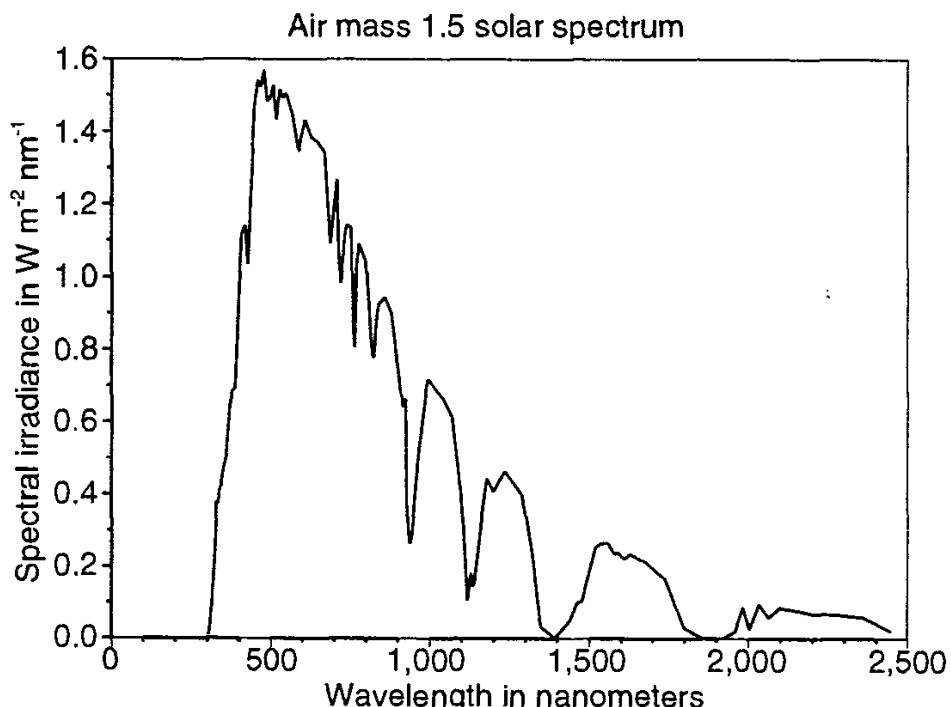


Figure 3.8 Spectral irradiance of terrestrial direct beam solar radiation.

the sun at sunrise and sunset.) The spectral distribution of daylight is important in the field of colorimetry, discussed in Chapter 11, and for many other applications, including the daylight illumination of building interiors [16–18]. The development of filtered lamps whose output spectral distributions simulate the shape of terrestrial solar radiation and/or daylight spectra is a problem of significance, but also of some difficulty due to the changing nature of these natural sources and to the difficulty of finding artificial sources that mimic them [19, 20].

Modern tungsten halogen lamps in quartz envelopes produce output spectra that are similar in shape to those of blackbody distributions. This is discussed in more detail in Section 9.11 on calibration of radiometers and photometers. A representative spectral distribution is shown in Figure 3.9. This lamp covers a wide spectral range, including the near UV, the visible, and much of the infrared portion of the spectrum.

Although quartz halogen lamps produce usable outputs in the ultraviolet region, at least down to 200 nm, the output at these short wavelengths is quite low and declines rapidly with decreasing wavelength. Deuterium arc lamps are available as sources of ultraviolet radiation that overcome the limitations of quartz halogen lamps in this spectral region, and they do so with little output above a wavelength of 500 nm except for a strong but narrow emission line at about 660 nm. More information is provided on Deuterium lamps in Section 9.11. The spectral irradiance from a deuterium lamp is plotted in Figure 3.10.

Tungsten halogen lamps have their strongest emission in the infrared, and the shape of the spectrum departs substantially from that of sunlight. As shown in Figure 3.11, xenon arc lamps, in contrast, have a more balanced output over the visible but they exhibit strong spectral “spikes” that pose problems in some applications. Short arc lamps, such as those using xenon gas, are the brightest manufactured sources, with the exception of lasers. Because of the nature of the arc discharges, these lamps emit a continuum of output over wavelengths covering the ultraviolet and visible portions of the spectrum. If the spectral lines are not a problem, xenon lamps produce spectra relatively similar in broad shape to that of direct beam solar radiation.

The spectral irradiance outputs of the three sources just mentioned cover the near UV, the visible, and the near IR. They are plotted along with the spectrum of a 50W mercury-vapor arc lamp in Figure 3.12. Mercury lamps emit strong UV and visible radiation, with strong spectral lines in the ultraviolet superimposed over continuous spectra. The arc region between the anode and cathode electrodes is so small that these lamps are effectively point sources, making them ideal for use with beam shaping and steering optics.

From (1.37) and (1.38), it is seen that photons in the ultraviolet range are more energetic than their longer wavelength cousins. Ultraviolet radiation has a number of health hazards. UV-B and UV-C cause sunburn (erythema) and pigmentation (tanning). UV-B produces vitamin D₃. Long-term exposure results in

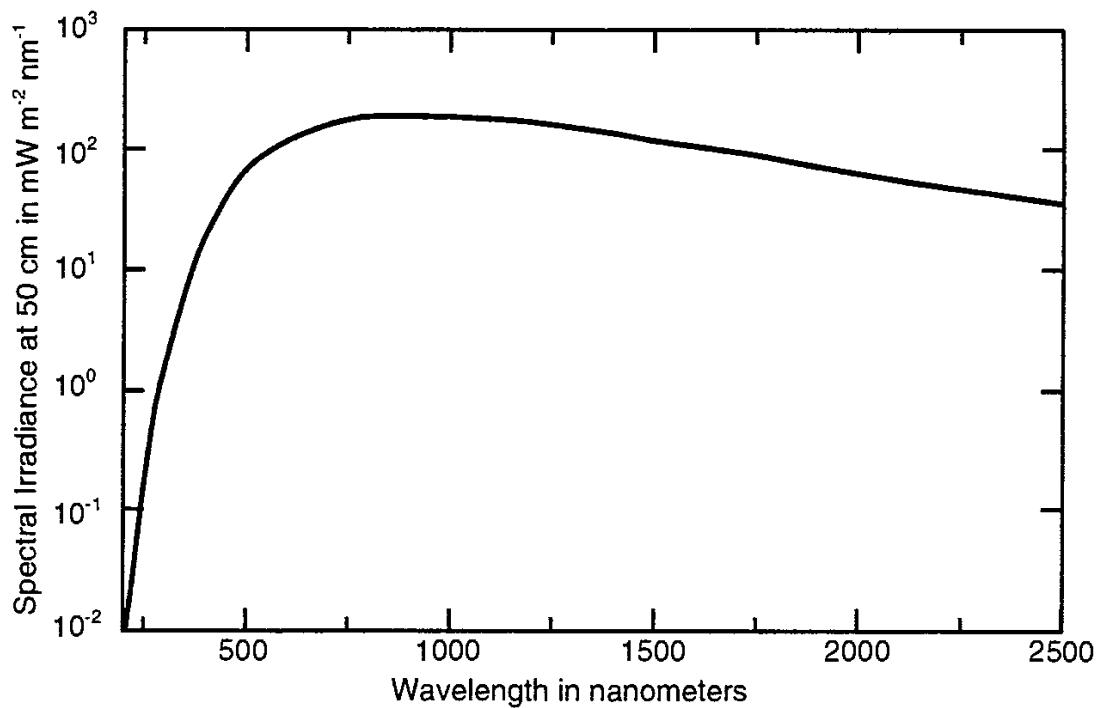


Figure 3.9 Spectral irradiance from a quartz halogen lamp 50 cm from the filament. (Courtesy Oriel Corporation.)

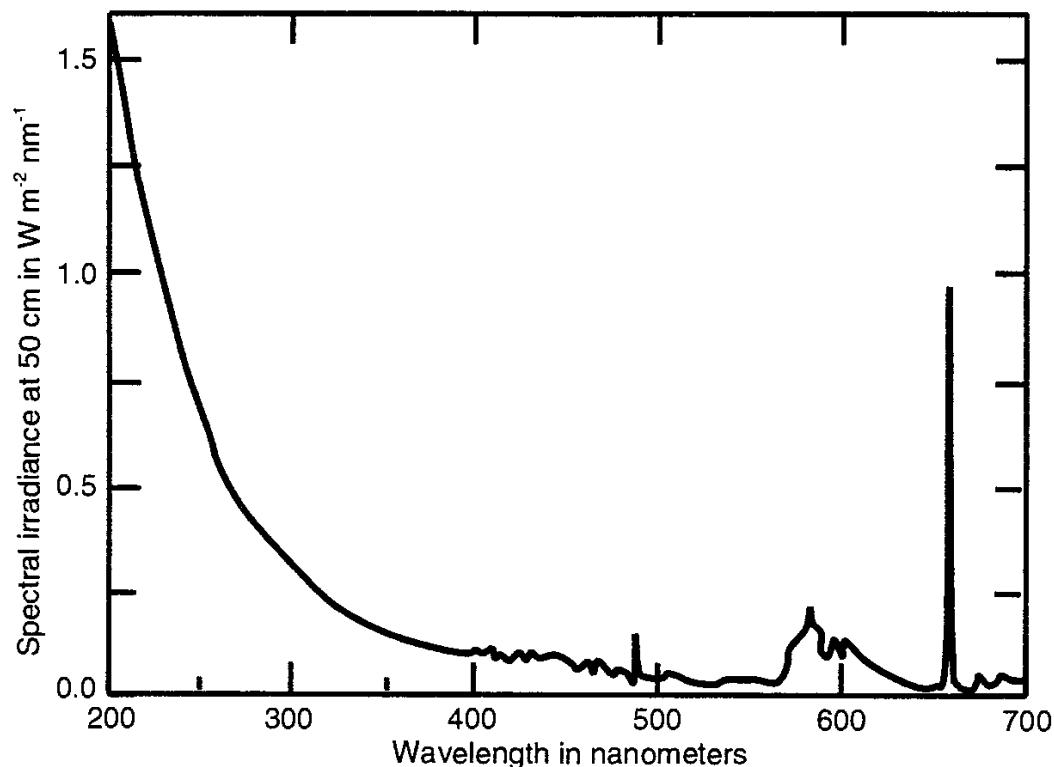


Figure 3.10 Spectral irradiance of a deuterium lamp. (Courtesy Oriel Corporation.)

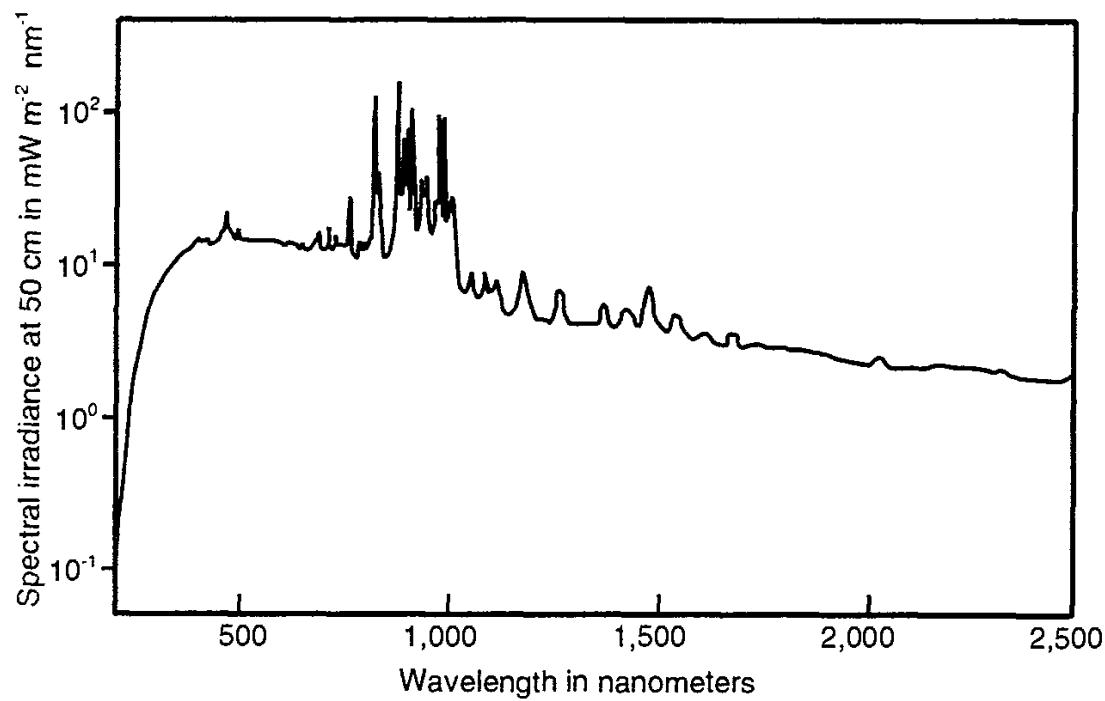


Figure 3.11 Spectral irradiance from a xenon gaseous discharge lamp. (Courtesy Oriel Corporation.)

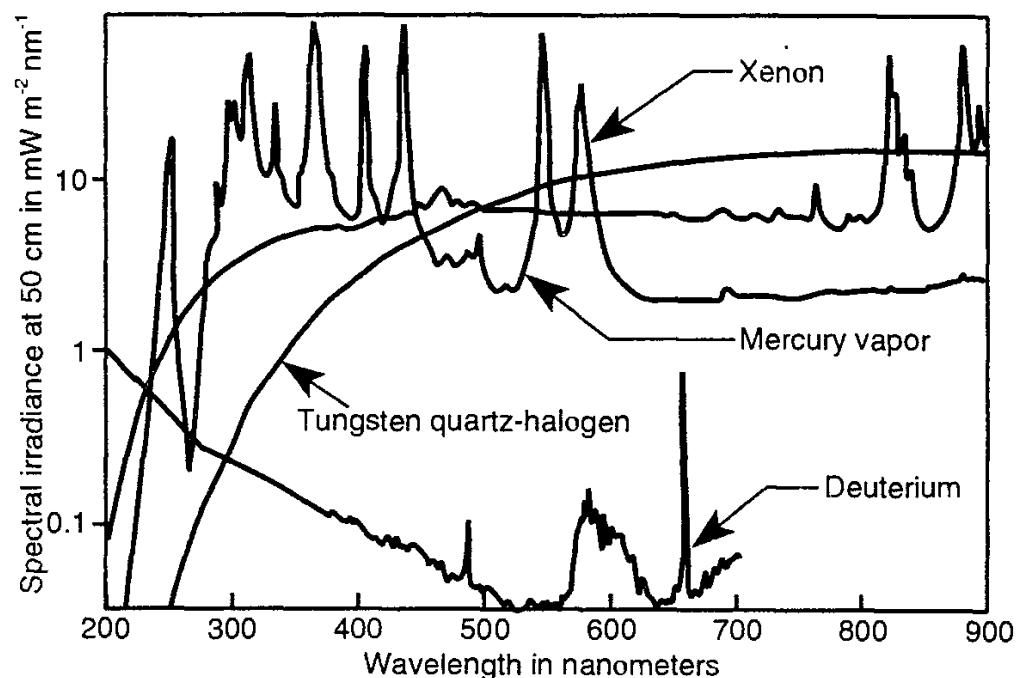


Figure 3.12 Spectral irradiance distributions for four laboratory sources of radiation. (Courtesy Oriel Corporation.)

loss of skin elasticity. There is a well-established link between exposure to wavelengths below 320 nm and skin cancer and concern is growing over possible longer-wavelength involvement. In the eye, corneal absorption of UV-B and UV-C may cause conjunctivitis and a corneal inflammation called photokeratitis. Prolonged exposure to longer ultraviolet radiation, particularly UV-A, may cause cataracts to form in the lens of the eye, and there are other possible eye effects from UV-A exposure [21]. In addition to these health effects, ultraviolet radiation produces degradation in the optical and physical properties of many materials. Care must therefore be taken in the use of laboratory sources of UV radiation and in prolonged exposure to UV radiation from the sun.

Attention is now turned to infrared sources. Although tungsten halogen lamps have substantial output in the near infrared, for sources with better coverage of IR-A and IR-B, one turns to a different technique. Typical outputs from several infrared laboratory sources are shown in Figure 3.13. The sources are basically electrical resistance heaters, ceramic and other substances that carry electrical current, which become hot due to ohmic heating, and which emit broadband infrared radiation with relatively high radiance.

In addition to these relatively broadband sources of radiation, there are numerous sources that emit radiation over more restricted spectral ranges. The light-emitting diode (LED) is one example. An LED is made of a semiconductor

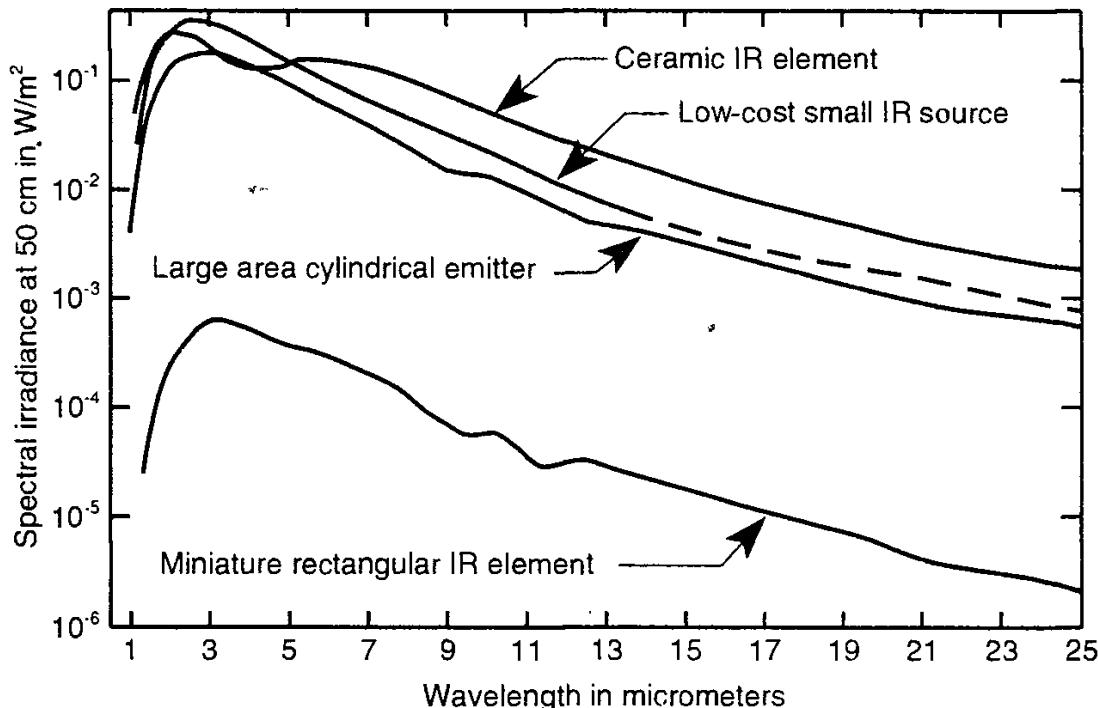


Figure 3.13 Spectral irradiance distributions from several sources of infrared radiation. (Courtesy Oriel Corporation.)

diode with a P-N junction designed so that electrical current through the junction in the forward bias direction produces the emission of optical radiation. The spectral range of emission is limited, but not so much to be considered fully monochromatic. A sample LED spectrum is shown in Figure 3.14.

An interesting characteristic of LEDs is that many can also be used as detectors of radiation. Mims has taken advantage of this fact to design a narrow-wavelength band radiometer to detect solar radiation without the need for auxiliary spectral filters [22]. The Mims radiometer is calibrated to determine stratospheric ozone concentrations from measurements of solar irradiance at six wavelengths from 555 to 940 nm. More information on semiconductor diodes is presented in Chapter 7 on detectors.

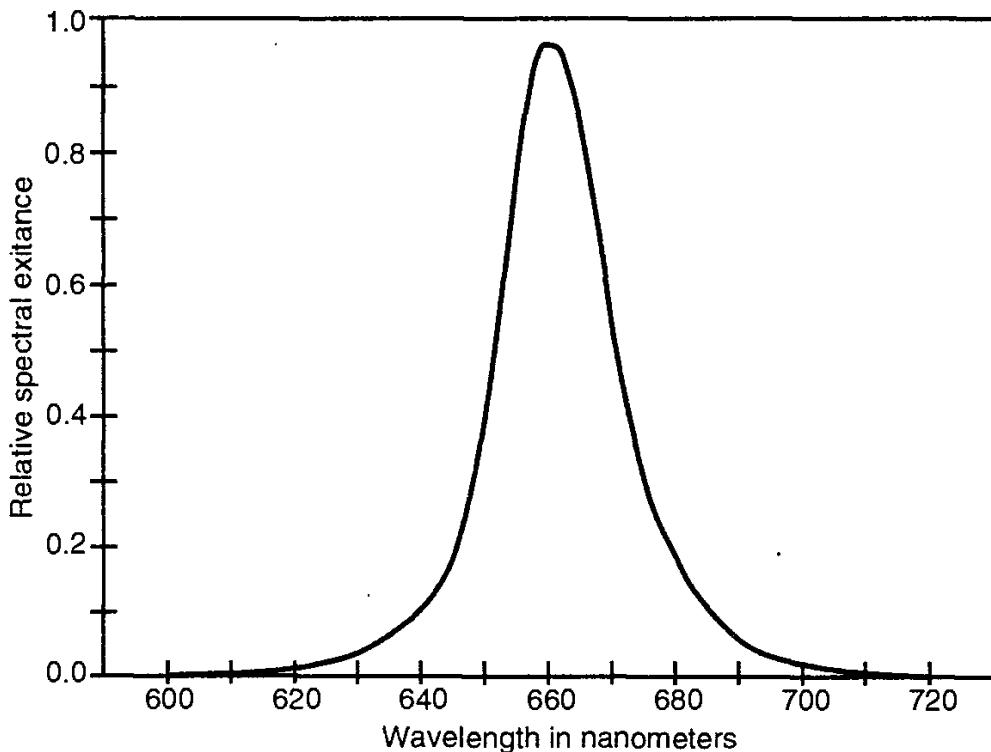


Figure 3.14 Relative spectral exitance of a representative light-emitting diode.

Laser sources deserve special mention. LASER stands for light amplification by stimulated emission of radiation. One important characteristic of these sources is their extremely narrow spectral output distribution, making them essentially monochromatic sources. Another is the very high spectral irradiances which they can produce. Finally, a consequence of the narrow spectral width of their emissions is a phenomenon called *coherence* in time and space whereby the phases of the oscillations in the electric and magnetic field strength vectors discussed in Section

1.7 are preserved over time and space. These important characteristics cannot be easily discussed in a few paragraphs and the reader is referred to any modern textbook on lasers or optics for further elucidation of the amazing properties of laser sources. One consequence of the way most lasers are fabricated is that their output beams are remarkably well collimated, with laser pulses having been successfully sent to the moon and back, with a beam spread sufficiently small to permit detection of the returned pulse by a sensitive detector placed at the focus of a powerful telescope.

One problem with highly coherent sources in radiometry and photometry is that not all of the relationships developed in this book governing flux levels can be considered strictly correct. The reason for this is the possibility with coherent monochromatic sources of *interference*. The superposition of two or more coherent monochromatic beams will produce a combined irradiance at a point that is not simply a sum of the irradiances of the two beams at that point. The combined irradiance in this case depends strongly upon the phase difference between the two beams. Laser illumination of a diffuse surface, for example, produces what is essentially a source having the equivalent of a very large number of coherent point sources, each with a different phase relationship, resulting in the familiar phenomenon called *speckle*. A person looking at an otherwise smooth and completely homogeneous surface illuminated by a single laser beam will perceive a speckled appearance in the light coming from the surface. The speckles are randomly distributed points in space where the sum total of all the light from the surface interferes *constructively* to produce bright spots as well as other points where *destructive* interference produces dark spots. For large sources and large detectors, the total flux transferred from source to detector will be the average over a large number of speckles and will therefore be approximated by the relationships developed in this chapter. Some of the problems involved with attempting classical radiometry and photometry calculations with radiation sources exhibiting strong wave-like effects, such as coherency, were discussed by Carter in connection with attempts to properly represent the exitance of sources using wave theory [23].

There are more sources of optical radiation, including fluorescent lamps, pulsed sources, electroluminescent lamps, and carbon arc lamps.

For more information concerning sources of electromagnetic radiation, the reader is referred to an excellent chapter on sources by Carlson and Clark [24], another by Grum and Becherer [28], an Oriel Corporation catalog [25], and the *IES Lighting Handbook Reference Volume* [26], all of which contain excellent discussions of the characteristics of a wide variety of light sources.

EXAMPLE PROBLEM 3.1

Problem: If the sun emitted as a perfect blackbody at 6,000K, what would be the solar radiance? What would be the spectral radiance at the peak of the human eye response, 555 nm?

Solution: Since a blackbody is Lambertian, the radiance is the same in all directions and we can use the radiance version of (3.1):

$$L_{bb} = \frac{\sigma T^4}{\pi} \quad (3.11)$$

We assume that for all practical purposes the sun is a circular disk parallel to the surface of the earth at the observer's location and we assume that its radiance is constant for all the points on the disk. Substituting the value provided above for σ , 3.1415926 for π , and 6,000 for T , we have $L = 2.34 \times 10^7 \text{ W}/(\text{sr} \cdot \text{m}^2)$ for the source radiance.

Substituting the desired values for T and λ into (3.6) and dividing by π , we have the spectral radiance at 555 nm calculated as follows:

$$\begin{aligned} L_{bb555} &= \frac{3.741 \times 10^8}{\pi(0.555)^5 \left[\exp\left(\frac{14,387.86}{6,000 \times 0.555}\right) - 1 \right]} \\ L_{bb555} &= \frac{1.191062 \times 10^8}{(0.555)^5 [\exp(4.3207) - 1]} . \\ L_{bb555} &= 3.0475 \times 10^7 \text{ W} \cdot \text{m}^{-2} \cdot \mu\text{m}^{-1} \cdot \text{sr}^{-1} \end{aligned} \quad (3.12)$$

or

$$L_{bb555} = 3.0475 \times 10^4 \text{ W} \cdot \text{m}^{-2} \cdot \text{nm}^{-1} \cdot \text{sr}^{-1} \quad (3.13)$$

EXAMPLE PROBLEM 3.2

Problem: What temperature is required for a blackbody to radiate its peak spectral radiance at the peak of human eye response, 555 nm?

Solution: From the Wien displacement law, we solve for temperature T

$$T = \frac{2,897.8 \text{ } \mu\text{m} \cdot \text{K}}{\lambda_m} \quad (3.14)$$

555 nm is 0.555 μm so that $T = 5,221\text{K}$, $4,948\text{C}$, and $8,938\text{F}$.

EXAMPLE PROBLEM 3.3

Problem: How many watts does a 1 ft^2 blackbody at 100°F radiate into a vacuum?

Solution: 100°F is about the same as 38°C or 311K so that from (3.1) we have $E = (5.67032 \times 10^{-8} \text{ W} \cdot \text{m}^{-2} \cdot \text{K}^{-4}) (311\text{K})^4 = 530.5 \text{ W/m}^2$ or $530.5/10.76 \text{ W/ft}^2 = 49.3 \text{ W/ft}^2$. Thus one square foot at 100°F radiates just under 50W.

EXAMPLE PROBLEM 3.4

Problem: The human body at rest in a room at ordinary temperatures produces approximately 70 W/m^2 of net heat through conduction, convection, respiration, and evaporation. If the surface area of the body is $20 \text{ ft}^2 = 1.86\text{m}^2$, then it produces about 130W net heat. If it radiates half as efficiently as a blackbody, what fraction of the 130W is radiation?

Solution: The normal body temperature of 98.6°F is equivalent to about 38°C or 311 K . Thus E is nearly equal to 49.3 W/ft^2 from Example Problem 3.3. Multiplying by 20 ft^2 yields 986W and half of this is 493W. If we were to stop our analysis at this point, we would have to conclude that the body emits a huge amount of radiation and for the net emission to be as low as 130W, then it must absorb the 363W difference by other mechanisms.

The problem with this is that the human body in this problem is not in a radiationless enclosure (one having absolute zero temperature or with zero emittance). Thus, in addition to radiating heat, it will also absorb heat from other emitting bodies around it. In reality, the surroundings are normally only a few degrees below body temperature and they also emit radiation, a portion of which is received and absorbed by the human body. Thus, we should consider the *net radiation exchange* between the human body and its surroundings. This general topic is discussed in the next chapter on radiation exchange.

For the time being, let us assume that the human is surrounded by surfaces emitting half as well as a blackbody, but at a temperature of 78°F , or $25.5^\circ\text{C} = 299\text{K}$. Thus, the net radiation exchange would be given approximately by $0.5 \sigma (T_{\text{hot}}^4 - T_{\text{cold}}^4) = 0.5 \times 5.67032 \times 10^{-8} \times (310^4 - 299^4) = 35.2 \text{ W/m}^2$. Multiplying this by the 1.86 m^2 of body area yields a net transfer of 65W, or about half of our presumed 130W total net heat production. This example is highly artificial, intended only to illustrate the equations developed in this chapter. For more comprehensive treatments of the role of radiation in human physiology, the reader is referred to other sources [27–30].

EXAMPLE PROBLEM 3.5

Problem: Let C_3 be the constant $2897.8 \mu\text{m}/\text{K}$ in (3.7). Solve this equation for λ_m and substitute this into (3.4) to obtain an expression for the peak spectral exitance $M_{\lambda m}$ of a blackbody at temperature T in degrees K.

Solution: Solving (3.6) for λ_m and substituting the result into (3.4) yields

$$M_{bb\lambda m} = \frac{C_1}{\left(\frac{C_3}{T}\right)^5 (e^{C_2/C_3} - 1)} \quad (3.15)$$

which can be simplified to

$$M_{bb\lambda m} = \frac{C_1 T^5}{C_3^5 C_4} \quad (3.16)$$

where C_4 is $\exp(C_2/C_3) - 1$, whose unitless value is 142.33. Thus

$$\begin{aligned} C_1 C_3^{-5} C_4^{-1} &= (3.74183 \times 10^8 \text{ W} \cdot \mu\text{m}^3 \cdot \text{m}^{-2}) (2.8978 \times 10^3 \mu\text{m} \cdot \text{K})^{-5} (142.33)^{-1} \\ &= 1.288 \times 10^{-11} \text{ W} \cdot \mu\text{m}^{-1} \cdot \text{m}^{-2} \cdot \text{K}^{-5} \end{aligned} \quad (3.17)$$

so that our final solution is

$$M_{bb\lambda m} = 1.288 \times 10^{19} T^5 \text{ W} \cdot \text{m}^{-2} \cdot \mu\text{m}^{-1} \cdot \text{K}^{-5} \quad (3.18)$$

Equation (3.18) can be used quickly to determine the peak spectral exitance of a blackbody at any temperature.

REFERENCES

- [1] Weast, R. C., ed. *CRC Handbook of Chemistry and Physics*, 66th ed., CRC Press: Boca Raton, FL, 1985, p. F-195.
- [2] Grum, F., and R. J. Becherer. *Optical Radiation Measurements, Volume 1 Radiometry*, Academic Press: New York, NY, 1979, Section 4.2.1.
- [3] Planck, M., *Ann. Phys.*, Vol. 4, No. 3, 1901, p. 553.
- [4] Grum, F., and R. J. Becherer. *Optical Radiation Measurements, Volume 1 Radiometry*, Academic Press: New York, NY, 1979, Section 4.3.3.
- [5] Driscoll, W. G., ed. *Handbook of Optics*, McGraw-Hill Book Co.: New York, NY, 1978.
- [6] Spiro, I. J., and M. Schlessinger. *Infrared Technology Fundamentals*, Marcel Dekker: New York, NY, 1989.
- [7] Andreic, Z.. "Distribution Temperature Calculations by Fitting the Planck Radiation Curve to a Measured Spectrum," *Applied Optics*, Vol. 31, No. 1, Jan. 1992, pp. 126–130.

-
- [8] Grum, F., and R. J. Becherer. *Optical Radiation Measurements, Vol. 1, Radiometry*, Academic Press: New York, NY, 1979.
 - [9] Harrison, T. R. *Radiation Pyrometry and its Underlying Principles of Radiant Heat Transfer*, John Wiley & Sons: New York, NY, 1960, Appendix A, pp. 207–209.
 - [10] Wysczecki, G., and W. S. Stiles. *Color Science: Concepts and Methods, Quantitative Data and Formulae*, 2nd ed., John Wiley & Sons: New York, NY, 1982.
 - [11] Conn, G.K.T., and D. G. Avery. *Infrared Methods*, Academic Press: New York, NY, 1960.
 - [12] Pivovovsky, M. and M. R. Nagel. *Tables of Blackbody Radiation Functions*, Macmillan: New York, NY, 1961.
 - [13] Hahn, D., J. Metzdorf, U. Schley, and J. Verch. *Seven-Place Tables of the Planck Function for the Visible Spectrum*, Academic Press: New York, NY, 1964.
 - [14] Levi, L. *Handbook of Tables of Functions for Applied Optics*, CRC Press: Boca Raton, FL, 1974.
 - [15] Siegel, R., and J. R. Howell. *Thermal Radiation Heat Transfer*, 2nd ed., Hemispherical Publishing/McGraw-Hill Book Co.: New York, NY, 1981.
 - [16] McCluney, R. “The Importance of the International Daylight Measurement Year,” *Lighting Design + Application*, May 1990, pp. 32–41.
 - [17] McCluney, R. “Bringing in the Sun,” *Glass Magazine*, Aug. 1986, pp. 60–63.
 - [18] McCluney, R. “Color-Rendering of Daylight from Water-Filled Light Pipes,” *Solar Energy Materials*, Vol. 21, 1990, pp. 191–206.
 - [19] Powell, I. “New Concept for a System Suitable for Solar Simulation,” *Appl. Opt.*, Vol. 19, 1980, pp. 329–334.
 - [20] Corrons, A., and A. Pons. “Daylight Simulator,” *Appl. Opt.*, Vol. 26, 1987, pp. 2867–2870.
 - [21] Oriel Corporation. “Light Sources, Monochromators & Spectrographs, Detectors & Detection Systems, Fiber Optics, Vol. II,” product catalog, Oriel Corporation, 250 Long Beach Blvd., P. O. Box 872, Stratford, CT 06497, 1994.
 - [22] Mims, F. M. III. “Sun Photometer With Light-Emitting Diodes as Spectrally Selective Detectors,” *Appl. Opt.*, Vol. 31, 1992, pp. 6965–6967.
 - [23] Carter, W. H., “Four Definitions for the Term Emittance,” *Appl. Opt.*, Vol. 19, 1980, pp. 3419–3420. The term “emittance” used by Carter refers to what is called exitance in this book, conceptually equivalent to irradiance.
 - [24] Carlson, F. E., and C. N. Clarke. “Light Sources for Optical Devices,” in *Applied Optics and Optical Engineering*, Vol. I, ed. by R. Kingslake, Academic Press: New York, NY, 1965.
 - [25] Oriel Corporation, “Light Sources, Monochromators & Spectrographs, Detectors & Detection Systems, Fiber Optics,” Volume II, Oriel Corporation, 250 Long Beach Blvd., P. O. Box 872, Stratford, CT 06497, 1994.
 - [26] Kaufman, J. E., ed. *IES Lighting Handbook – 1984 Reference Volume*, IESNA, 345 E. 45 St., New York, 1984.
 - [27] American Society of Heating, Refrigerating, and Air Conditioning Engineers. *1989 ASHRAE Handbook of Fundamentals*, ASHRAE, 1791 Tullie Circle, N. E., Atlanta, GA 30329, 1989.
 - [28] Fanger, P. O. *Thermal Comfort*, Robert E. Krieger Publishing Co.: Malabar, FL, 1982.
 - [29] Fanger, P. O. *Thermal Comfort Analysis and Applications in Environmental Engineering*, McGraw-Hill Book Co.: New York, NY, 1970.
 - [30] Gagge, A. P., and J. D. Hardy. “Thermal Radiation Exchange of the Human Body by Partition Calorimetry,” *Journal of Applied Physiology*, Vol. 23, 249\8 1967.

Chapter 4

Source/Receiver Flux Transfer Calculations

4.1 INTRODUCTION

The first two chapters were devoted to definitions and terminology. The third dealt with blackbody radiation and other sources. Real problems in radiometry and photometry generally involve determinations of the quantity of flux transferred from one place to another, or from one surface to another. For example, if one wishes to calculate the illumination on a table top produced by a rectangular luminaire in the ceiling of a room, or if the irradiance from a heat lamp received by a person seated at a table in an outdoor restaurant in New Orleans is needed, the principles of flux transfer can be employed. The study of flux transfer begins in this chapter, first showing the basic equations and procedures for performing the calculations and then providing some examples.

Only the geometrical aspects of flux transfer are of interest in this treatment, including the famous “inverse square law” governing the decline of irradiance and illuminance with distance from a point source. The effects of absorption and scattering of radiation as it propagates through the medium from a source to a receiver are dealt with in later chapters, as is the effect of changing refractive index in the medium.

All uses of flux quantities in this chapter refer to both their radiant (subscript e) and luminous (subscript v) versions. The subscripts are left off for simplicity. When the terms radiance and irradiance are used in this chapter, the discussion applies equally to luminance and illuminance, respectively.

4.2 GEOMETRY AND DEFINITIONS

We begin with the drawing of Figure 4.1 and the definition of radiance L in (1.13) of Chapter 1 (or of luminance, using the same equation but replacing L with L_v and Φ with Φ_v):

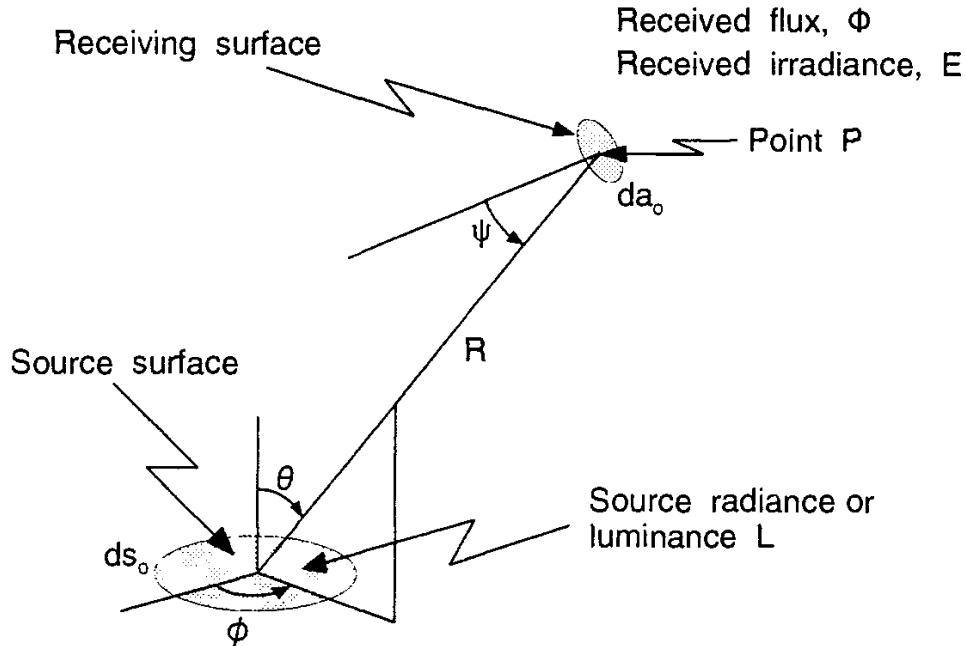


Figure 4.1 Source/receiver geometry.

$$L = \frac{d^2\Phi}{d\omega \, ds_o \cos \theta} \quad (4.1)$$

where θ is the angle made by the direction of emerging flux with respect to the normal to the surface of the source, ds_o is an infinitesimally small element of area at the point of definition in the source, and $d\omega$ is an element of solid angle from the point of definition in the direction of interest. In Figure 4.1 are shown an infinitesimally small element ds_o of area at a point in a source, an infinitesimal element da_o of area at point P on a receiving surface, the distance R between these points, and the angles θ and ψ between the line of length R between the points and the normals to the surfaces at the points of intersection, respectively.

The element $d\omega$ of solid angle subtended by element of projected receiver area $da_o = da_o \cos \psi$ at distance R from the source is

$$d\omega = \frac{da}{R^2} = \frac{da_o \cos \psi}{R^2} \quad (4.2)$$

so that, solving (4.1) for $d^2\Phi$ and using (4.2), the element of flux received at point P from the element ds_o of area of the source is given by

$$d^2\Phi = L \frac{ds_o \cos \theta \, da_o \cos \psi}{R^2} \quad (4.3)$$

with the total flux Φ received by area A_o from source area S_o being given by

$$\Phi = \int_{S_o} \int_{A_o} L \frac{ds_o \cos \theta da_o \cos \psi}{R^2} \quad (4.4)$$

This is the fundamental equation describing the transfer of radiation from a source surface to a receiving surface. Most problems of flux transfer involve this integration over finite areas of the source and receiver. The problem can be quite complex analytically because in general L , θ , ψ , and R will be functions of position in both the source and the receiver surfaces. The general dependency of L on direction is also embodied in this equation, since the direction from a point in the source to a point in the receiver generally changes as the point in the receiver moves over the receiving surface. These complications in the application of (4.4) to potential problems will become clearer in the examples and special cases to follow.

Before we begin work on this most general version of the problem, there are two important simplifications which address a large class of problems in radiometry and photometry, and which also happen to make analysis easier.

The first simplification results when the source of radiance is Lambertian and, in addition, has the same value for all points in the source surface. This makes L constant over all ranges of integration, both the integration over the surface area and over the hemispherical solid angle of emerging directions from each point on the surface. In these cases, the radiance can be removed from any integrals over these variables and the remaining integrals are geometric in character.

The second simplification arises when one doesn't want the total flux over the whole receiving surface—only the flux per unit area at a point on that surface—the irradiance or illuminance E at point P in Figure 4.1. In this case, we can divide both sides of (4.3) by the element of area in the receiving surface, da_o , to get

$$dE = \frac{d^2\Phi}{da_o} = L \frac{\cos \theta \cos \psi ds_o}{R^2} \quad (4.5)$$

This equation is the counterpart of (4.3) when it is the irradiance E of the receiving surface that is desired. For the total irradiance at point P , one must integrate this equation over the portion S_o of the source surface contributing to the flux at P

$$E = \int_{S_o} L \frac{\cos \theta \cos \psi ds_o}{R^2} \quad (4.6)$$

Of course, when L is constant over direction, it can be removed from this integral and one is left with a simpler integration to perform. Equation (4.6) is the coun-

terpart to (4.4) when it is the irradiance E at the receiving point that is of interest rather than the flux.

Now we look at some special cases of (4.4) and (4.6) that simplify the problem and help us to better understand the physics of the processes involved. They will also help us describe the most general case and point the way toward solving more general flux transfer problems.

4.2.1 Case 1

If the source area S_o is small with respect to the distance R to the point P of interest (i. e., if the maximum dimension of the source is small compared with R , as illustrated in Figure 4.2), then R^2 , $\cos \psi$, and $\cos \theta$ don't vary much over the range of integration shown in (4.6) and they can be removed from the integral. If L doesn't vary over S_o , then it also can be removed from the integral, even if L is direction dependent, because the range of integration over direction is so small; that is, only the one direction from the source to point P in the receiver is of interest. We are left with an approximate version of (4.6) for small homogeneous sources some distance from the point of reception:

$$E \approx \frac{LS_o \cos \theta \cos \psi}{R^2} \quad (4.7)$$

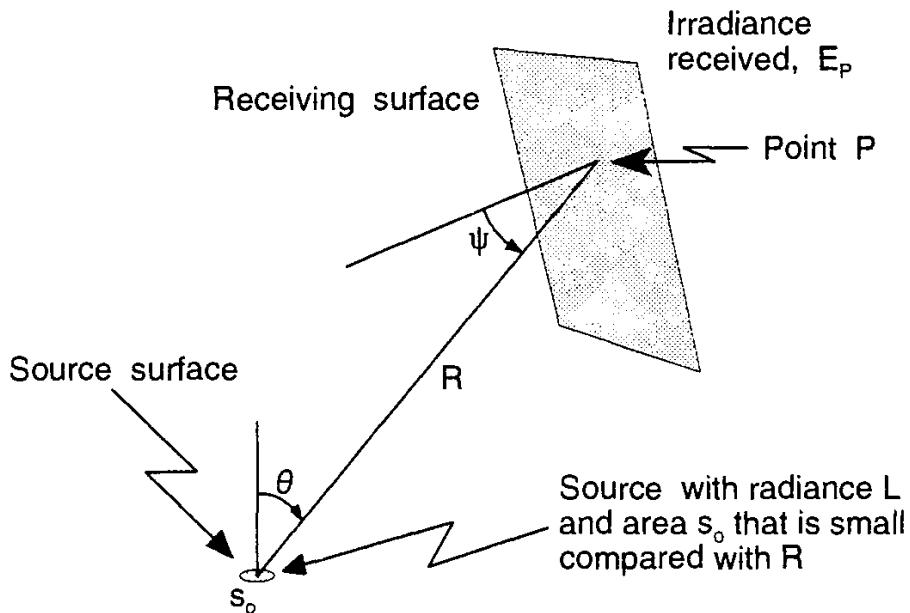


Figure 4.2 Source and receiver configuration where the source area is small with respect to the distance between them.

If the source and receiving surfaces face each other directly, so that θ and ψ are zero, both of the cosines in this equation have values of unity and the equation is still simpler in form.

Equation (4.7) shows clearly what is called the *inverse square law* governing the reduction in irradiance/illuminance with distance from a point source. To be strictly correct, we are not here dealing with a truly point source, just one that is very small with respect to the distance R . Thus, we are treating the source as if it were a point. This is an important point, because many real and finite sources of light and other radiation we deal with can be very closely approximated as if they were point sources if the distance R is great enough. An ordinary light bulb, for example, can be treated as a point source if the distance from it is greater than about 3 meters. Automobile headlights can be so treated if the distance is greater than 50 meters or so. A criterion can be suggested for deciding if a finite source can be approximated by a point source. If the maximum lateral dimension x of a source is less than, say, 10% of the distance R to it, then the ratio of the squares of x and R will be less than 1% and this can serve as an approximate criterion for when a finite source can be treated as a point source.

Looking back at Section 1.3 on the definition of solid angle, we can see that the inverse square dependence exhibited in (4.7) arises from the fact that the area used to define a solid angle increases with the square of the distance from the point of definition. Thus, if the solid angle is filled with constant flux, this same flux will be distributed over a larger and larger area as the distance from the point increases. Since this area *increases* with the square of the distance, the *flux density* (irradiance) *decreases* with the square of the distance and the inverse square law results.

Equation (4.7) also contains in it Lambert's cosine law, the variation in irradiance with the cosine of the angle ψ of essentially collimated (parallel rays) radiation received by a surface. This idea was discussed in some detail in Section 1.4.

4.2.2 Case 2

We can write the general equation (4.5) in a different form. Note that the element of solid angle $d\Omega$ subtended by the *source* at point P on the receiver is given by

$$d\Omega = \frac{ds_o \cos \theta}{R^2} \quad (4.8)$$

so that (4.5) becomes

$$dE = L d\Omega \cos \psi \quad (4.9)$$

To obtain the total irradiance at point P , this expression must be integrated over the whole solid angle subtended by the source at point P . The result is

$$E = \int L \cos \psi d\Omega \quad (4.10)$$

Thus it is seen that in real problems involving determination of E at a point in a receiving surface, one must perform an integration over the solid angle subtended by the *source* at the point of interest in the receiver. Equation (4.10) is useful when one wishes to determine the illuminance at a point on a horizontal table from a window to the side. If the luminance L of the window from a point in its aperture to the point of observation on the table varies over the aperture, this variation must be known before the integration in (4.10) can be performed. If the luminance of the aperture is known to be constant over the solid angle range of integration, then L can be removed from the integrand and the resulting equation is a simplified version of (4.10):

$$E = L \int \cos \psi d\Omega \quad (4.11)$$

Note that everything to the right of the radiance L in (4.11) is purely geometric in nature. This leads to a subject dealt with in Section 4.3 called the *configuration factor*. The configuration factor is essentially the purely geometric portion of equations in flux transfer, such as (4.11), where the geometric portion can be separated out from the variations in source radiance.

4.2.3 Case 3

Now let us look at the opposite extreme from Case 1, where the source is very large with respect to the distance R to the receiver. Let's consider an infinite plane. To get this result, we perform the integration of (4.10) over the solid angle subtended by the source at the receiver. If the radiance L is Lambertian and constant over the whole infinite plane source, then L can be removed from the integral, giving (4.11) and the following expression for the irradiance at the receiving point from an infinite plane homogenous Lambertian source:

$$E = L \int_{2\pi} \cos \psi d\Omega \quad (4.12)$$

Equation (4.12) is mathematically (but not conceptually) the same as (1.17), with L constant over the range of integration. The result of the integration is the same as (1.19): $E = \pi L$.

This is a very interesting result. It says that the irradiance on a receiving surface from a homogeneous infinite Lambertian source is constant, independent of distance from that source (as long as there is no intervening absorbing or scattering medium), because the solid angle of the source and the radiance/luminance as seen from the point P of observation are always the same, regardless of distance from the source plane. $E = \pi L$ is also true for the irradiance at a point receiving flux from a hemispherical source of any radius and constant radiance.

The radiance is constant with distance also. If one points a radiance meter at a wall having uniform Lambertian radiance and walks away from the wall, the radiance reading will stay the same until the field of view of the meter exceeds the area of uniformity of the wall, or until the wall fails to fill the meter's field of view. More will be said about the propagation of radiance along a ray from a source in the next chapter, but it is interesting here to see the invariance of radiance with distance from a Lambertian and uniform source as long as the source fills the field of view of the radiance meter and the intervening medium is perfectly transparent.

The irradiance received on a surface from a *finite* source will lie somewhere between the extremes of Case 1, exhibiting the $1/R^2$ fall-off of a point source and Case 3, being constant for a uniform infinite plane, independent of distance R , as seen in the next case.

4.2.4 Case 4

Let's look at a slightly more difficult case with a finite source and an observation point some distance away. Suppose one is interested only in the irradiance at the single point P in a surface parallel to the source. The situation is diagrammed in Figure 4.3, where the flux to P comes from a finite source whose center is distance H away. To simplify the mathematics, a planar source with circular symmetry is selected. The geometry needed to solve the problem is shown in Figure 4.3.

We look at infinitesimal element ds_o of area within a planar source having the shape of a circular disk of radius r_o . The point P of observation lies on the perpendicular to the source passing through its center, where the origin of an x - y coordinate system in the plane of the source is placed. Element dE of irradiance at point P , due to the element ds_o in the source, distance R away, will be given by (4.5), in the following form

$$dE = \frac{L ds_o \cos^2 \theta}{R^2} \quad (4.13)$$

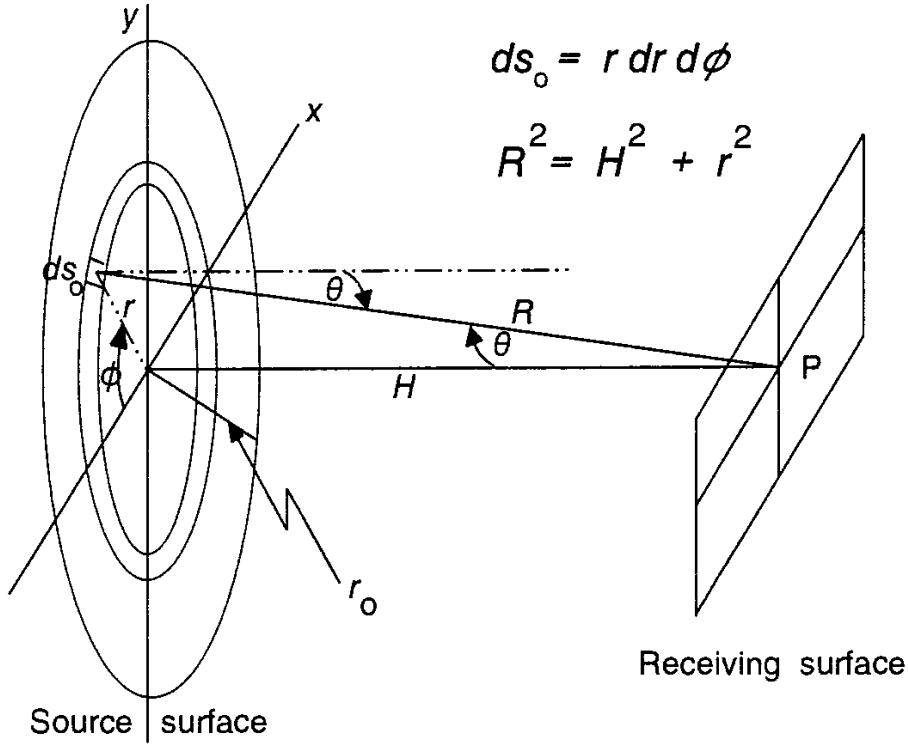


Figure 4.3 Source and receiver configuration where the source is not small with respect to the distance between them.

Due to the geometry of this problem, the angle ψ in (4.5) is equal to angle θ . (Hence the $\cos^2 \theta$ term.) Looking at the figure we also see that

$$ds_o = r dr d\phi \quad (4.14)$$

and

$$R^2 = H^2 + r^2 \quad (4.15)$$

so that the integral to be evaluated is

$$E = L \int_0^{2\pi} \int_0^{r_o} \frac{r \cos^2 \theta}{H^2 + r^2} dr d\phi \quad (4.16)$$

where r_o defines the maximum extent of the source. However,

$$r = H \tan \theta \quad (4.17)$$

so we must either express θ in terms of r or r in terms of θ to evaluate this integral.

Choosing the latter, we replace r with $H \tan \theta$ and dr with $d(H \tan \theta)$ to obtain

$$E = 2\pi L \int_0^{\theta_o} \frac{H \tan \theta \cos^2 \theta}{H^2 + H^2 \tan^2 \theta} d(H \tan \theta) \quad (4.18)$$

Now

$$d(H \tan \theta) = H \sec^2 \theta d\theta \quad (4.19)$$

so that

$$E = 2\pi L \int_0^{\theta_o} \frac{\tan \theta}{1 + \tan^2 \theta} d\theta \quad (4.20)$$

We can transform the integrand to a form more easily integrated by replacing $\tan \theta$ with $\sin \theta / \cos \theta$ and then multiplying the integrand by $\cos^2 \theta / \cos^2 \theta$ so that we have

$$\frac{\tan \theta}{1 + \tan^2 \theta} = \sin \theta \cos \theta \quad (4.21)$$

Equation (4.20) becomes

$$E = 2\pi L \int_0^{\theta_o} \sin \theta \cos \theta d\theta = \pi L \sin^2 \theta_o \quad (4.22)$$

This is a reasonable result. It goes to zero as θ_o goes to zero and it approaches πL as θ_o approaches 90 deg (the source expands to infinity). In the latter case, the source is an infinite plane and we have the same result as in Case 3.

4.2.5 Case 5

Let's look at the geometry shown in Figure 4.4. Here we have a planar source of width $2w_1$ and height $2h_1$ parallel to a flat receiver of width $2w_2$ and height $2h_2$ with the center of each area lying on the perpendicular line of length Z connecting them. The unit vectors \mathbf{u}_1 and \mathbf{u}_2 are perpendicular to elemental areas ds_1 and ds_2 at points P_1 and P_2 in the source and receiver surfaces, respectively. The angles between the \mathbf{u}_1 and \mathbf{u}_2 directions and the line of length R connecting P_1 and P_2 are γ_1 and γ_2 , respectively.

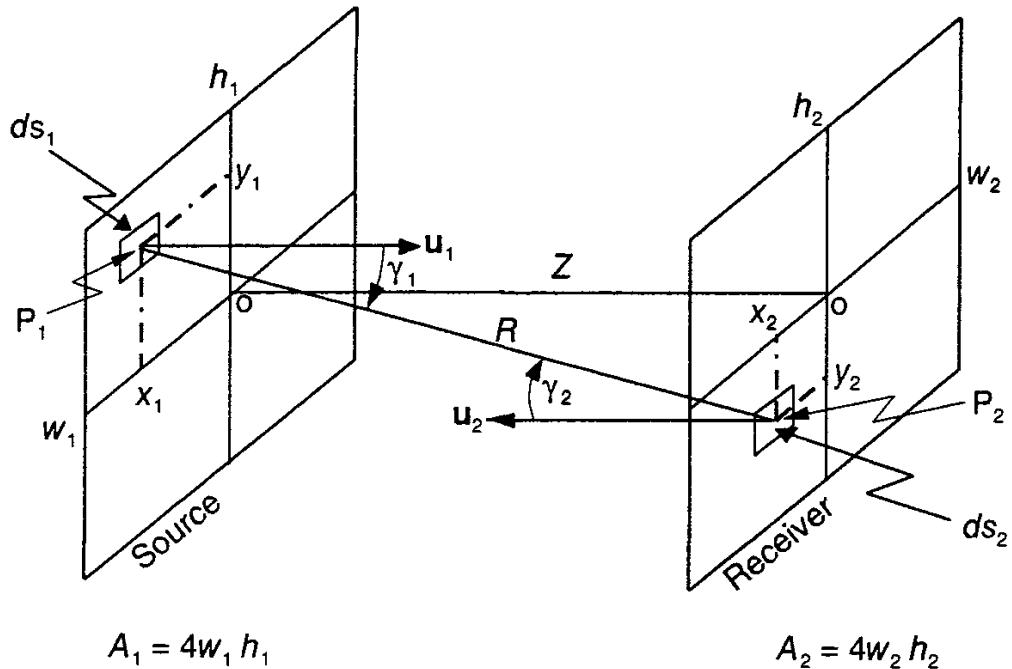


Figure 4.4 Geometry for Case 5.

If the radiance L is isotropic and homogeneous over the source, it can be removed from the integrals. Integrating (4.3) for this case and geometry yields

$$\Phi = L \int_{-w_1}^{+w_1} \int_{-h_1}^{+h_1} \int_{-w_2}^{+w_2} \int_{-h_2}^{+h_2} \frac{\cos \gamma_1 \cos \gamma_2 dx_1 dy_1 dx_2 dy_2}{(x_1 - x_2)^2 + (y_1 - y_2)^2 + Z^2} \quad (4.23)$$

The denominator of the integrand in (4.23) is R^2 , expressed in the rectangular coordinates appropriate to the problem. The remaining tasks are to express the cosines of γ_1 and γ_2 in terms of the coordinates of P_1 and P_2 and the distance Z and then to integrate the resulting expression. This will be left as an exercise for the student. Equation (4.23) illustrates the mathematical complexities that result when finite source and receiver areas are involved that cannot be approximated as point sources.

4.2.6 Case 6

Now we turn to an even more general problem where the source and receiver geometries are more complicated: rectangular source and receiving surfaces that are inclined with respect to each other. The situation could be that of calculating the illuminance on a wall or table from a finite source such as a window in a room with black walls (so that there are no interreflections in the room), or the infrared

radiation received by a flat test object in an experimental chamber, one wall of which is hot with the other walls designed to emit little or no thermal radiation.

The geometry of the problem is as shown in Figure 4.5. The distance R between the element ds_o of area in the source and the observation point P in the receiver is given by

$$R^2 = (x_s - x_r)^2 + (y_s - y_r)^2 + (z_s - z_r)^2 \quad (4.24)$$

Let γ_s and γ_r be the angles made by the line R with the normals to the source and receiver planes, respectively. With some work, these angles can be written as functions of x_s , y_s , x_r , and y_r . The element of area ds_o in the source is given by

$$ds_o = \frac{dx_s dy_s}{\cos \theta_s} \quad (4.25)$$

so that (4.6) can be written as

$$E_p = L \int \int \frac{\cos \gamma_s(x_s, y_s, x_r, y_r) \cos \gamma_r(x_s, y_s, x_r, y_r) dx_s dy_s}{[(x_s - x_r)^2 + (y_s - y_r)^2 + (z_s - z_r)^2] \cos \theta_s} \quad (4.26)$$

with the angle γ_s and γ_r being shown as functions of x_s , y_s , x_r , and y_r . L has been

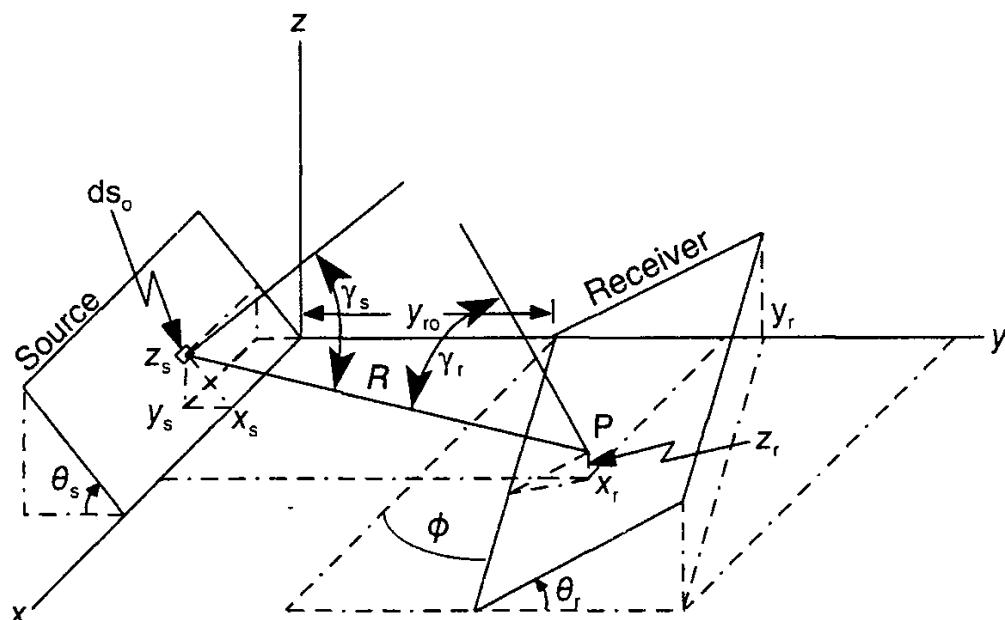


Figure 4.5 Radiation exchange geometry for nonparallel planar source and receiver.

taken outside the integral in accordance with the assumption of source homogeneity and angular isotropy used in this case.

Once the functional dependences of θ_s , γ_s , and γ_r on x_s , y_s , x_r , and y_r have been determined, this integral can be evaluated, either analytically if the resulting integrand is integrable, or numerically if it is not. Even if the integral in (4.26) can be evaluated analytically, one might still choose to evaluate it numerically.

With L removed from the integral, (4.26) should in general be easier to evaluate. However, this is not necessarily the case. Many seemingly simple geometries result in functional dependencies in the integrand of (4.26) that make this integral impossible to evaluate in closed form (analytically). Other techniques have been developed to deal with these situations, including numerical evaluation as sketched in the Appendix, and a novel approach called the Monte Carlo method, which is outlined in Section 4.2.8.

Whenever the radiance L can be taken out of the integral, what is left behind depends only on the geometry of the situation. A version of the remaining integral, called the configuration factor, is described in Section 4.3.

4.2.7 Case 7

We now return to the most general flux transfer problem, where the source is large and arbitrarily oriented, and the flux is received by a finite portion of a receiving surface. For this we need to integrate (4.3) over both source S_o and receiver A_o areas:

$$\Phi = \int_{S_o} \int_{A_o} \frac{L \cos \theta \cos \psi ds_o da_o}{R^2} \quad (4.27)$$

This integral can be very difficult to evaluate for complex geometries, where all the parameters in the integrand are complicated functions of the variables of integration. Such integrations are best performed numerically by finite element techniques, a generalization of the procedure outlined in the Appendix, or by the Monte Carlo method mentioned previously. Siegel and Howell [1] present a comprehensive analysis of radiant transfer and point out that the chief difficulties with problems in this area are not with the concepts involved but “in the geometrical and algebraic manipulations and the integrations that must be carried out to determine the configuration factors for specific geometries. These difficulties are minimized by the availability in the literature of fairly extensive formulas, graphs, and tabulations of configuration factors that have already been calculated.” Siegel and Howell provide these factors for a number of geometries and offer references to many more in the literature.

4.2.8 Monte Carlo Method

The Monte Carlo method mentioned above is an alternative to the evaluation of integrals. It is named for the famous gambling casino because it is based upon a kind of random process, a conceptual process performed on a computer in which photons of radiation are fired from a point in the source surface in random directions over the solid angle subtended by the receiver at that point. The number of photons fired in each of a finite (but very large) number of directions in this solid angle is weighted to match the angular distribution of radiance $L_p(\theta, \phi)$ at point P in the source.

The computer program performing these calculations is made to keep track of how many photons are received at each of a large number of small patches of area spread over the receiving surface. After a sufficiently large number of photons has been traced from source to receiver by this process (the limit being determined by the size and speed of the computer used and the investigator's need for accuracy), the process is terminated and the fraction of photons fired from the source that reach each different patch of the receiver (or a single one of interest) is determined. Since the energy content of each photon is known, the flux distribution over the second surface can easily be determined from these fractions.

One of the beauties of the Monte Carlo approach is that it can be used for the case of interreflections between surfaces. The main objective of this chapter has been to describe a formalism for treating the case of radiation transfer from one surface in space to another one. In real situations, however, the radiation received by the second surface is in general partially reflected, absorbed, and transmitted by that surface. Some of the reflected radiation can then propagate according to its own reflected radiance distribution L_r , back to the first surface, where it can be partially reflected again. There is in general an infinite number of multiple reflections between two surfaces in "sight" of each other.

With the Monte Carlo method, each photon from surface 1 incident upon surface 2 can be made to produce a very large number of reflected photons, directed (weighted) according to the angular reflection properties of that surface, and filling the solid angle subtended by surface 1 at the point of reflection on surface 2. This process is repeated for each subsequent reflection and the program keeps track of where all the photons go. To keep the process from proceeding on indefinitely, the program is made to determine after each subsequent reflection if the quantity of flux involved is less than some lower limit set by the investigator. If it is, the process is stopped with confidence that a specified large fraction of the desired flux has been accounted for. The accuracy of the results obtained this way is limited by the accuracies of the input radiance distributions and optical property distributions as well as by the number of photons that can be tracked in a reasonable period of time. The faster the computer, the greater the number of photons that can be tracked in a given time.

The details of setting up Monte Carlo calculations can be found in modern textbooks on statistical analysis and in various papers in this field [2–4]. Another important use of Monte Carlo calculations arises in the field of radiative transfer, to deal with flux propagation in a scattering medium where multiple scatterings take place. Radiative transfer is discussed in Sections 5.8 and 6.5.2.

4.3 CONFIGURATION FACTOR

In analyzing complicated radiation transfer problems, it is frequently helpful to introduce what was referred to previously as the configuration factor. Alternate names for this factor are the *view, angle, shape, interchange, or exchange factor* [1]. This purely geometric factor is defined to be the fraction of total flux from the source surface that is received by the receiving surface. It is given the symbol F_{s-r} , or F_{1-2} , indicating flux transfer from source to receiver or from surface 1 to surface 2. In essence, it indicates the details of how flux is transferred from a source area of some known form to a reception area.

In many problems, one is most concerned with the magnitude and spectral distribution of the source radiance and the corresponding spectral irradiance at some other point, rather than with the geometrical aspects of the problem that are embodied in the shape factor. It is very convenient in such cases to separate the spectral variations from the geometrical variations, when possible. When such separation is possible, all geometrical information can be put into the configuration factor. This is where the configuration factor comes in. Once this factor has been determined for a situation with nonchanging geometry, it remains constant and attention can be focused on the variable portions of (4.6), (4.25), or (4.26).

We can obtain an expression for the configuration factor by dividing (4.24) for the flux Φ_r on the receiver by the total flux Φ_s emitted by the source, as follows:

$$F_{s-r} = \frac{\Phi_r}{\Phi_s} = \frac{\int_{S_o} \int_{A_o} \frac{L \cos \theta \cos \psi ds_o da_o}{R^2}}{\int_{S_o} \int_{2\pi} L \cos \theta d\omega ds_o} \quad (4.28)$$

This is the general expression for configuration factor. If the source is Lambertian and homogeneous, then L can be removed from the integrals, resulting in

$$F_{s-r} = \frac{\int_{S_o} \int_{A_o} \frac{\cos \theta \cos \psi ds_o da_o}{R^2}}{\pi S_o} \quad (4.29)$$

which is the conventional form of the configuration factor, and, as desired, it has no radiation components. In the denominator of (4.29) the factor of π comes from the integral over a hemispherical solid angle and S_o comes from the integral over source area. The configuration factor F_{s-r} , given by (4.29), expresses the purely geometrical part of the problem. For homogeneous Lambertian sources, of radiance L_s , the flux Φ_{s-r} , is given by

$$\Phi_{s-r} = \pi L_s S_o F_{s-r}, \quad (4.30)$$

There are three other classes of configuration factor. The first is the trivial one for the geometry between two arbitrarily oriented and spaced infinitesimal areas. The second is for a finite source and an infinitesimally small receiver. The third is for an infinitesimal source and a finite receiver. A special notation has been devised by Siegel and Howell [1] to distinguish between these four configuration factors.

We begin with the case of infinitesimal source and receiver areas. Since both surfaces are infinitesimally small, no integration is needed and the configuration factor, which is designated dF_{ds-dr} , can be easily obtained by dividing (4.3) by the element $d\Phi_s$ of flux leaving the source in all directions, given by:

$$d\Phi_s = \pi L ds_o, \quad (4.31)$$

resulting in the configuration factor

$$dF_{ds-da} = \frac{\cos \psi \cos \theta da_o}{\pi R^2} = \frac{\cos \theta d\omega_r}{\pi} \quad (4.32)$$

where $d\omega_r$ is the infinitesimally small solid angle subtended by the element of receiver area da_o at the element of source area distance R away.

The notation used here for differential configuration factors follows that of Siegel and Howell [1], who point out that the differential notation dF is redundant with the subscript notation. The differential notation is nevertheless retained to keep the mathematical form of equations such as (4.32) consistent in having a differential quantity on both sides of the equation.

Next comes the configuration factor, denoted dF_{s-dr} , from a finite source to an infinitesimal receiving area and the related factor dF_{ds-r} , for going from an infinitesimal source to a finite receiver. It is easy to show reciprocity in differential-element configuration factors [1], as stated by the equation $dF_{ds-dr} = dF_{dr-ds}$. This means that it hardly matters which surface is the source and which is the receiver when infinitesimals are involved.

When finite areas are being treated, there are reciprocity relations for flux transfer in both directions. One frequently sees configuration factors written with

the alternative notation F_{1-2} , F_{2-1} , F_{12} , and F_{21} . For finite sources the reciprocity relation is [1]

$$A_s F_{s-r} = A_r F_{r-s} \quad (4.33)$$

Further information on configuration factors and their use can be found in the classic text by Siegel and Howell [1].

4.4. NET EXCHANGE OF RADIATION

To this point, the focus has been on describing the quantity of flux transported from one surface to another without considering multiple reflections or the emission of other sources back to the first one. This subject was discussed briefly in Section 3.6 and when describing the Monte Carlo method in the previous section. In the real world, there are generally many surfaces that exchange radiation with each other. All surfaces above a temperature of absolute zero emit and absorb radiation, and therefore exchange energy with each other. Those with reflectances that are nonzero also reflect and/or scatter a portion of the incident radiation, and some of this will be received by the other source surfaces.

In this section, the interest is on the thermal radiant part of the exchange between surfaces, not in the reflected and scattered radiation that is exchanged between them. Thus the surfaces are treated as blackbodies, without reflection, to simplify the problem and focus attention on the radiant exchange mechanisms. All received radiation is absorbed and turned into heat within the surface, possibly causing the temperature to rise. The surface also emits radiation, depending upon its temperature, as described in the previous chapter. Since both of the surfaces are emitting and absorbing energy in exchange with each other, one is interested in describing the net radiant exchange of energy between them.

We begin with the differential quantity of flux $d^2\Phi_{1-2}$ transferred from surface 1 to surface 2, given by (4.3) in the form

$$d^2\Phi_{1-2} = L_{bb1} \frac{ds_1 \cos \theta_1 ds_2 \cos \theta_2}{R^2} \quad (4.34)$$

and the flux $d^2\Phi_{2-1}$ transferred from surface 2 to surface 1, given by

$$d^2\Phi_{2-1} = L_{bb2} \frac{ds_2 \cos \theta_2 ds_1 \cos \theta_1}{R^2} \quad (4.35)$$

The net flux transferred between the surface elements, $d^2\Phi_{1 \leftrightarrow 2}$ is the difference between these two quantities:

$$d^2\Phi_{1 \leftrightarrow 2} \equiv d^2\Phi_{1-2} - d^2\Phi_{2-1} = (L_{bb1} - L_{bb2}) \frac{\cos \theta_1 \cos \theta_2}{R^2} ds_1 ds_2 \quad (4.36)$$

Since the surfaces are blackbodies, this quantity can be written as

$$d^2\Phi_{1 \leftrightarrow 2} = \sigma(T_1^4 - T_2^4) \frac{\cos \theta_1 \cos \theta_2}{\pi R^2} ds_1 ds_2 \quad (4.37)$$

For finite blackbodies exchanging radiation, the flux from surface 1 to surface 2 was given in (4.30). Since $L_1 = \sigma T_1^4 / \pi$, (4.30) becomes

$$\Phi_{1-2} = \sigma T_1^4 S_1 F_{1-2} \quad (4.38)$$

A similar relationship is given for Φ_{2-1} , so that the net transfer between the two surfaces is

$$\Phi_{1 \leftrightarrow 2} = \sigma(T_1^4 - T_2^4) S_1 F_{1-2} \quad (4.39)$$

where S_1 is the area of the source and we have used reciprocity relation (4.33).

4.5 SUMMARY

One can solve radiation transfer problems with the following procedure. Start with a point in the source and a point in the receiver, separated by the distance R . Integrate the appropriate one of equations (4.3), (4.5), (4.9), (4.13), or (4.26). Most of the work is involved in first setting up the geometry and then determining the functional relationships amongst the variables. If any of the parameters inside the integrals vary significantly over any of the ranges of integration, then those parameters must stay inside the integrals. If the integrand is not integrable, then the problem can be solved numerically by any of several available methods. This is generally the case, for example, when the directional dependence of the radiance L is measured data that cannot be fit by simple functions.

In the most general case, of two surfaces fairly close together (R less than or comparable to the dimensions of the surfaces), with the source radiance/luminance being non-constant in both position in the source and direction to the receiver, L , R , θ , and ψ will vary over the ranges of integration and the problem can be mathematically very complex. Situations such as these are treated by Siegel and Howell [1] who present two valuable appendices. Their Appendix B contains ref-

erences to about 200 configuration factors that have been worked out by other authors and are available in the literature. Their Appendix C contains drawings and equations for the configuration factors for a number of common geometries. The configuration factor for Example Problem 4.1 comes from Siegel and Howell's Appendix C.

EXAMPLE PROBLEM 4.1

Problem: The reciprocity relation for configuration factors between a differential and a finite area is [1] $A_2 dF_{2-d1} = dA_1 F_{d1-2}$. The configuration factor for going from plane element of area dA_1 on a table in a room to rectangular plane area A_2 on a wall of the room perpendicular to the differential area, as shown in Figure 4.6, is given by [1]

$$F_{d1-2} = \frac{1}{2\pi} \left(\tan^{-1} \frac{1}{Y} - \frac{Y}{\sqrt{X^2 + Y^2}} \tan^{-1} \frac{1}{\sqrt{X^2 + Y^2}} \right) \quad (4.40)$$

where

$$X = \frac{a}{b} \quad Y = \frac{c}{b} \quad (4.41)$$

If area A_2 is that of a window having uniform Lambertian luminance L_1 and dA_1 is at a point on the table, even with the bottom of the window, find the illuminance E_2 at this point on the table.

Solution: Using the reciprocity relation, we can obtain the configuration factor for going from the window, surface 2, to the table, surface 1:

$$dF_{2-d1} = \frac{F_{d1-2} dA_1}{A_2} \quad (4.42)$$

The element of flux Φ_{2-d1} at surface 1 is given by (4.30), in the form

$$d\Phi_{2-d1} = \pi L_2 A_2 dF_{2-d1} \quad (4.43)$$

which, using (4.42), is

$$d\Phi_{2-d1} = \pi L_2 F_{d1-2} dA_1 \quad (4.44)$$

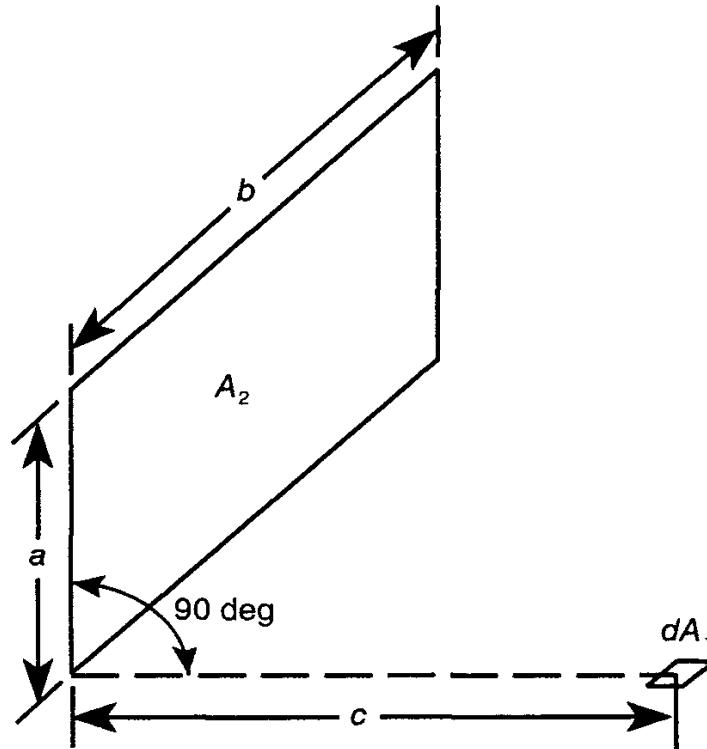


Figure 4.6 Geometry for configuration factor between a differential source element and a finite receiver plane perpendicular to the source area.

The illuminance at dA_1 is defined to be $d\Phi_{2-dA_1}/dA_1$, so that our solution to the problem is

$$E_2 = \pi L_2 F_{dA_1-dA_2} \quad (4.45)$$

with $F_{dA_1-dA_2}$ being given by (4.40). It is seen with this example that what would otherwise be a very difficult and mathematically complicated problem is rendered very straightforward through the use of configuration factors and the reciprocity relation.

REFERENCES

- [1] Siegel, R., and J. Howell. *Thermal Radiation Heat Transfer*, McGraw-Hill Book Co.: New York, NY, 1972.
- [2] Raeside, D. E., "An Introduction to Monte Carlo Methods," *Amer. J. Phys.*, Vol. 42, 1974, pp. 20–26.
- [3] Finn, G. D. "Probabilistic Radiative Transfer," *J. Quant. Spectrosc. Radiat. Transf.*, Vol. 11, 1971, pp. 203–222.
- [4] House, L. L., and L. W. Avery. "The Monte Carlo Technique Applied to Radiative Transfer," *J. Quant. Spectrosc. Radiat. Transf.*, Vol. 9, 1969, pp. 1579–1591.

Chapter 5

The Invariance of Radiance and the Limits of Optical Concentration

5.1 INTRODUCTION

This chapter begins the first serious discussion of the propagation of radiation through material media. In introducing new concepts, it is always easiest to limit the discussion initially to the simplest of cases. Thus we begin here with transparent media, such as idealized glass, that have no absorption, no scattering, and are perfectly isotropic in all respects; the principles of radiometry and photometry and optics are the same for all directions.

In dealing with flux transfer between surfaces separated by nonabsorbing, nonscattering, isotropic, nonrefracting media, as well as flux transfer in optical systems and other transparent or semitransparent media, the concept of radiance (or luminance) invariance is invaluable. Radiance is the most versatile of the basic definitions presented in Chapter 1. One reason for this is the invariance along rays in homogeneous media of a simple quantity that contains radiance. That invariance is demonstrated and its use in some important applications is described. In addition to introducing this important concept, readers following the derivation of radiance invariance should better understand the concepts of rays, pencils of rays, and the difference between idealized theoretical beams of radiation and real beams with nonzero angular spread and nonzero cross-sectional areas. As in the previous chapter, wherever the term radiance is used in this chapter, it is meant to refer to both radiance and luminance unless otherwise specified.

5.2 RADIANCE IS A FIELD QUANTITY

In general, radiance depends on the point in space where it is evaluated and on the direction through that point. It is in general a function of position in space and direction from that position. Radiance is a ray-associated field quantity [1] (see Figure 5.1). A ray of radiation is propagating to the right. Think of an element of

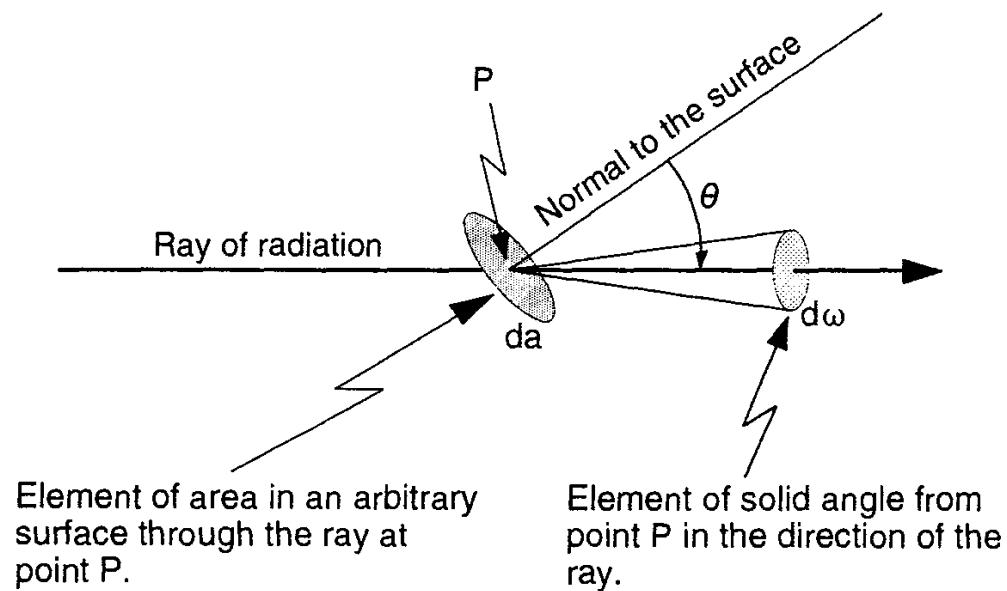


Figure 5.1 Geometry for definition of radiance along a ray.

area da at a point P of evaluation on this ray. The area can be oriented in any direction relative to the ray, its normal making angle θ with respect to the ray direction at point P. Also think of an element of solid angle from point P and in the direction of the ray through point P.

The radiance L at point P along this ray is defined by the equation

$$L(x, y, z, \theta, \phi) \equiv \frac{d^2\Phi(x, y, z, \theta, \phi)}{da \cos \theta d\omega} \quad (5.1)$$

where (x, y, z) are coordinates describing the position of the point P in space, (θ, ϕ) are coordinates describing the direction of the ray relative to the element of area da at point P, and $da \cos \theta$ is the element of projected area perpendicular to the ray direction (θ, ϕ) .

5.3 PENCILS OF RAYS

The next step in this treatment is to find a way to assign a meaning to the idea of an element of flux associated with a ray of radiation. The treatment of Nicodemus [1, 2] is followed, beginning with a couple of geometric constructions, shown in Figures 5.2 and 5.3. The first illustrates the definition of a pencil of rays, a set of all rays that intersect at a common point. The envelope of the pencil of rays can be the familiar shape of the right circular cone seen at the end of a well-sharpened

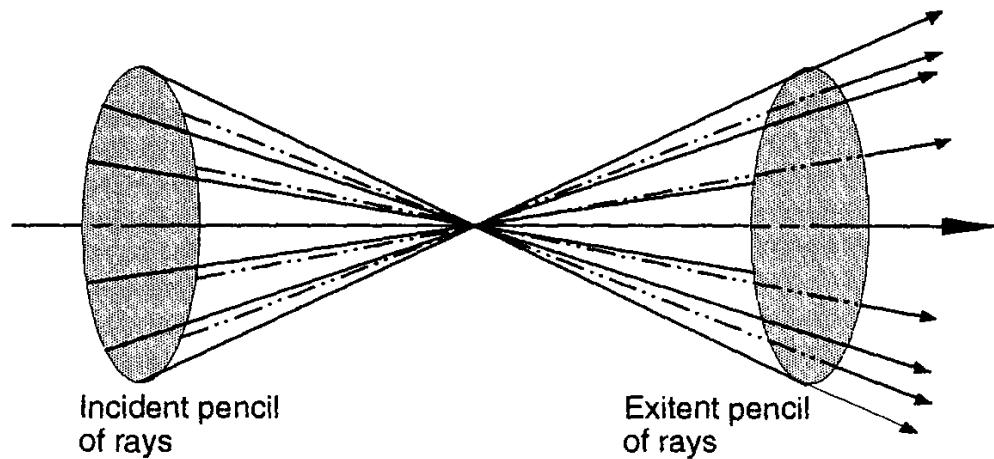


Figure 5.2 Illustration of pencils of rays.

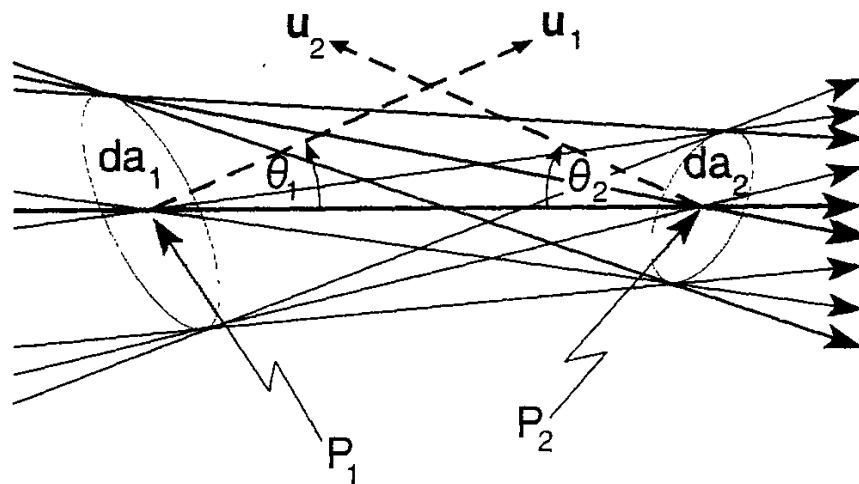


Figure 5.3 Illustration of an elementary beam of radiation.

wooden pencil. It can also be more complex than this, but only simple cone-shaped pencils of rays are dealt with here.

One speaks of an *incident* pencil of rays converging to a point or an *exitent* pencil of rays diverging from a point.

5.4 ELEMENTARY BEAM OF RADIATION

This next concept may be unfamiliar to many readers. It can be difficult both to explain and to understand. Perhaps some of the difficulties can be overcome with a short discussion of the reason the concept is needed. The primary reason for

defining an *elementary beam* of radiation is that it facilitates the elucidation of some mathematical subtleties important for precise discussions of flux propagation along a ray of radiation through a medium.

A truly mathematical ray of flux has zero solid angle and zero cross-sectional area associated with it. It can therefore carry no flux. In radiometry, one is primarily interested in the propagation of *finite, measurable* quantities of flux in space and through various media. A language (and the associated mathematical concepts) is required in order to associate an infinitesimally small amount of flux with any ray of radiation. The elementary beam concept facilitates this.

Figure 5.3 illustrates the considerations involved in the definition of an elementary beam of radiation. Shown in the figure is the collection of all rays leaving all points of one elemental area da_1 centered about point P_1 on a ray that intersect a second elemental area da_2 centered about another point P_2 on that same ray. In Figure 5.3, we have shown only nine of the rays in this elementary beam. They go from different parts of area da_1 to different parts of area da_2 . There would be an infinite number of such rays if elemental areas da_1 and da_2 were finite. The unit normal vectors \mathbf{u}_1 and \mathbf{u}_2 perpendicular to the elemental areas at points P_1 and P_2 are also shown in Fig. 5.3, as are the angles θ_1 and θ_2 between these normals and the central ray. (Those readers whose knowledge of and/or interest in differential calculus is limited may wish to skip ahead to the text beginning with (5.6) in Section 5.5 for the conclusion of the following treatment of radiance invariance. On the other hand, this derivation can enhance understanding of important concepts and the effort to understand it should thereby be rewarded.)

The gap between the idea of “perfect ray” not capable of carrying finite quantities of radiant or luminous flux and an “elementary beam of radiation” capable of carrying infinitesimally small quantities of flux is bridged as follows. The elementary beam from P_1 to P_2 is considered to be the collection of all ray directions along which radiation flows or is propagated between the two surface elements da_1 and da_2 . The elemental solid angle subtended by da_2 at point P_1 is given by

$$d\omega_{12} = \frac{\cos \theta_2 da_2}{R^2} \quad (5.2)$$

and the solid angle subtended by da_1 at P_2 is

$$d\omega_{21} = \frac{\cos \theta_1 da_1}{R^2} \quad (5.3)$$

where R is the distance from P_1 to P_2 . Integrating $d\omega_{12}$ and $d\omega_{21}$ over finite areas results in finite solid angles that can contain finite quantities of flux.

5.5 RADIANCE INVARIANCE

Because the elements of area da_1 and da_2 are infinitesimal, there is no significant difference in radiance L among any of the rays in the elementary beam or among any of the rays within any of the pencils of rays diverging from any point in da_1 or converging to any point in da_2 . Thus, we may assume that all of the rays leaving da_1 toward da_2 are of radiance L_1 and that those same rays all arrive at da_2 with radiance L_2 .

The flux in the exitant elementary beam leaving da_1 is given by

$$d^2\Phi_1 = L_1 \frac{\cos \theta_1 da_1 \cos \theta_2 da_2}{R^2} \quad (5.4)$$

Similarly, the flux in the incident elementary beam arriving at da_2 is given by

$$d^2\Phi_2 = L_2 \frac{\cos \theta_2 da_2 \cos \theta_1 da_1}{R^2} \quad (5.5)$$

If there are no losses in the medium between da_1 and da_2 , then, by the principle of conservation of energy, we must have $d^2\Phi_1 = d^2\Phi_2$. This means that

$$L_1 = L_2 \quad (5.6)$$

Because no restriction was placed on where P_1 and P_2 are located along the ray, (5.6) must apply to any pair of points along a ray in a lossless, isotropic, uniform medium. This is called *radiance invariance*. It is true also of luminance. The radiance or luminance is constant along a ray of propagation within a homogeneous, lossless, isotropic medium.

This invariance of radiance along rays from source to receiver through an isotropic, nonabsorbing, and nonscattering medium was implicitly imbedded in the discussion of Chapter 4. There are numerous applications in radiometry and photometry in which having radiance invariance available as a general principle is enormously helpful. For example, if we want to know the irradiance or illuminance of a real image of an object formed by a lens, then all we have to know is how the radiance or luminance varies across the object and across the solid angle subtended by the lens at each point of the object and we shall know the radiance/luminance at every point in the image, if we can determine the effect the lens has on our principle of radiance invariance. This is discussed in the next two sections.

5.6 RADIANCE INVARIANCE AT AN INTERFACE

This discussion begins with an analysis of radiance along a ray crossing a smooth (specular) interface between two media of differing refractive indices. Again the derivation of Nicodemus [1] is followed. The situation is as shown in Figure 5.4. The angles of incidence θ_1 and refraction θ_2 are shown for a ray incident upon an infinitesimal area da at an interface between two media of differing refractive indices n_1 and n_2 .

If ρ is the reflectance for the given angle of incidence at the interface and it is assumed that there is no absorption at the interface, then $(1 - \rho)L_1$ is the portion of the incident radiance L_1 that enters the second medium. The element of flux incident on da in the element of incident solid angle $d\omega_1$ is $d\Phi_1$ and $(1 - \rho)d\Phi_1$ is the portion of this flux that enters the second medium. The refracted elementary beam fills element $d\omega_2$ of solid angle, which differs slightly from $d\omega_1$ because the rays bounding it are refracted by slightly different amounts owing to their slightly different angles of incidence. The radiance of the refracted ray is L_2 and the elements of flux can be written in terms of the radiances as follows:

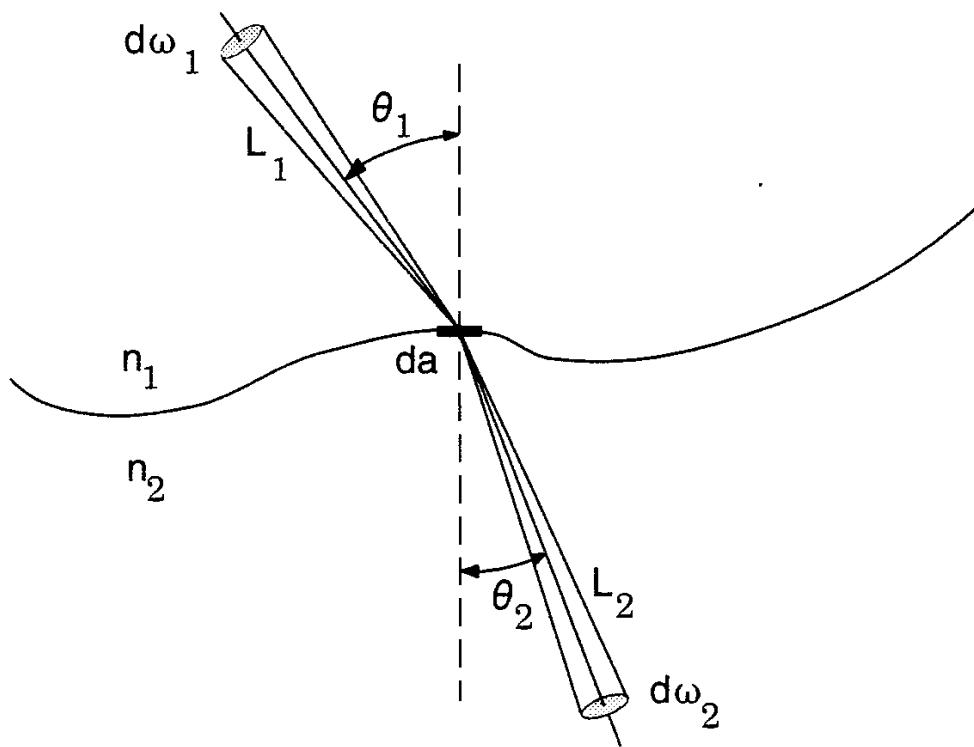


Figure 5.4 Refraction at a smooth interface. After: Nicodemus [1].

$$d\Phi_1 = L_1 \cos \theta_1 da d\omega_1 = L_1 da \cos \theta_1 \sin \theta_1 d\theta_1 d\phi \quad (5.7)$$

and

$$d\Phi_2 = L_2 \cos \theta_2 da d\omega_2 = L_2 da \cos \theta_2 \sin \theta_2 d\theta_2 d\phi \quad (5.8)$$

The element $(1 - \rho)d\Phi_1$ of incident flux that makes it into the second medium is given by

$$(1 - \rho)d\Phi_1 = (1 - \rho)L_1 \cos \theta_1 da d\omega_1 = (1 - \rho)L_1 da \cos \theta_1 \sin \theta_1 d\theta_1 d\phi \quad (5.9)$$

For conservation of energy (5.8) equals (5.9) so that their ratio is unity:

$$\frac{(1 - \rho)L_1 \cos \theta_1 \sin \theta_1 d\theta_1}{L_2 \cos \theta_2 \sin \theta_2 d\theta_2} = 1 \quad (5.10)$$

Now, when a ray crosses an interface between media of differing refractive indices n_1 and n_2 , the angle of incidence θ_1 and the angle of refraction θ_2 are related by Snell's law:

$$n_1 \sin \theta_1 = n_2 \sin \theta_2 \quad (5.11)$$

Differentiating both sides of this expression with respect to angle yields

$$n_1 \cos \theta_1 d\theta_1 = n_2 \cos \theta_2 d\theta_2 \quad (5.12)$$

so that

$$\frac{\sin \theta_1}{\sin \theta_2} = \frac{\cos \theta_1 d\theta_1}{\cos \theta_2 d\theta_2} = \frac{n_2}{n_1} \quad (5.13)$$

Combining (5.13) with (5.10) yields

$$\frac{(1 - \rho)L_1}{n_1^2} = \frac{L_2}{n_2^2} \quad (5.14)$$

If we define L'_1 to be the incident radiance after the reflected portion has been taken away, that is

$$L'_i = (1 - \rho)L_i \quad (5.15)$$

then we may write (5.14) as

$$\frac{L'_i}{n_i^2} = \frac{L_2}{n_2^2} \quad (5.16)$$

We draw an important conclusion from this. When a ray crosses an interface between two media of differing refractive indices, ignoring reflection losses at the interface, the radiance is altered somewhat, but the quantity L/n^2 , which Nicodemus calls *basic radiance*, remains invariant. If reflection losses are not ignored, (5.14) can be used to relate the radiance L_2 emerging from the interface to that L_i incident upon it.

According to Nicodemus [1] “more sophisticated proofs show that the invariance of basic radiance is a completely general geometric property, even along a ray traversing a nonuniform, nonisotropic medium in which the index of refraction varies continuously from point to point.” A consequence of this is that a ray entering a medium of different refractive index will have its radiance altered according to (5.16), but upon emerging back into the original medium its original radiance will be restored, neglecting absorption, scattering, and reflection losses.

5.7 RADIANCE THROUGH A LENS

Now the case of radiance through a lens is addressed again. Since the radiance that is altered according to (5.16) when a ray enters a lens has its radiance restored when it emerges into the original medium later, we see that the radiance associated with every ray contributing to a point in an image is the same as when that ray left the object on the other side of the lens (ignoring reflection and transmission losses in the lens). Since this is true of all rays making up an image point, the radiance of the image formed by a perfect lens equals the radiance of the object (source), apart from absorption, scattering, and reflection losses. Including these losses, we can define the *radiance transmittance T* as the ratio of exitant to incident radiance along a ray through the lens. Assuming the radiance transmittance to be the same for all rays, then the radiance L_i of the image is

$$L_i = TL_o \quad (5.17)$$

for corresponding pairs of directions from an object point to the corresponding image point. L_o is the radiance of the object. Chapter 8 contains a more detailed discussion of radiance and irradiance in simple optical systems.

5.8 RADIANCE IN ABSORBING AND SCATTERING MEDIA

There is another interesting application of radiance invariance: the case of radiation propagation in absorbing and scattering media. This application is called *radiative transfer* and a considerable body of theory has been developed in this field—for dealing with radiation in planetary atmospheres, in bodies of water, in stars, and in a variety of other media, including biological tissues. If we think of a single ray propagating through such a medium, its radiance will vary only as the refractive index varies if there are no sources or losses along its path. The radiance will *decrease* as energy is removed by absorption by the medium and by scattering away from the ray by particles in the medium and particles (atoms and molecules) of the medium itself. The radiance will *increase* by scattering into this ray of radiation from other rays that cross this ray. The net result can be either a steady increase or a decrease in the radiance of a ray, or a constancy of the ray's radiance along the path, depending upon the nature of the absorption and scattering characteristics of the medium.

The radiance can also increase if the medium is self-luminous, as in a star, in a medium exhibiting fluorescence, or in an electrical plasma (an ionized gas emitting radiation). The portion of the radiance of a ray that is increasing is called *path radiance* (because it is the radiance due to sources along the path itself rather than of some object or source of interest). Path radiance can reduce the contrast of images formed in or through scattering media.

If the medium is sufficiently thin, then one may make an assumption that only single scattering takes place, and the radiative transfer equations are simplified. If the medium is dense with scattering particles, then the radiation field is made up of many multiple scatterings and the problem is very difficult to solve with conventional approaches.

Radiative transfer is in general a very difficult and mathematically involved topic. The fundamental processes are frequently described by combined integral and differential equations that are very difficult or impossible to solve in closed form without simplifying assumptions. In many cases, to get enough simplifying assumptions to make the problem tractable, the problem is so oversimplified that the solution is meaningless.

The concepts of radiometry and photometry are very important for radiative transfer studies, in such diverse fields as hydrological, atmospheric, and ocean optics, as well as for studies of radiation propagation in planetary atmospheres and stars, in turbid solids including biological tissues, and in particulate suspensions in chemistry.

For more information about radiative transfer, the reader is referred to several excellent texts on the subject [3–6].

5.9 CONCENTRATING RADIANCE METER

In Problem 1.2 in Chapter 1, it was shown how a radiance meter can be made out of a detector placed at the bottom of a tube with a well-defined aperture at the top. The arrangement has been given the name “Gershun tube.” In many applications, the source radiance is very faint, and the flux detection system (the detector and its associated electronics) may have difficulty distinguishing the signal to be measured from background noise. This problem will be discussed more fully in Chapter 7 on detectors and Chapter 8 on measurement systems. One way to overcome the problem, however, is to increase the signal level, by designing the meter so that more flux reaches the detector. The material presented previously will now be used to address this design problem, as an illustration of the use of (5.16) and related concepts. The limits to the flux gathering power that can be obtained with such a design will be explored as an introduction to the interesting question: What is the maximum concentration possible with optical systems? The next section examines more generally the problem of concentrating radiant flux and discusses the ultimate limits to this concentration that can be achieved with optical systems.

The problem with the design used in Problem 1.2 is that as the aperture area A_a is increased to accept more flux, over a larger solid angle, the area A_t of the target increases along with the solid angle of the field of view of the instrument. In many measurement situations this increased field of view is undesirable. The field of view needs to be kept small while still gathering adequate quantities of flux. In most measurements of radiance, high angular resolution is desired, meaning that the meter should not average over too large a solid angle. The solid angle field of view might need to be limited to a cone whose half-angle is below a few degrees, for example. One needs a way of increasing the flux on the detector without increasing the field of view.

An obvious choice is to place a telescopic lens in the entrance aperture A_a of Figure 1.14, choosing a lens whose focal length f equals the distance X to the detector D . The resulting new radiance meter design is shown schematically in Figure 5.5. Points on the detector are in the focal plane of the lens and therefore receive (in the perfect lens assumption) rays that are parallel to each other before entering the lens and converging to the detection point. These rays are indicated by the solid and dashed lines in Fig. 5.5.

Looking at the figure, it is seen that for a perfect lens or other imaging system, there is no angular spread in the rays approaching the lens that reach P . As a result, the solid angle field of view of the meter corresponding to detector point P is zero.

Contrast this setup with the lensless radiance meter of Figure 1.14, where the solid angle field of view seen by a point at the center of the detector is defined by the distance X and the area A_a of the aperture. Increasing A_a to gather more flux

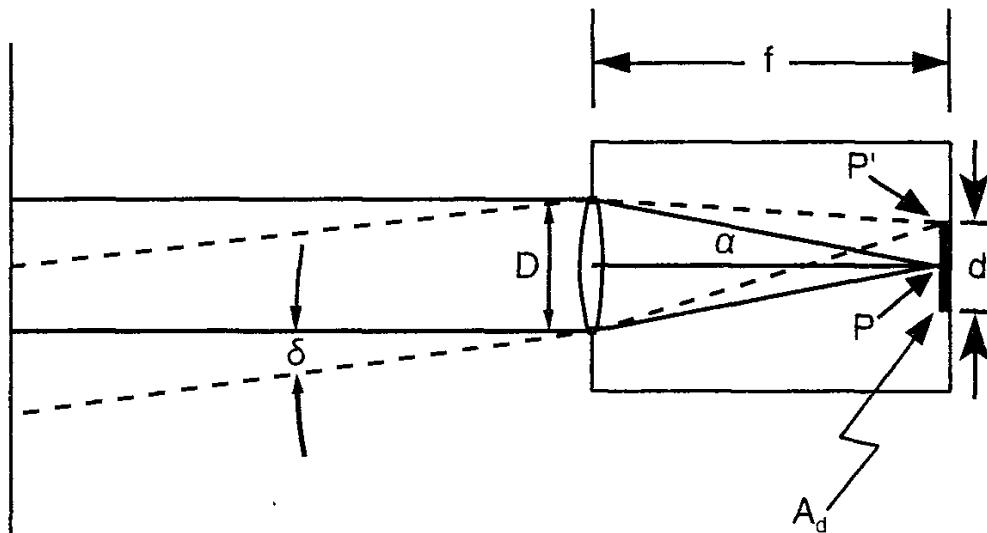


Figure 5.5 Conceptual diagram of a Gershun tube radiance meter with a lens added.

also increases the angular field of view. In the new case of the telescopic radiance meter depicted in Figure 5.5, as the lens area A_a is increased, the number of parallel rays entering the lens increases, but the angular field of view does not increase.

Thus, the greater the lens diameter D (keeping its focal length constant), the greater the concentration of rays reaching the center of the detector. It appears from this that the new design meets the stated objective of increasing the flux without increasing the field of view. There is a small conceptual problem, however. A geometrical point in a detector cannot receive a measurable quantity of flux. There is in reality no measurable flux in an infinitesimally small solid angle field of view. This difficulty is resolved by looking at some other points in the detector, such as point P' at the detector's edge shown in Figure 5.5. Peripheral rays reaching this point are indicated by the dashed lines. It is seen that the effect of a finite detector area is to expand the field of view by the half-angle δ (called the *acceptance angle*) subtended by the detector radius at the center of the lens of focal length f . Thus,

$$\delta = \tan^{-1} \frac{d}{2f} \quad (5.18)$$

where d is the detector diameter. (It is assumed that the lens aperture is circular, with diameter D , as is the detector, but with diameter d .) The full angular extent of the field of view is 2δ . Because the detector area is finite, it can receive finite, measurable quantities of flux and the radiance meter can measure finite quantities

of flux over a finite solid angle field of view Ω_d and our conceptual problem is avoided.

In order to analyze the performance of the new radiance meter, an equation is needed for the total flux Φ_d reaching the detector from the defined solid angle field of view Ω_d and lens aperture area A_a . It will be helpful to look at the drawing shown in Figure 5.6. For a thin lens it is known that rays passing through the center of the lens are undeviated in direction of propagation. We see from the construction in Figure 5.6 that the solid angle field of view Ω_d of the radiance meter is given by (1.5) in Chapter 1, with α being replaced by δ as given in (5.18). The solid angle field of view of the “lensed” radiance meter is determined *not* by the size of the lens aperture but by the size of the detector in relation to the focal length of the lens.

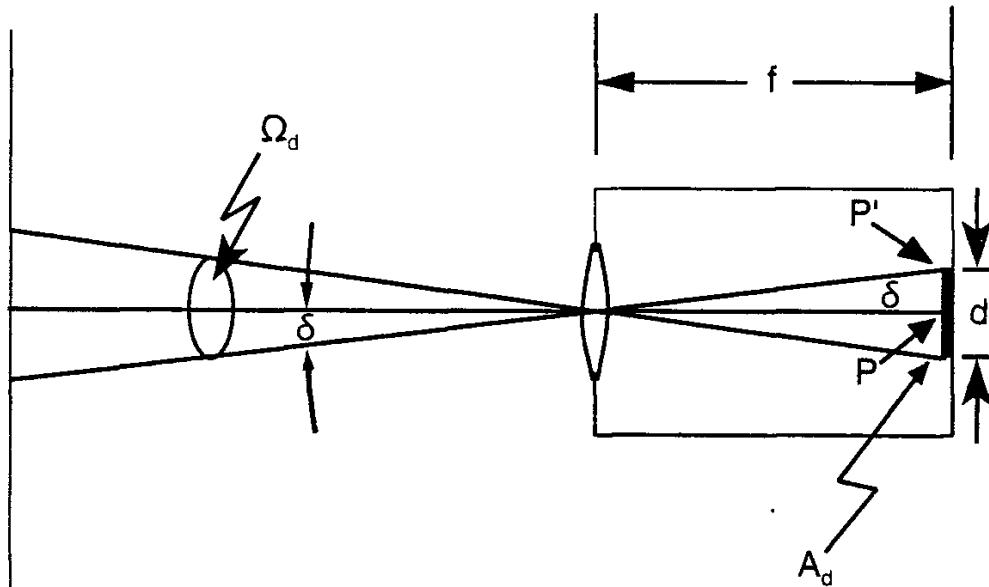


Figure 5.6 Diagram of radiance meter showing the relationship of the solid angle field of view Ω_d to the diameter d of the detector.

Let L_1 be the radiance incident upon the lens, assumed constant over A_a , whose average over Ω_d we wish to determine from the flux Φ_d measured by the detector. Solving (5.16) for L_2 , setting n_1/n_2 equal to 1 because both sides of the lens are the same medium, and replacing L'_1 with $T_L L_1$, where T_L is the transmittance of the lens, one obtains

$$L_2 = T_L L_1 \quad (5.19)$$

Next, (4.22) is used to get the irradiance E_d on the detector due to radiance L_2 emerging from the lens toward the detector:

$$E_d = \pi L_2 \sin^2 \alpha \quad (5.20)$$

with

$$\alpha = \tan^{-1} \frac{D}{2f} \quad (5.21)$$

where D is the lens aperture diameter and f is its focal length.

Combining (5.19) and (5.20) yields

$$E_d = (\pi T_L \sin^2 \alpha) L_1 \quad (5.22)$$

with total flux on the detector given by

$$\Phi_d = E_d A_d = (\pi A_d T_L \sin^2 \alpha) L_1 \quad (5.23)$$

and the solid angle field of view of the radiance meter given by

$$\Omega_d = 2\pi(1 - \cos \delta) \quad (5.24)$$

$$\Omega_d = 2\pi \left[1 - \cos \left(\tan^{-1} \frac{d}{2f} \right) \right]$$

Solving (5.23) for L_1 , we have

$$L_1 = \frac{\Phi_d}{\pi A_d T_L \sin^2 \alpha} \quad (5.25)$$

The detector and its associated electronics produces an electrical signal proportional to Φ_d and hence to L_1 . Knowing the value of the constant in the denominator on the right side of (5.25), one can determine the radiance entering the lens from the measured flux Φ_d on the detector. The calibration of radiometers is discussed in Chapter 9.

The ratio f/D , where D is the lens aperture diameter, is called the *f-number* or *F/#* of the lens. A lens is said to be *fast* (meaning it has more flux collecting power and produces shorter exposure times on photographic films) when this number is small. Thus the *speed* of a lens is inversely proportional to its *f-number*.

Returning now to the problem of increasing the flux on a radiance meter's detector, as the lens area increases, the solid angle field of view Ω_d of the scene to the left of the lens shown in Figure 5.6 remains unchanged, but the solid angle of flux incident on each point of the detector to the right of the lens as shown in Figure 5.5 increases. If the lens diameter is taken to infinity (and its f-number to zero), the solid angle incident upon point P increases to 2π sr. This gives the theoretical maximum possible (but practically unrealizable) concentration of flux on the detector.

In this case, $\sin^2 \alpha$ in (5.23) becomes 1.0 and the theoretical maximum flux $\Phi_{d\max}$ one can send to the detector for a given detector diameter, and hence solid angle field of view, is given by

$$\begin{aligned}\Phi_{d\max} &= \pi A_d T_L L_1 = \pi f^2 \Omega_d T_L L_1 \\ &= \pi^2 \left(\frac{d}{2}\right)^2 T_L L_1 = \pi^2 f^2 \tan^2 \delta T_L L_1\end{aligned}\quad (5.26)$$

using (5.18). This shows the direct relationship between the size of the field of view of a lensed radiance meter and the maximum theoretical flux concentration.

5.10 THE LIMITS OF OPTICAL CONCENTRATION

Examine the generalized situation diagrammed in Figure 5.7. It shows an idealized optical system with no absorption, reflection, and scattering losses, having circular entrance aperture area A_1 that takes rays distributed over the solid angle Ω_2 having cone half-angle δ and focuses them onto circular receiving area A_2 at the focus of an idealized imaging system of focal length f . The solid angle of rays from the lens approaching the center of the receiving area is Ω_1 . Let Φ_1 be the flux through the optical system within solid angle Ω_2 . Let Φ_2 be the flux from the lens reaching the receiver. An idealized system was specified, so that $\Phi_1 = \Phi_2$. Let L_1 be the radiance incident on the optical system from the left, in a medium of refractive index n_1 and L_2 be the radiance of rays from the optical system to the receiver in medium of refractive index n_2 .

From (4.21), with θ_o replaced by α , it is known that the irradiance E_2 on the receiver is given by

$$E_2 = \pi L_2 \sin^2 \alpha \quad (5.27)$$

Similarly, removing the radiance from the integral in (4.10) and integrating over

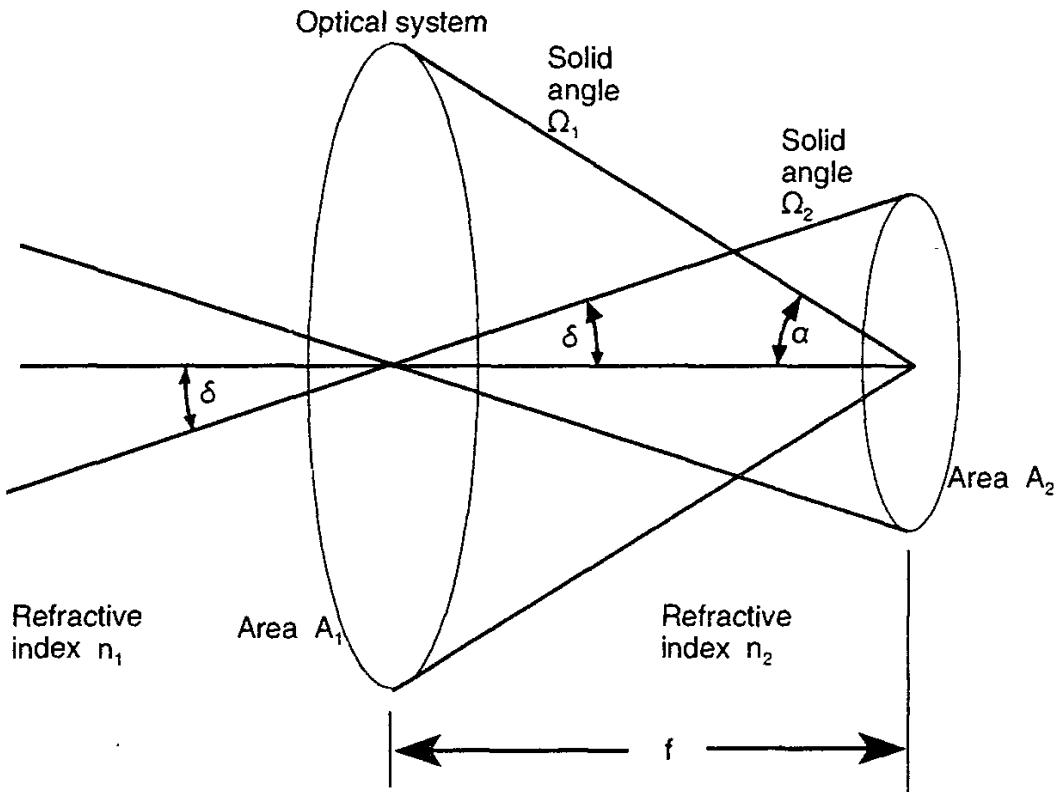


Figure 5.7 Diagram of idealized optical system for determining the concentration ratio.

the solid angle Ω_2 gives the following expression for the irradiance E_1 incident on the optical system due to flux within Ω_2 :

$$\begin{aligned}
 E_1 &= \int_{\Omega_2} L_1 \cos \theta d\omega \\
 E_1 &= 2\pi \int_0^\delta L_1 \cos \theta \sin \theta d\theta \\
 E_1 &= \pi L_1 \sin^2 \delta
 \end{aligned} \tag{5.28}$$

The concentration ratio C is defined to be the ratio of the receiver irradiance E_2 to the incident irradiance E_1 . Accordingly one obtains

$$C = \frac{E_2}{E_1} = \frac{L_2 \sin^2 \alpha}{L_1 \sin^2 \delta} \tag{5.29}$$

Replacing L'_1 with L_1 in (5.16), according to our assumption of no losses due to

reflection, absorption, or scattering, and using the resulting equation in (5.29) yields the result

$$C = \frac{n_2^2 \sin^2 \alpha}{n_1^2 \sin^2 \delta} \quad (5.30)$$

If the medium to the left of the optical system is air with refractive index nearly equal to 1.00, then the final expression for the concentration ratio is

$$C = \frac{n_2^2 \sin^2 \alpha}{\sin^2 \delta} \quad (5.31)$$

The maximum concentration of flux on the receiver will occur when the size of the optical system entrance aperture area increases to infinity, causing the angle α to increase to 90 deg so that $\sin \alpha$ goes to 1.0. Thus the theoretical maximum concentration ratio C_{\max} will be given by

$$C_{\max} = \frac{n_2^2}{\sin^2 \delta} \quad (5.32)$$

Since $E = \Phi/A$, we have

$$C = \frac{E_2}{E_1} = \frac{A_1}{A_2} \quad (5.33)$$

showing the definition of concentration ratio in terms of optical system entrance and exit apertures that is used by many authors.¹

Another example can be suggested. Suppose one has a sensor in air that produces a noise-free signal proportional to the incident illuminance E_v only for illuminance levels above 1 lux. Equation (5.32) can be used to determine the

¹Roland Winston's group at the University of Chicago in 1988 used an oil-filled concentrator with a refractive index of 1.53 to achieve a record concentration of sunlight to 56,000 times the irradiance at the earth's surface, or 44 W/mm², over half the theoretical maximum computed in Example Problem 5.2, assuming a terrestrial solar irradiance of 1000 W/m². This was reported by Gleckman in 1988 [7]. In 1990, a sapphire concentrator with $n = 1.76$ employing total internal reflection was used to increase concentration to 84,000, or 72 W/mm², 15% greater than the irradiance at the surface of the sun itself [8–9]. In Example Problem 5.1, we calculate the average radiance over the solar disk to be about $19.8 \times 10^6 \text{ W} \cdot \text{m}^{-2} \cdot \text{sr}^{-1}$. Assuming a Lambertian surface and multiplying this by π yields an approximate value for the irradiance of the sun's surface: 62 W/mm², in good agreement with 15% lower than the figure quoted by Winston [7].

minimum detectable illuminance we could achieve with an acceptance half-angle of 10 deg. The answer is 33 lux.

Clearly, as the acceptance angle δ decreases, the theoretical maximum concentration ratio increases. At the opposite extreme, the limiting case of $\delta = 90$ deg, the maximum concentration is 1.0. This is another way of saying that you cannot concentrate totally diffuse radiation (a hemisphere of isotropic flux onto a planar surface). There is one exception to this rule, arising from the presence of n_2^2 in (5.32). Even if $\sin^2 \delta = 1$ in this equation, n_2^2 can exceed unity. Glass with a refractive index of 1.5, for example, could provide a concentration ratio of $1.5^2 = 2.25$, if one could avoid the reflection losses at the interface between air and glass. Ries [10] offers a comprehensive review of the thermodynamic limitations to maximum achievable concentrations of electromagnetic radiation.

Equation (5.32) is a general result. It has been derived by other authors using the principles of thermodynamics [10–12] and the conservation of *etendue* or “throughput.” [13] Both imaging and nonimaging concentrators of radiation are useful in a wide variety of fields, including high-energy physics, infrared astronomy [14], solar energy, the optics of visual receptors, stray-radiation shields, and the optical pumping of solid-state lasers. It is interesting to note that the cone receptors in the human eye, discussed in Chapter 2, have a shape corresponding approximately to that of a nonimaging concentrator designed to collect light coming from the solid angle subtended by the fully dilated iris [15].

A considerable body of literature exists on the design of nonimaging concentrators for solar energy collection [16].

EXAMPLE PROBLEM 5.1

Problem: The irradiance from the sun at the mean radius of the earth’s orbit around the sun is $E_e = 1353 \text{ W/m}^2$ (also called the *solar constant*). The sun subtends a diameter of about 0.535 deg at the earth. Assuming the sun to be a circular disk facing the earth, what is the average radiance L_s of the sun over the solar disk? If the sun were a perfect black body at 6,000K, what would be the approximate solar constant?

Solution: The half-angle of the sun is 0.2675 deg. Using (1.5) in Chapter 1, the sun subtends a solid angle of $\Omega_s = 6.85 \times 10^{-5} \text{ sr}$, or 68.5 microsteradians. Set $\theta = 0$ and $\psi = 0$ in (4.6). Convert the integral over source area S_o to an integral over the solid angle Ω_s of the solar disk, replacing ds_o/R^2 with $d\omega_s$, an element of solid angle subtended by an elemental area ds_o on the solar disk:

$$E_e = \int_{\Omega_s} L_s d\omega_s \quad (5.34)$$

where L_s is the radiance from a point on the solar disk in the earth's direction. Since it is the average radiance L_s over the solid angle subtended by the solar disk that is desired, L_s can be removed from the integral and the result solved for L_s to obtain

$$L_s = \frac{E_e}{\Omega_s} \quad (5.35)$$

Dividing the extraterrestrial irradiance $E_e = 1353 \text{ W/m}^2$ by the sun's solid angle Ω_s yields the average radiance of the sun

$$L_s = 19.8 \times 10^6 \text{ W} \cdot \text{m}^{-2} \cdot \text{sr}^{-1} \quad (5.36)$$

as viewed from the earth's orbit. This value is the approximate one to use for all directions within the solid angle Ω_s subtended by the sun at the point of interest in the orbiting plane. The radiance is zero outside this solid angle (ignoring the relatively small amount of radiance emanating from the solar corona, which extends out some distance from the photosphere, the astronomical name for the solar disk).

Turning now to the problem of the sun as a blackbody, (5.35) is solved for the irradiance E_{bbe} near the earth:

$$E_{bbe} = L_{bbs} \Omega_s \quad (5.37)$$

in terms of the radiance L_{bbs} of the sun as a blackbody at 6,000K. In Problem 3.1, the radiance of a 6,000K blackbody is calculated to be $23.4 \times 10^6 \text{ W} \cdot \text{sr}^{-1} \cdot \text{m}^{-2}$. Multiplying this by the solid angle of the sun yields $1,603 \text{ W/m}^2$ for E_e , reasonably close to the solar constant to validate the approximation of the sun as a blackbody source at 6,000K.

Before leaving this problem, it is interesting to determine the temperature T_{bbs} of the sun as a blackbody that gives precisely the known solar constant and its corresponding average solar radiance. Solving (3.11) for T in terms of $L_s = 19.8 \times 10^6 \text{ W} \cdot \text{m}^{-2} \cdot \text{sr}^{-1}$ from (5.36) yields a temperature of 5,755K.

EXAMPLE PROBLEM 5.2

Problem: Let δ be the half-angle of the sun, 0.268 deg. Let there be an irradiance detector inside water having a refractive index of 1.33. Calculate the theoretical maximum concentration ratio for this case if the terrestrial solar direct beam irradiance is 1000 W/m^2 .

Solution: From (5.32), the concentration ratio is $1.33^2/\sin^2(0.268 \text{ deg}) = 80,850$. Taking the $1,000 \text{ W/m}^2$ irradiance value, it is seen that the maximum

irradiance one could hope to achieve in these circumstances is about 81,000 times this, 81 MW/m² or 81W/mm².

REFERENCES

- [1] Nicodemus, F. E. "Radiance," *Am. J. Phys.*, Vol. 31 (1963), pp. 368-377.
- [2] Nicodemus, F. E. *Self-Study Manual on Optical Radiation Measurements, Part 1—Concepts*, Chs. 1-3, NBS Technical Note 910-1, U. S. Department of Commerce, National Bureau of Standards (now National Institute of Standards and Technology), March 1976.
- [3] Chandrasekhar, S. *Radiative Transfer*, Dover Publications: New York, NY, 1960.
- [4] Preisendorfer, R. W. *Radiative Transfer on Discrete Spaces*, Pergamon Press: New York, NY, 1965.
- [5] Kattawar, G. W., and G. N. Plass. "Radiative Transfer in the Earth's Atmosphere Ocean System: II. Radiance in the Atmosphere and Ocean," *Jour. Phys. Oceanog.*, Vol. 2, 1972, p. 46.
- [6] Plass, G. N., and G. W. Kattawar. *Appl. Opt.*, Vol. 8, 1969, p. 455.
- [7] Gleckman, P. *Appl. Opt.*, Vol. 27, 1988, p. 4385.
- [8] Cooke, D., et al. *Nature*, Vol. 346, 1990, p. 802.
- [9] Winston, R., et al. "Ultra-High Solar Flux and Applications to Laser Pumping," *Solar '93, Proceedings of the 1993 Annual Conference, American Solar Energy Society*, Washington, DC, April 25-28 1993, S. Burley and M. E. Arden, eds.
- [10] Ries, H. "Thermodynamic Limitations of the Concentration of Electromagnetic Radiation," *J. Opt. Soc. Am.*, Vol. 72, 1982, p. 380.
- [11] Winston, R., "Light Collection Within the Framework of Geometrical Optics," *J. Opt. Soc. Am.*, Vol. 60, 1970, p. 245.
- [12] Rabl, A. "Comparison of Solar Concentrators," *Solar Energy*, Vol. 18, 1976, p. 93.
- [13] Welford, W. T., and R. Winston. *The Optics of Nonimaging Concentrators*, Academic Press: New York, NY, 1978.
- [14] Rabl, A., and R. Winston. "Ideal Concentrators for Finite Sources and Restricted Exit Angles," *Applied Optics*, Vol. 15, 1976, p. 2880.
- [15] Winston, R., and J. M. Enoch. "Retinal Cone Receptor as an Ideal Light Collector," *J. Opt. Soc. Am.*, Vol. 61, 1971, p. 1120.
- [16] Welford, W. T., and R. Winston. *High Collection Nonimaging Optics*, Academic Press: New York, NY, 1989.

Chapter 6

Optical Properties of Materials

6.1 INTRODUCTION

Central to radiometry and photometry is the interaction of radiation with matter. One aspect of this concern is dealt with in Chapter 3 on blackbody radiation. Others are found in Section 4.4 and Sections 5.6 through 5.8. The discussion of blackbody radiation is confined to the spectral emissive properties of a particular type of material, having perfect absorption and no reflection (a blackbody). The remaining deal with materials having other idealized properties that simplify calculations.

This chapter begins a discussion of the properties of real materials and their abilities to emit, reflect, refract, absorb, transmit, and scatter radiation. This is a large field. Only the rudiments of the subject can be addressed here, dealing mostly with terminology and basic concepts. Some representative (if synthetic) applications will be dealt with in the example problems at the end of the chapter. For more information on the optical properties of matter, the reader is directed to available texts on optics and optical engineering, as well as other literature on material properties.

The treatment of material optical properties in this chapter begins with a discussion of the optical properties of interfaces. Next is discussed the interior optical properties of transparent or translucent materials, away from interfaces. Then interfaces and interior properties are combined into a description of the properties of complete objects, macroscopically homogeneous materials having plane parallel interfaces to air, to vacuum, or to another transparent medium such as water.

This chapter also addresses the directional dependence of the optical properties of plane parallel plates of materials, transparent and semitransparent, and the surfaces of opaque materials, discussing in some detail the solid angles associated with various optical property definitions. Then come discussions of broadband directional properties, spectral dependences, broadband spectral properties, and a separate section on spectral selectivity.

6.2 TERMINOLOGY

As mentioned in the Preface and Section 1.2, the terminology of radiometry and photometry is approaching a state of uniformity, thanks in part to the *International Lighting Vocabulary* [1] published jointly by the International Lighting Commission (*Internationale de l'Eclairage*, or CIE) and the International Electrotechnical Commission (*Commission Electrotechnique Internationale*, or CEI). This trend has been extended to the subdiscipline dealing with the optical properties of materials, and a proper terminology can now be identified for the processes of reflection, transmission, and emission of radiant flux by or through material media. Although symbols have been standardized for most of these properties, there are a few exceptions, discussed in this section.

To begin the discussion, the CIE definitions [1] for *reflectance*, *transmittance*, and *absorptance* are provided:

Reflectance (for incident radiation of a given spectral composition, polarization and geometrical distribution) (ρ): Ratio of the reflected radiant or luminous flux to the incident flux in the given conditions (unit: 1).

Transmittance (for incident radiation of given spectral composition, polarization and geometrical distribution) (τ): Ratio of the transmitted radiant or luminous flux to the incident flux in the given conditions (unit: 1).

Absorptance (α): Ratio of the absorbed radiant or luminous flux to the incident flux under specified conditions (unit: 1) [1].

These definitions make explicit the point that radiation incident upon a surface can have nonconstant distributions over the directions of incidence, over polarization state, and over wavelength (and frequency). Thus, when one wishes to measure the optical properties named above, it must be specified how the incident radiation is distributed in wavelength and direction for the measurement to make any sense. For the moment, we suspend discussion of polarization effects. The wavelength dependence of radiometric properties of materials is indicated with a functional lambda λ thus: $\tau(\lambda)$, $\rho(\lambda)$, and $\alpha(\lambda)$. The directional dependencies are indicated by specifying the spherical angular coordinates of the incident and emergent beams of radiation. The angular and spectral distributions of these optical properties are discussed below in some detail. For now, only the notational conventions are addressed, after which the concepts of greybodies and nonblackbodies, postponed from Chapter 3, are introduced.

As pointed out by Siegel and Howell [2], in other fields it is common to assign the ending *-ivity* to *intensive*, inherent, or bulk properties of materials. The ending *-ance* is reserved for the *extensive* properties of a fixed quantity of substance, for

example a portion of the substance having a certain length or thickness. Sometimes the term *intrinsic* is used instead of intensive and *extrinsic* instead of extensive. An example of this terminology is that of the *resistance* of 3 cm of a conductor having a *resistivity* of 10 ohms per cm. The resistance of the 3 cm length is 30 ohms. According to this usage, *reflectance* is reserved for the fraction of incident flux reflected (under defined conditions of illumination and reception) from a finite and specified portion of material, such as a 1-cm thickness of fused silica glass with parallel, sand-blasted (roughened) surfaces in air. The *reflectivity* of a material, such as BK7 glass, would refer to the ratio of reflected to incident flux for the perfectly smooth (polished) *interface* between an optically infinite thickness of the material and some other material, such as air or vacuum. The infinite thickness is specified to insure that reflected flux from no other interface can contribute to that reflected by the interface of interest.

The CIE *International Lighting Vocabulary* offers the following definitions of reflectivity, spectral transmissivity, and spectral absorptivity:

Reflectivity (of a material) (ρ_∞): Reflectance of a layer of the material of such a thickness that there is no change of reflectance with increase in thickness (unit: 1).

Spectral transmissivity (of an absorbing material) ($\tau_{i,o}(\lambda)$): Spectral internal transmittance of a layer of the material such that the path of the radiation is of unit length, and under conditions in which the boundary of the material has no influence (unit: 1).

Spectral absorptivity (of an absorbing material) ($\alpha_{i,o}(\lambda)$): Spectral internal absorptance of a layer of the material such that the path of the radiation is of unit length, and under conditions in which the boundary of the material has no influence (unit: 1).

In the previous paragraph, reflectivity refers to both interface and bulk properties of a semi-infinite material. Ref. [1] does not currently offer a definition for spectral reflectivity, but one can be formed by analogy with the definitions of spectral transmissivity and absorptivity by adding the subscript λ to ρ_∞ . The definitions of spectral transmissivity in the previous paragraph refer to the internal (i) properties of a material, away from any interfaces with other materials.

The previous definition of reflectivity includes the effects of volume scattering from within the medium of the material. In some cases, it is desirable to restrict the definition of reflectivity to include only interface effects. Some treatments use a bar over the reflectivity symbol ρ , as in $\bar{\rho}$, to indicate the interface contribution to the reflectivity of a substance. This notation will be continued in this text and will be extended to include interface transmissivity $\bar{\tau}$ as well.

Spectral transmissivity and absorptivity, as defined above, conform to the previously stated principle that the property of interest is an inherent or intrinsic

property of a material substance, without reference to interface effects. The concept of absorptivity is discussed more extensively in Section 6.4 on bulk medium optical properties.

The following notation convention is followed in the remainder of this text: Lower case Greek symbols will be used for *-ivity* quantities, while upper-case Roman symbols will be used for *-ance* quantities:

$\bar{\rho}$ = reflectivity of an interface;

ρ = reflectivity of a pure substance, including both bulk and interface processes;

R = reflectance of an object;

$\bar{\tau}$ = transmissivity of an interface;

τ = (internal) linear transmissivity of (a unit length of) a transparent or partially transparent substance, away from interfaces;

T = transmittance of an object;

α = (internal) linear absorptivity of (a unit length of) a transparent or partially transparent substance, away from interfaces;

A = absorptance of an object.

6.3 SURFACE AND INTERFACE OPTICAL PROPERTIES

6.3.1 Conductor Optical Properties

A perfect conductor, characterized by infinitely great conductivity, has an infinite refractive index and penetration of electromagnetic radiation to any depth is prohibited, producing perfect reflectivity [3]. Real conductors such as aluminum and silver do not have perfect conductivities nor do they have perfect reflectivities. However, their reflectivities are quite high over rather broad spectral ranges. They are therefore very useful in radiometric and photometric applications. Unprotected mirrors made of these materials, however, tend to degrade in exposure to air over time and they are seldom used without protective overcoatings. The normal incidence spectral reflectances of optical quality glass mirrors coated with aluminum, with aluminum having a magnesium fluoride protective overcoat, with aluminum having a silicon monoxide overcoat, with silver having a protective dielectric coating, and with gold are shown in Figure 6.1. The reflectance of these surfaces, already quite high at visible and infrared wavelengths, increases with incidence angle, approaching unity at 90 deg.

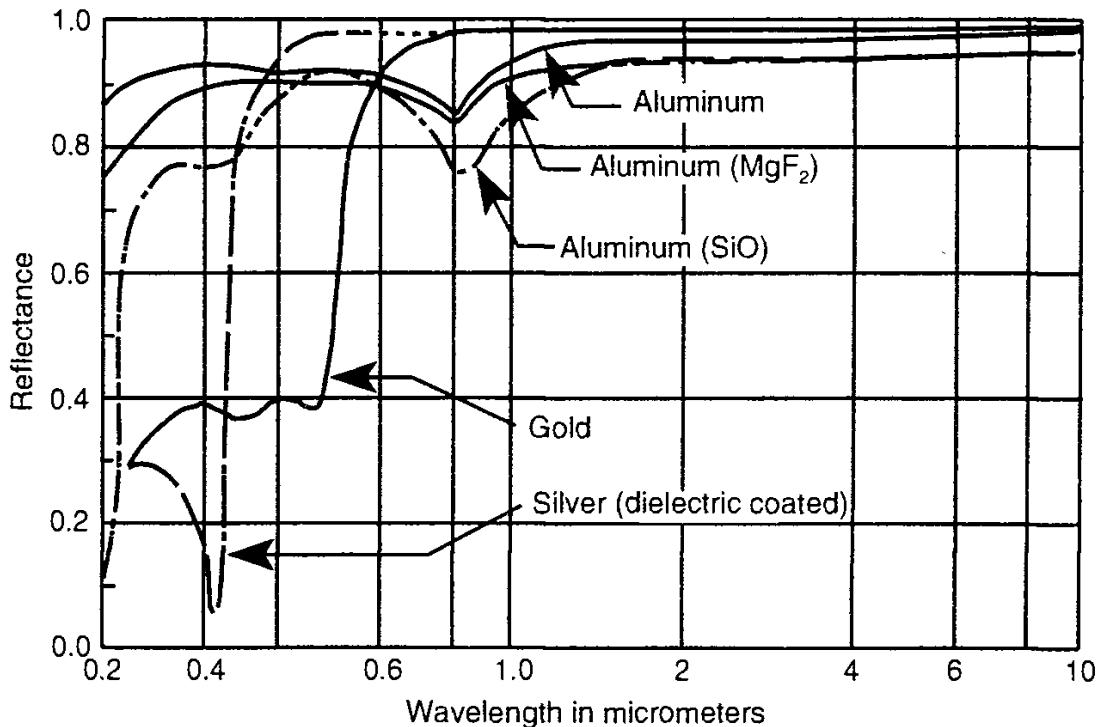


Figure 6.1 Spectral reflectances of commercially available metallic mirror coatings.

6.3.2 Nonconductor Optical Properties

Consider the extremely thin surface region of a perfectly smooth homogeneous and isotropic dielectric material, that is, its interface with another medium such as air, water, or a vacuum, an interface that is normally too thin to absorb significant quantities of the radiation incident on it. Absorption is not considered in this discussion since it is considered to be a bulk or volume characteristic of the material. Radiation incident upon an interface between two different materials is split into two parts. Some is reflected, and the rest is transmitted. The fraction of incident flux that is reflected is called the interface reflectivity $\bar{\rho}$, and the fraction transmitted is the interface transmissivity $\bar{\tau}$. The variations of $\bar{\tau}$ and $\bar{\rho}$ with angle of incidence are considered in Section 6.3.4. For the conservation of energy [4] at an interface

$$\bar{\tau} + \bar{\rho} = 1 \quad (6.1)$$

Of course, when the bulk medium optical properties are considered, the situation is more complicated than this, since the transmitted flux can be absorbed and “reflected” and/or scattered by the medium below the interface by direction and wavelength-dependent processes.

When both interface and interior optical processes are considered together, the spectral and directional variations in transmissivity and reflectivity become still

more important, and the absorptivity of the medium also comes into play. The wavelength dependence of the optical properties of materials is indicated with a functional λ notation thus: $\tau(\lambda)$, $\alpha(\lambda)$, and $\rho(\lambda)$. The direction of an element of solid angle $d\omega$ is indicated using the spherical angular coordinates (θ, ϕ) , illustrated in Figure 1.8 of Chapter 1 and repeated here as Figure 6.2. Using these coordinates, the directional dependence of optical properties is indicated with the functional notation: $\tau(\theta, \phi)$ and $\rho(\theta, \phi)$, and the combined spectral and directional properties thus: $\tau(\lambda, \theta, \phi)$ and $\rho(\lambda, \theta, \phi)$. More is said about these optical properties subsequently. It will be seen that these properties depend in general upon both *incident* and *emergent* beam directions.

6.3.3 Surface Emission Properties

The emissive properties of non-blackbody surfaces are characterized by a quantity called by various authors either the *emittance* or the *emissivity* ϵ of the surface. Emittance is the ratio of the actual emission of thermal radiant flux from a surface to the flux that would be emitted by a perfect blackbody emitter (discussed in Chapter 3) at the same temperature. According to the terminology guidelines given in Section 6.1, the term *emissivity* should be reserved for the surface of an infinitely

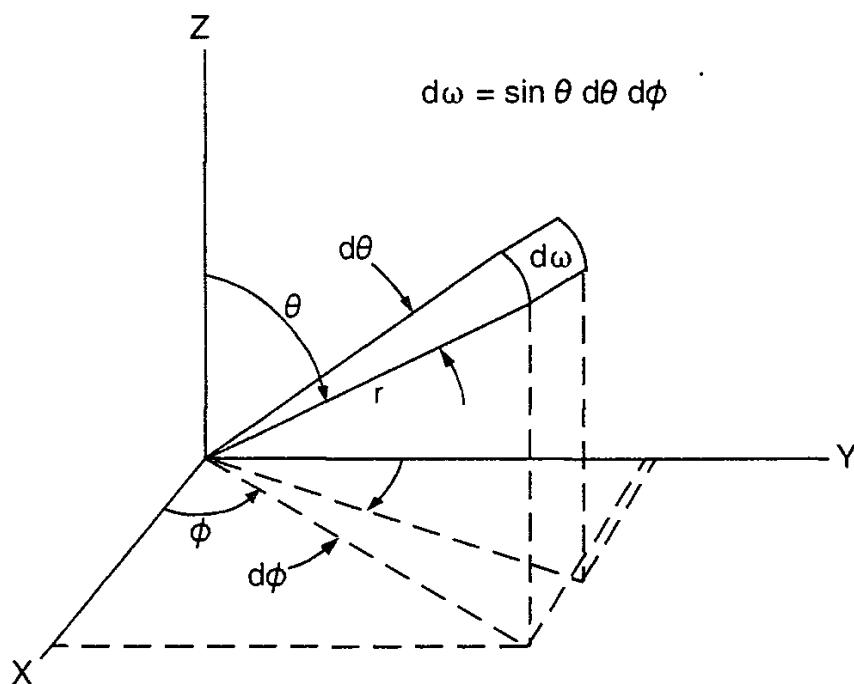


Figure 6.2 Coordinate system for specifying the direction and size of infinitesimal element $d\omega$ of solid angle.

thick slab of pure material with a polished surface, while emittance would apply to a finite thickness of an actual object. As pointed out by Siegel and Howell [2], for substances opaque at the wavelengths of emission, the intrinsic and extrinsic versions of ϵ are the same, leading to two names for the same quantity. The emission of radiant flux from a substance opaque to the emitted spectrum is a surface phenomenon, not a volume one.

The 1987 CIE *International Lighting Vocabulary* [1] uses only the term “emissivity” for ϵ . However, the 1993 *Handbook of Fundamentals* of the American Society of Heating, Refrigerating, and Air Conditioning Engineers (ASHRAE) [5] uses “emittance” for ϵ and points out that “emittance, reflectance, absorptance, and transmittance refer to actual pieces of material. Emissivity, reflectivity, absorptivity, and transmissivity refer to the properties of materials that are optically smooth and thick enough to be opaque.” A new American Society for the Testing of Materials (ASTM) standard for the measurement and calculation of emittance of architectural flat glass products [6] follows this terminology, but uses the roman lower case letter e for emittance and the Greek letter epsilon, ϵ for emissivity. Although this terminology is used through the remainder of this book, the Greek letter ϵ will be used for *both* emittance and emissivity in order to avoid any conflict with the symbol E reserved for irradiance or the symbol e for the base of the natural logarithms.

Surfaces that are not perfect blackbodies are called greybodies if their emittance is constant, independent of wavelength. They are called nonblackbodies if their emittance varies with wavelength. Greybodies do not necessarily appear grey in color, but they do not absorb all radiation incident upon them and therefore do not have to appear totally black. The exitance of a greybody is given by the following modified form of the Stefan-Boltzmann law:

$$M = \epsilon\sigma T^4 \quad (6.2)$$

where ϵ is the emittance of the surface. The emittance is 1.0 for a blackbody. Surfaces of real materials do not emit as much as a blackbody at the same temperature; they have emittances ϵ that are less than 1.0. Typical emittances for a variety of materials are listed in Table 6.1. Only the far infrared emittance values are given. Emittances at other wavelengths may be different.

As was the case for reflectance and transmittance, emittance is in general a directional quantity and can be specified as $\epsilon(\theta, \phi)$. The directional emittance at normal incidence ($\theta = 0$) is called the *normal emittance*. The average of the directional emittance over the whole hemispherical solid angle is called the *hemispherical emittance*. The emittances shown in Table 6.1 are for hemispherical emittance into a vacuum. Emission into an imperfect vacuum or a solid was discussed in Section 3.6. (It can be noted that many solids have strong absorptance spectra over portions of the spectrum involved in long-wavelength blackbody or greybody

Table 6.1
Hemispherical Emittance Values for Typical Materials

Material	Emittance From 4 to 16 μm
White paint	0.90
Black asphalt and roofing tar	0.93
Light concrete	0.88
Pine wood	0.60
Stainless steel	0.18 to 0.28
Galvanized sheet metal	0.13 to 0.28
Aluminum sheet metal	0.09
Polished aluminum	0.05–0.08

thermal emission at modest temperatures and many materials are opaque at these wavelengths. Emitted radiation at long wavelengths may not penetrate very far into such materials.)

The normal emittances of several materials are listed in Table 6.2. The emittances of many additional materials are given in the *Handbook of Chemistry and Physics* [7], pages E-390 through E-392.

As shown in Table 6.2, the emittance depends upon the temperature of the emitting surface. The values listed in Table 6.2 are each for a specific temperature or range of temperatures. Most of these data were taken from the *Handbook of Chemistry and Physics* cited previously. This handbook contains optical property data for a variety of other materials as well.

The emitted radiances of real surfaces in general vary with wavelength as well as direction. To account for these effects, the spectral directional emittance $\epsilon(\lambda, \theta, \phi)$ of a surface is defined as the ratio of flux emitted at wavelength λ in a direction specified by the angles θ and ϕ from a surface at some temperature to that emitted by a blackbody at the same temperature and wavelength and in the same direction. In terms of the emitted spectral radiance $L_{\lambda\text{nb}}$ of the nonblackbody and the emitted spectral radiance of a blackbody $L_{\lambda\text{bb}}$, the spectral directional emittance can be defined by the relation

$$\epsilon(\lambda, \theta, \phi) \equiv \frac{L_{\lambda\text{nb}}(\theta, \phi)}{L_{\lambda\text{bb}}(\theta, \phi)} \quad (6.3)$$

The symbol \equiv is used here to mean *equivalency*, to emphasize that the equations involved are defining equations. *Hemispherical emittance* is defined by the relation

$$\epsilon(\lambda) = \frac{M_{\lambda\text{nb}}}{M_{\lambda\text{bb}}} \quad (6.4)$$

Table 6.2
Normal Emittance Values for Typical Materials

Material	Emittance From 4 to 16 μm
Highly polished aluminum	0.04–0.06 at 50–500C
Oxidized aluminum	0.11 at 200C
Polished copper	0.02 at 50–100C
Planed oak wood	0.895 at 27C
Polished steel	0.52–0.56 at 750–1,050C
Brick	0.85–0.93
Carbon filament	0.53 at 1,000–1400C
Smooth glass	0.937
Uncoated window glass	0.84
Polished silver	0.02–0.03 at 200–600C
Tungsten	0.032 at 300K 0.101 at 900K 0.207 at 1,600K 0.334 at 3,000K
Thick lampblack carbon	0.96 at 20–400C

in terms of the spectral exitances of a nonblackbody ($M_{\lambda\text{nb}}$) and a blackbody ($M_{\lambda\text{bb}}$). It is pointed out in Section 3.1 that $L_{\lambda\text{bb}}(\theta, \phi)$ is independent of direction, so that $M_{\lambda\text{bb}} = \pi L_{\lambda\text{bb}}(\theta, \phi)$ and

$$\epsilon(\lambda, \theta, \phi) = \frac{\pi}{M_{\lambda\text{bb}}} L_{\lambda\text{nb}}(\theta, \phi) \quad (6.5)$$

When the hemispherical emittance is integrated over all wavelengths, it is called the *total emittance*. Many surfaces are spectrally selective in emittance. They have different spectral emittance values in different parts of the spectrum. To illustrate this point, the spectral hemispherical emittance of tungsten at 2,800K is shown in Table 6.3 for several wavelengths ranging from 0.25 to 2.6 mm. This data was taken from the CRC *Handbook of Chemistry and Physics* [7].

Table 6.3
Spectral Emittance of Tungsten at 2,800K

Wavelength	0.25 μm	0.50 μm	1.0 μm	1.8 μm	2.6 μm
Emittance	0.411	0.448	0.367	0.274	0.224

Temperatures at the surface of the earth and inside human-built structures seldom exceed the range from, say, -27 to 123°C , or 250 to 400K . From Figure 3.1 in Chapter 3, it is seen that the predominant emission from a blackbody below 400K occurs at wavelengths above about $4 \mu\text{m}$. The hemispherical emittance for most ordinary materials is therefore usually quoted for wavelengths above $4 \mu\text{m}$ or so.

The emittance shown in (6.2) is the total hemispherical emittance. If the spectral emittance is strongly varying over a portion of the emitted spectrum of interest, then (6.2) is not appropriate. Following is the modified form of Planck's radiation law (3.3), for nonblackbodies:

$$L_{\text{nb}\lambda}(\theta, \phi) = \frac{2hc^2\epsilon(\lambda, \theta, \phi)}{\lambda^5(e^{hc/\lambda kT} - 1)} \quad (6.6)$$

The nonblackbody version of (3.5) for the spectral exitance $M_{\text{nb}\lambda}$ is

$$M_{\text{nb}\lambda} = \frac{C_1}{\lambda^5(e^{C_2/\lambda T} - 1)} \int_0^{2\pi} \int_0^{\pi/2} \epsilon(\lambda, \theta, \phi) \cos \theta \sin \theta d\theta d\phi \quad (6.7)$$

where

$$C_1 = 2\pi hc^2 = 3.741832 \times 10^8 \text{ W} \cdot \mu\text{m}^4 \cdot \text{m}^{-2}$$

$$C_2 = hc/k = 14,387.86 \mu\text{m} \cdot \text{K}$$

$\epsilon(\lambda, \theta, \phi)$ is the spectral directional emittance at a point on the surface of the radiator.

6.3.4 Angular Dependence of Dielectric Optical Properties

Formulas for the reflectivities of the parallel and perpendicular polarization components at an infinitesimally thin interface between two *dielectric* media of differing refractive indices, n_1 and n_2 , such as air and glass, were worked out by the French physicist Augustin Fresnel (pronounced "Frehnnel") in the early 1800s. Fresnel's formulas for the specular reflectivities of the polarization components parallel \parallel and perpendicular \perp the plane of incidence (the plane containing a ray incident on the surface and the normal to the surface at the point of incidence) are as follows:

$$\bar{\rho}_{\parallel} = \frac{\tan^2(\theta_1 - \theta_2)}{\tan^2(\theta_1 + \theta_2)} \quad (6.8)$$

$$\bar{\rho}_{\perp} = \frac{\sin^2(\theta_1 - \theta_2)}{\sin^2(\theta_1 + \theta_2)} \quad (6.9)$$

where θ_1 is the angle of incidence and θ_2 is the angle of refraction in the medium, the two angles being related by Snell's law, (5.11) in Chapter 5. (Polarization of radiation is discussed in Section 1.8.) The term *specular* (discussed at length in Section 6.7.2) refers to reflection (and transmission) according to the laws of geometrical optics; that is, Snell's law and the equality of angles of incidence and reflection from a surface. At normal incidence there is no difference between the parallel and perpendicular polarization directions and

$$\bar{\rho} = \frac{(n_1 - n_2)^2}{(n_1 + n_2)^2} \quad (6.10)$$

where n_1 and n_2 are the refractive indices on either side of the interface.

As mentioned previously, the overbar on the ρ is used here and subsequently to indicate reflectivity at an *interface*, not including any effects from within the medium. Equation (6.10) can be derived by using the approximation that as x goes to zero, $\sin x \approx x$. Thus, (6.9) can be replaced by $\bar{\rho} = (\theta_1 - \theta_2)^2/(\theta_1 + \theta_2)^2$ and Snell's law can be solved for $\sin \theta_2$ in terms of n_1 , n_2 , and $\sin \theta_1$. The small-angle approximation is again used and, following a rearrangement of terms, (6.10) is obtained.

For the total reflected flux, including both polarization components, assuming unpolarized (or circularly polarized) incident radiation, one takes the average of the parallel and perpendicular components of the reflectance, given above

$$\bar{\rho} = \frac{\bar{\rho}_{\parallel} + \bar{\rho}_{\perp}}{2} \quad (6.11)$$

Through the use of Snell's law and several trigonometric identities, one can write formulas for the parallel and perpendicular components of the reflectance in the following form:

$$\bar{\rho}_{\parallel} = \left[\frac{n_1 \cos \theta_2 - n_2 \cos \theta_1}{n_1 \cos \theta_2 + n_2 \cos \theta_1} \right]^2 \quad (6.12)$$

$$\bar{\rho}_{\perp} = \left[\frac{n_1 \cos \theta_1 - n_2 \cos \theta_2}{n_1 \cos \theta_1 + n_2 \cos \theta_2} \right]^2 \quad (6.13)$$

To get the corresponding transmittances, (5.14) is solved for the radiance transmissivity $\bar{\tau} = L_2/L_1$ to obtain

$$\bar{\tau} = \left(\frac{n_1}{n_2}\right)^2 (1 - \bar{\rho}) \quad (6.14)$$

This result can be applied to the separate polarization components to obtain

$$\bar{\tau}_{\parallel} = (1 - \bar{\rho}_{\parallel}) \left(\frac{n_2}{n_1}\right)^2 \quad (6.15)$$

$$\bar{\tau}_{\perp} = (1 - \bar{\rho}_{\perp}) \left(\frac{n_2}{n_1}\right)^2 \quad (6.16)$$

These equations can be recast in terms of $n = n_2/n_1$ and the angle of incidence θ_i only, using the identity $\cos^2 x + \sin^2 x = 1$. The results are:

$$\bar{\rho}_{\perp} = \left[\frac{n \cos \theta_i - \sqrt{1 - n^2 \sin^2 \theta_i}}{\cos \theta_i + n \sqrt{1 - n^2 \sin^2 \theta_i}} \right]^2 \quad (6.17)$$

$$\bar{\tau}_{\parallel} = (1 - \rho_{\parallel}) n^2 \quad (6.18)$$

$$\bar{\rho}_{\parallel} = \left[\frac{n \sqrt{1 - n^2 \sin^2 \theta_i} - \cos \theta_i}{n \sqrt{1 - n^2 \sin^2 \theta_i} + \cos \theta_i} \right]^2 \quad (6.19)$$

$$\bar{\tau}_{\perp} = (1 - \rho_{\perp}) n^2 \quad (6.20)$$

As before, for unpolarized and circularly polarized light,

$$\bar{\rho} = \frac{\bar{\rho}_{\parallel} + \bar{\rho}_{\perp}}{2} \quad (6.21)$$

$$\bar{\tau} = \frac{\bar{\tau}_{\parallel} + \bar{\tau}_{\perp}}{2} \quad (6.22)$$

Figure 6.3 shows the interface reflectivities and transmissivities given in (6.17) through (6.22), plotted versus angle of incidence, for air as the first medium and a material with refractive index 1.80 as the second one. A high refractive index was chosen for these plots to make the shapes of the curves clearer to see. Ordinary plate glass has a refractive index closer to 1.5 in the visible portion of the spectrum.

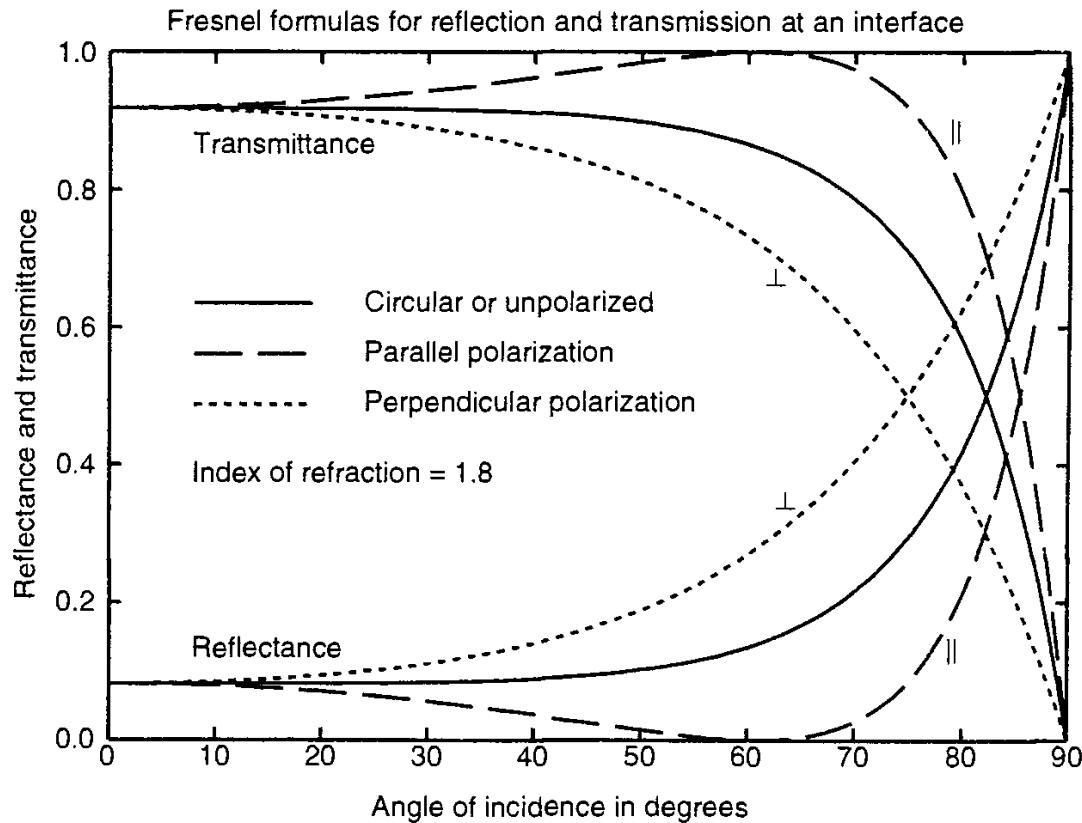


Figure 6.3 Interface reflectivities and transmissivities for parallel (||) and perpendicularly (⊥) polarized light and unpolarized light incident from air to a medium with refractive index 1.8.

Since we are assuming a homogeneous medium without scattering effects, the transmitted ray emerges at the angle θ_2 specified by Snell's law and the reflected ray emerges at the same angle θ_1 as the incident one, but on the other side of the normal to the surface at the point of entry.

Figure 6.4 shows the reflectivity and transmissivity for rays going the *other* direction, from glass to air, calculated with (6.17) through (6.22) by reversing the refractive indices (replacing n with $1/n$). One can see in these plots the effect of what is called *total internal reflection*. When rays go into a medium with lower refractive index, there is a special angle, called the *critical angle*, given by

$$\theta_c = \sin^{-1} \frac{n_2}{n_1} \quad n_2 < n_1 \quad (6.23)$$

Use of this angle in Snell's law produces an angle of refraction of 90 deg. The reflectivity is 1.0 for all angles of incidence greater than this angle and the phenomenon of total internal reflection (TIR) is observed—the reflectivity at each reflection is total, 100%. TIR is the principle of operation of modern light pipes called fiber optics that are used to transmit coded optical radiation very reliably

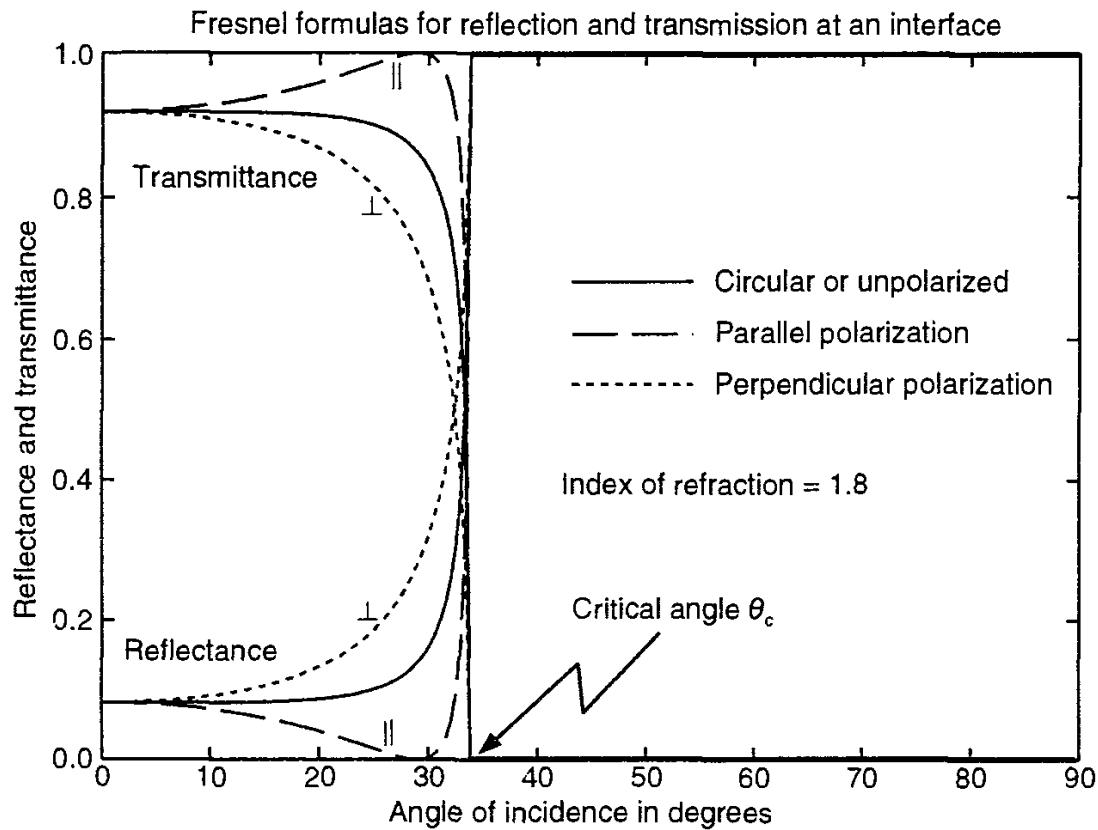


Figure 6.4 Reflectivities and transmissivities for incidence from glass with a refractive index of 1.8 to air with a refractive index of 1.0.

over great distances. Light pipes are used in a variety of other applications as well. They are currently being used more and more for the piping of illumination inside buildings and for remote measurements of parameters ranging from water pressure and the refractive index of liquids to the concentration of ground water pollutants [8–10].

It is also noted that the transmissivity of the parallel component in Figure 6.4 becomes 1.0 and the reflectivity becomes zero at a certain angle, called the *polarizing angle* or *Brewster's angle* where the angle between the reflected and refracted rays is 90 deg, that is when $\tan^2(\theta_1 + \theta_2) = \infty$. This effect can be used to convert unpolarized or circularly polarized light to partially linearly polarized light by illuminating a sheet of transparent material at this angle. Due to surface films, surface strains, and so forth, in real materials the transmitted light may not be completely linearly polarized. It may have a small unpolarized component or be elliptically polarized with a very high eccentricity [11].

Equations (6.8) through (6.22) are useful for a variety of applications and are used in Section 6.4 to describe the optical properties of a parallel plate of transparent material such as glass.

The refractive indices of several different materials are listed in Table 6.4. The refractive index varies with wavelength, a phenomenon known as *dispersion*. Dispersion provides the principle of operation of the prism spectrometer. A ray of light incident at an angle on a wedge-shaped prism of glass is bent by refraction according to Snell's law. Since the refractive index of the glass varies with wavelength, different wavelengths are bent by different amounts and the result is a spreading with angle of the incident spectrum. For incident white light, the effect over the visible portion of the spectrum is to produce a dispersed spread of colors emanating from the prism. This effect is discussed at greater length in Section 8.12.2.1.

6.3.5 Rough Surfaces

There is considerable interest in the optical properties of rough surfaces. Most materials have surfaces with significant non-specular components and their scattering properties are of some importance in several fields. An excellent introduction to the subjects of surface roughness and light scattering from surfaces is provided by Bennett and Mattsson [12].

Table 6.4
Refractive Indices of Common Optical Materials

Material	Refractive Index at Wavelength
Pure water at 20°C	1.3403 at 434 nm
Pure water at 20°C	1.3330 at 589 nm
Borosilicate Crown glass	1.517 at 589.3 nm
	1.52262 at 486.1 nm
	1.51461 at 656.3 nm
	1.4893 at 2.325 μ m
Light flint glass	1.575 at 589.3 nm
Extra dense flint glass	1.689 at 589.3 nm
Lanthanum flint glass	1.720 at 589.3 nm
Polymethyl methacrylate (Lucite)	1.491 at 589.3 nm
Polycyclohexyl methacrylate	1.506 at 589.3 nm
Fused quartz (fused silica, SiO_2)	1.4585 at 589.3 nm
	1.4656 at 450 nm
	1.4564 at 656.3 nm
Irtran 4	2.786 at 489 nm
	2.491 at 1 mm
	2.408 at 10 mm

6.4 BULK MEDIUM OPTICAL PROPERTIES

The previous sections, and parts of Chapter 5, deal with the main optical properties characterizing the fate of a ray of radiance or luminance when it is incident upon an *interface* between two different media. Now the optical properties governing radiation propagation *within* an absorbing and scattering medium (without sources) are discussed. In the next section, we combine both medium and interface properties for a treatment of the optical properties of a finite object with plane parallel boundaries (interfaces with other media).

Consider two planes inside a homogeneous medium, these planes being separated by the infinitesimal distance dx , as shown in Figure 6.5. The spectral radiance along a ray propagating perpendicular to these two planes is $L_\lambda(x)$ in the first plane and $L_\lambda(x + dx)$ in the second plane. Since there are no sources along this ray and multiple scattering effects are not considered, $L_\lambda(x + dx)$ will be less than $L_\lambda(x)$. One writes

$$dL_\lambda = L_\lambda(x) - L_\lambda(x + dx) \quad (6.24)$$

and defines the spectral volume extinction (or attenuation) coefficient $c(\lambda)$ as follows

$$c(\lambda) \equiv -\frac{1}{L_\lambda} \frac{dL_\lambda}{dx} \quad (6.25)$$

with $c(\lambda)$ having the dimensions of reciprocal length, such as m^{-1} . As before, the symbol \equiv is used to mean *equivalency*, emphasizing that the equations involved are defining equations. Solving (6.25) for dL_λ/L_λ , replacing x with x' , dx with dx' , and integrating from $x' = 0$ to $x' = x$ yields

$$\int_0^x \frac{dL_\lambda(x')}{L_\lambda} = - \int_0^x c(\lambda) dx' \quad (6.26)$$

$$L_\lambda(x) = L_\lambda(0)e^{-c(\lambda)x} \quad (6.27)$$

Equation (6.27) is sometimes referred to as Beer's law and it is assumed in this form of the equation that $c(\lambda)$ is spatially constant from 0 to x . (If $c(\lambda)$ is not constant from $x' = 0$ to $x' = x$, then the exponent in (6.27) must be replaced by the negative integral on the right side of (6.26).)

It can be noted that in any experiment to determine $c(\lambda)$ over any sizable path length, great care must be taken to avoid the detection of multiply scattered

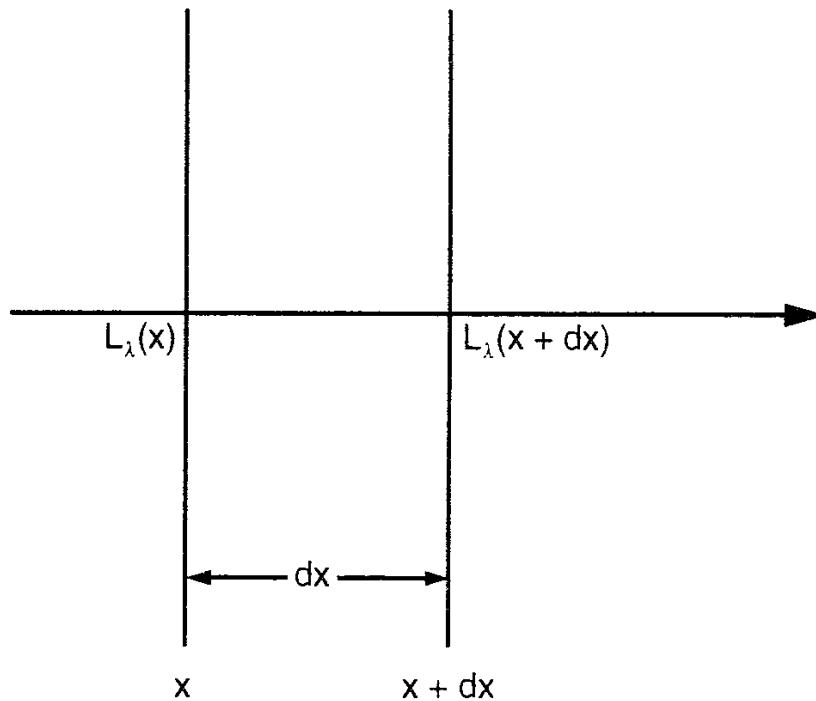


Figure 6.5 The spectral radiance L_λ at two parallel planes separated by distance dx in an absorbing and scattering medium.

radiation, radiation scattered out of the beam that is subsequently scattered back into it.

If one has a collimated beam propagating along the direction indicated by the ray in Figure 6.5, meaning that the flux is confined to a very small solid angle around all the rays in the beam parallel to that ray, then $L_\lambda(x)$ in (6.27) can be integrated over the small solid angle to obtain the spectral irradiance $E_\lambda(x)$ on a plane perpendicular to the ray, and the more conventional version of (6.27)

$$E_\lambda(x) = E_\lambda(0)e^{-c(\lambda)x} \quad (6.28)$$

Opaque media are those for which $c(\lambda)$ is so large (or the thickness t of the medium is so great) that $E_\lambda(t)/E_\lambda(0) = 0$ in (6.28). Nearly any opaque medium can be made partially transparent by slicing it thin enough. (Although the term “opaque” has a visible light connotation, the term is frequently used to describe objects with extremely high extinction coefficients in other spectral regions.)

The coefficient $c(\lambda)$ in (6.27) and (6.28) (or its integral over wavelength) is called either the *extinction coefficient* or the *attenuation coefficient* by most authors. It is sometimes given the name *turbidity*. However, this term has been applied to a large variety of different optical properties in the fields of hydrology, oceanography, water pollution [13–19], and atmospheric physics. Because of the confusion

of terminology, it is recommended that the term turbidity be limited to visual observations of the appearance of scattering media. When one is really interested in the particulate concentration of a medium, one should specify this concentration rather than some optical property that is related to the particle concentration. If it is an optical property of the medium that is of interest, then the specific optical property itself should be used. Additional optical properties of scattering media are defined next.

The element dL_λ of spectral radiance in (6.24) that is lost due to absorption and scattering is separated into an amount due to absorption only $dL_{a,\lambda}$ and an amount due to scattering only $dL_{b,\lambda}$. By analogy with (6.25), we have definitions for the *spectral absorption coefficient* $a(\lambda)$:

$$a(\lambda) \equiv -\frac{1}{L_\lambda} \frac{dL_{a,\lambda}}{dx} \quad (6.29)$$

and the *spectral scattering coefficient* $b(\lambda)$:

$$b(\lambda) \equiv -\frac{1}{L_\lambda} \frac{dL_{b,\lambda}}{dx} \quad (6.30)$$

Because $dL_\lambda = dL_{a,\lambda} + dL_{b,\lambda}$, clearly

$$c(\lambda) = a(\lambda) + b(\lambda) \quad (6.31)$$

As with $c(\lambda)$, the units of $a(\lambda)$ and $b(\lambda)$ are reciprocal length. The combined extinction coefficient $c(\lambda)$ is an intrinsic property of the medium and should be independent of the measuring instrument or method. In reality, it is difficult to prevent some singly and multiply scattered radiation from reaching the detector in actual measurements. Techniques exist, however, for minimizing the amount of scattered radiation detected, and to therefore come close to a measurement of the true extinction coefficient. The above symbols and terminology are common in hydrologic optics [20–21]. Different symbols are used in atmospheric optics [22–23]. The CIE uses the symbol $\mu(\lambda)$ for $c(\lambda)$, the symbol $s(\lambda)$ for $b(\lambda)$, and the symbol $a(\lambda)$ for $a(\lambda)$ [1] and omits the negative signs in (6.25), (6.29) and (6.30), resulting from a reversal of the signs of the terms on the right-hand side of (6.24).

The total scattering coefficient $b(\lambda)$ represents radiation lost from a ray by scattering in all directions. The angular distribution of this singly scattered radiation is described by a function called the spectral volume scattering function (VSF), $\beta(\lambda, \theta)$.

Consider the small element of volume dv shown in Figure 6.6. Let E_λ be the irradiance of a collimated beam of radiation incident on this volume in the direction of the x axis and $d^2\Phi_\lambda$ be the element of spectral flux scattered out of dv into the element of solid angle $d\Omega$ at the scattering angle θ . The spectral volume scattering function $\beta(\lambda, \theta)$ is defined by

$$\beta(\lambda, \theta) = \frac{d^2\Phi_\lambda(\theta)}{E_\lambda d\Omega dv} \quad (6.32)$$

and is called the *scattering indicatrix* by some authors. The quotient $d^2\Phi_\lambda(\theta)/d\Omega$ in this equation is by definition the element $dI_\lambda(\theta)$ of spectral intensity scattered from the infinitesimal volume dv in the direction specified by angle θ and may be replaced by it in (6.32) to yield

$$\beta(\lambda, \theta) = \frac{dI_\lambda(\theta)}{E_\lambda dv} \quad (6.33)$$

Still another expression can be obtained for the definition of the VSF by writing dv as the product of an infinitesimal distance dx along the axis of the collimated incident beam and an infinitesimal area da perpendicular to this direction as follows:

$$\beta(\lambda, \theta) = \frac{d^3\Phi_\lambda(\theta)}{E_\lambda d\Omega dadx} \quad (6.34)$$

or, equivalently,

$$\beta(\lambda, \theta) = \frac{dL_\lambda(\theta)\cos(\theta)}{E_\lambda dx} \quad (6.35)$$

This last form shows the VSF to be the radiance per unit length and per unit incident irradiance scattered in the direction θ , weighted by $\cos \theta$. Equation (6.35) is used in Section 6.5.2.

The origin of the $\cos \theta$ term in (6.35) is illustrated in Figure 6.7. The radiance is defined in terms of the projected area $da \cos \theta$ so that (6.34) is multiplied by unity in the form $\cos \theta / \cos \theta$ placing a $\cos \theta$ in the denominator for the definition of radiance. Once $d^3\Phi_\lambda(\theta)/(d\Omega da \cos \theta)$ is replaced by $dL_\lambda(\theta)$ in (6.34), $\cos \theta$ is left in the numerator, as shown in (6.35).

Solving (6.32) for $d^2\Phi_\lambda(\theta)$ and integrating over a finite volume and solid angle yields the flux scattered from that volume into the selected solid angle. The VSF describes how the radiation scattered by an element of volume is distributed in

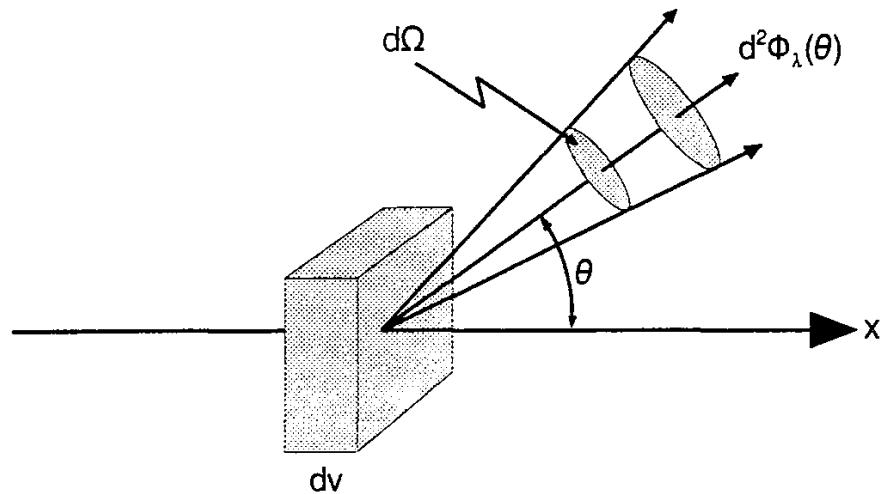


Figure 6.6 Geometry for the definition of the volume scattering function.

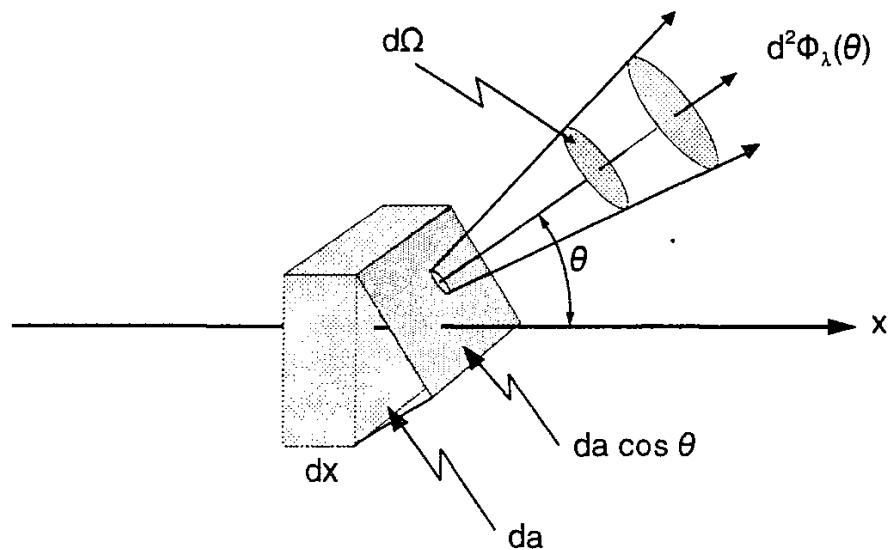


Figure 6.7 Geometry to illustrate the origin of the $\cos \theta$ term in (6.31).

direction. As defined above, the VSF presumes the scattering to be axially symmetric, depending only on the scattering angle θ . If the scattering particles are oriented in such a way as to produce an azimuthal dependence, the defining equation can be expanded to include this dependency.

As might be expected, the total scattering coefficient is the integral of the VSF over a full 4π sr spherical solid angle

$$\begin{aligned} b(\lambda) &= \int_{4\pi} \beta(\lambda, \theta) d\omega \\ b(\lambda) &= 2\pi \int_0^\pi \sin \theta \beta(\lambda, \theta) d\theta \end{aligned} \quad (6.36)$$

There are several theories of scattering that can be used to determine $\beta(\lambda, \theta)$. The major theories are distinguished by the size of the scattering particle in comparison with the wavelength of the radiation being scattered. Rayleigh theory is applicable to very small particles, at the atomic or molecular level, and yields a λ^{-4} wavelength dependence and a $(1 + \cos^2 \theta)$ angular dependence. This theory has been used for scattering by molecules in the earth's atmosphere and in pure water. Rayleigh's formula, as developed by Kondratyev [24], is

$$\beta(\lambda, \theta) = \frac{\pi^2(n^2 - 1)^2}{2N\lambda^4} (1 + \cos^2 \theta) \quad (6.37)$$

where $\beta(\lambda, \theta)$ is the VSF, N is the number of particles per unit volume, and n is the refractive index of the medium that is composed of the particles (air for the atmosphere or water). It is this $1/\lambda^4$ dependence in molecular scattering that is responsible for the blueness of the sky. Light scattered by air molecules is strongest in the blue portion of the spectrum. This effect is also responsible for the orange color of the sunset. Light reaching the eye from the sun near sunrise or sunset has much of the blue light stripped from it, leaving light with a decidedly reddish-orange color. The blue color of water is due not only to Rayleigh scattering but also to spectral molecular absorption effects, which are strongly spectrally selective, with wavelengths in the blue portion being absorbed the least. Molecular scattering redirects the transmitted and spectrally modified incident sunlight back to the eye [25, 26].

For spherically shaped particles, much larger than the wavelength of incident radiation, the Lorentz-Mie scattering theory [27] applies. This theory yields a much stronger angular dependence, with a large forward-scattering component, but a weaker wavelength dependence. This slighter wavelength dependence is responsible for the lack of color in clouds containing water droplets that are large compared with the wavelengths of light. The apparent red or orange color of some clouds at sunrise and sunset is due to the scattering of colored sunlight from the clouds.

For scattering media where multiple scattering effects are significant, additional complex theory is required if one desires accurately to describe the large-scale bulk optical properties of such media. One such approach is the Monte Carlo

method, described in a different context in Section 4.2.8. Others are described in Section 5.8 and at the end of Section 6.5.2.

6.5 PROPERTIES OF PLANE PARALLEL PLATES

The interfaces between two media, on either side of a finite portion of one of them, define an object that can reflect, transmit, and absorb radiation incident on it. When a ray of radiation passes through a finite distance of a material, flux can be removed from the ray by absorption and scattering processes. This section deals with the optical properties of a plane parallel “plate” of transparent or semitransparent material. Following the terminology conventions of Section 6.2, the following definitions are introduced:

Spectral transmittance, $T(\lambda)$, is the fraction of flux incident upon an object that is transmitted by the object, for defined directions of incidence and emergence, at the wavelength λ , and for defined incident and emergent polarization states.

Spectral reflectance, $R(\lambda)$, is the fraction of flux incident upon an object that is reflected by the object, for defined directions of incidence and emergence, at the wavelength λ , and for defined incident and emergent polarization states.

Spectral absorptance, $A(\lambda)$, is the fraction of flux incident upon an object that is absorbed by the object, for a defined direction of incidence, at the wavelength λ , and for defined incident polarization states.

Discussion of the directional dependencies of the above defined quantities is postponed until additional concepts are introduced. The interface optical properties $\bar{\tau}$, $\bar{\rho}$, and ϵ discussed previously are combined below with the intrinsic properties of spectral absorptance and scattering to deduce the extrinsic optical properties T , R , and A defined above, for plates of materials, specific whole objects with surfaces that are flat and parallel to each other and have the same medium on both sides (usually air or vacuum). Before considering more general cases, the simple case of nonscattering media is examined first.

6.5.1 Nonscattering Media

Consider a collimated beam of radiation (a bundle of approximately parallel rays propagating in the same direction with the associated flux contained in a small but measurable solid angle) in air incident upon a plate of homogeneous and isotropic material. The spectral transmittance $T(\lambda)$ of the plate is defined to be the fraction of collimated incident spectral irradiance at the wavelength λ that emerges from the other side of the plate, i.e., the ratio of transmitted-to-incident spectral *irradiances*. It can also be defined as the ratio of the transmitted-to-incident spectral

radiances along one of the rays incident on the plate. As is shown in subsequent sections, additional definitions are useful.

Similarly, the spectral reflectance $R(\lambda)$ is the fraction of incident spectral radiance or irradiance that is reflected by the plate. The spectral absorptance $A(\lambda)$ is the fraction that is absorbed inside the plate. Except for the case of fluorescence or other effects that shift radiant energy from one wavelength to another, all the incident flux must be accounted for. Thus the sum of the transmitted, reflected, and absorbed fractions of the incident radiation must be unity:

$$T(\lambda) + R(\lambda) + A(\lambda) = 1 \quad (6.38)$$

In general each of these properties is dependent upon the direction of incidence and the polarization state of the radiation incident on the parallel plate of material. The angular dependence of these and other optical properties is discussed in Sections 6.6 and 6.7. Polarization was discussed in Section 6.3. For the remainder of this section, it is assumed that the incident beam is unpolarized and that the total flux of the emergent beam is collected, regardless of polarization state.

Before moving on to consider scattering media, an important relationship between emittance and absorptance is established and the transmittance and reflectance of a plane parallel plate are calculated.

There is a connection between emittance, absorptance, and reflectance. For opaque materials, a good absorber is generally a good emitter and a good reflector is a poor emitter, on a wavelength-by-wavelength basis. It is important to emphasize that this is only true in general at a specific wavelength, as indicated in the following relationship, called Kirchhoff's law [28,29]:

$$A(\lambda, \theta, \phi) = \epsilon(\lambda, \theta, \phi) \quad (6.39)$$

where θ and ϕ are angles defining the directional dependence of the spectral absorptance $A(\lambda)$ and spectral emittance $\epsilon(\lambda)$. For opaque bodies, $T(\lambda) = 0$ and

$$\begin{aligned} A(\lambda) + R(\lambda) &= 1 \\ \epsilon(\lambda) + R(\lambda) &= 1 \end{aligned} \quad (6.40)$$

showing explicitly the relationship between spectral emittance, absorptance, and reflectance.

Grum and Becherer [29] derive (6.39) as follows. Equation (6.3) is used for the spectral directional emittance $\epsilon(\lambda, \theta, \phi)$ and $L_{\lambda o}(\theta, \phi)/L_{\lambda i}(\theta, \phi)$ is used for the spectral directional absorptance $A(\lambda, \theta, \phi)$, of a nonblackbody in an isothermal cavity at thermal equilibrium. To maintain equilibrium, the emitted radiance $L_{\lambda nb}(\theta, \phi)$ must

equal the absorbed radiance $L_{\lambda a}(\theta, \phi)$ at each wavelength and in each direction. This means that

$$\epsilon(\lambda, \theta, \phi)L_{\lambda bb} = A(\lambda, \theta, \phi)L_{\lambda i} \quad (6.41)$$

However, a black isothermal enclosure at thermal equilibrium must be at the same temperature as the body enclosed. Thus, $L_{\lambda i}(\theta, \phi) = L_{\lambda bb}(\theta, \phi)$ and (6.39) results.

Section 6.9 discusses spectral broadband properties of materials, optical properties integrated over large wavelength ranges. Spectral weighting functions are frequently used in these integrations. Anticipating that discussion here, it is noted that equations (6.39) through (6.41) are not necessarily true if the quantities involved are integrated over different wavelength ranges or if different weighting functions are used in the integrating process. These expressions are true only at a specific wavelength, or in special circumstances when each quantity in these equations is integrated over the same wavelength range, using the same weighting function [29]. These considerations are discussed further in Section 6.9 on broadband spectral properties. Table 6.5 specifies the conditions where the equivalence shown in (6.39) holds for ranges of both wavelength and angle of incidence or emergence.

The need for such specific qualifications is illustrated as follows. A surface that looks like an excellent reflector in the visible portion of the spectrum may have a high emittance over most of the infrared spectrum, or vice versa. Spectrally selective solar absorbers work on this latter principle—they have high absorptance over the solar spectrum but low emittance over the long wavelength infrared portion of the spectrum, thereby reducing the escape of the thermal IR radiation (which occurs at longer wavelengths) by emission from the surface. More examples of spectral selectivity are given in Section 6.10.

The previous discussion assumed that the ray of interest is incident on a well-defined, planar interface between two different media and that the angle of incidence is with respect to the normal to the plane at the point of intersection of the ray. The results are applicable, at least on a local basis, to interfaces that are not strictly planar if one can determine the angle of incidence of the ray at the point of incidence on the interface.

For interfaces that are patterned or corrugated, one would have to apply the interface equations on a point-by-point basis over the area of interest. For interface surfaces that are rough, meaning that the surface normal varies in its direction on a random basis over distances across the interface that are quite small in comparison to the portion of the interface that is of interest, the above equations are not directly applicable. Reflection and transmission from such surfaces is often called diffuse to distinguish them from specular reflection and transmission at strictly smooth, planar surfaces where the surface normals are parallel over the whole area of the interface.

Table 6.5
Required Conditions for the Equality of Average Emittance and Absorptance

<i>Quantity</i>	<i>Equality</i>	<i>Required Conditions</i>
Directional spectral emittance	$\epsilon(\lambda, \theta, \phi) = \alpha(\lambda, \theta, \phi)$	None, except thermal equilibrium
Directional total emittance	$\epsilon(\theta, \phi) = \alpha(\theta, \phi)$	1. Spectral distribution of incident flux is proportional to that of a blackbody, or 2. $\epsilon_\lambda(\theta, \phi) = \alpha_\lambda(\theta, \phi)$ is independent of wavelength
Hemispherical spectral emittance	$\epsilon(\lambda) = \alpha(\lambda)$	1. Incident radiant flux is independent of angle, or 2. $\epsilon(\lambda, \theta, \phi) = \alpha(\lambda, \theta, \phi)$ is independent of angle
Hemispherical total emittance	$\epsilon = \alpha$	1. Incident radiant flux is independent of angle <i>and</i> has spectral distribution proportional to that of a blackbody, or 2. Incident radiant energy is independent of angle <i>and</i> $\epsilon(\lambda, \theta, \phi) = \alpha(\lambda, \theta, \phi)$ is independent of wavelength, or 3. Incident radiant flux at each angle has spectral distribution proportional to that of a blackbody <i>and</i> $\epsilon(\lambda, \theta, \phi) = \alpha(\lambda, \theta, \phi)$ is independent of angle, or 4. $\epsilon(\lambda, \theta, \phi) = \alpha(\lambda, \theta, \phi)$ is independent of angle and wavelength

(After Grum and Becherer [32]).

Diffusely transmitting and reflecting materials are discussed in Section 6.7.2. The angular dependence of optical properties and their integrations over ranges of angles are discussed in the next two sections. A comprehensive description of their spectral dependences is presented in Sections 6.8 and 6.9.

Before beginning those discussions, consider the case of a ray incident from air upon a parallel plate of transparent nonscattering material such as clear glass of known thickness w and refractive index n . The situation is illustrated in Figure 6.8. The Fresnel formulas presented in Section 6.3.4 can be used to determine the interface transmittance $\bar{\tau}$ and reflectance $\bar{\rho}$. A ray of unity radiance is incident at

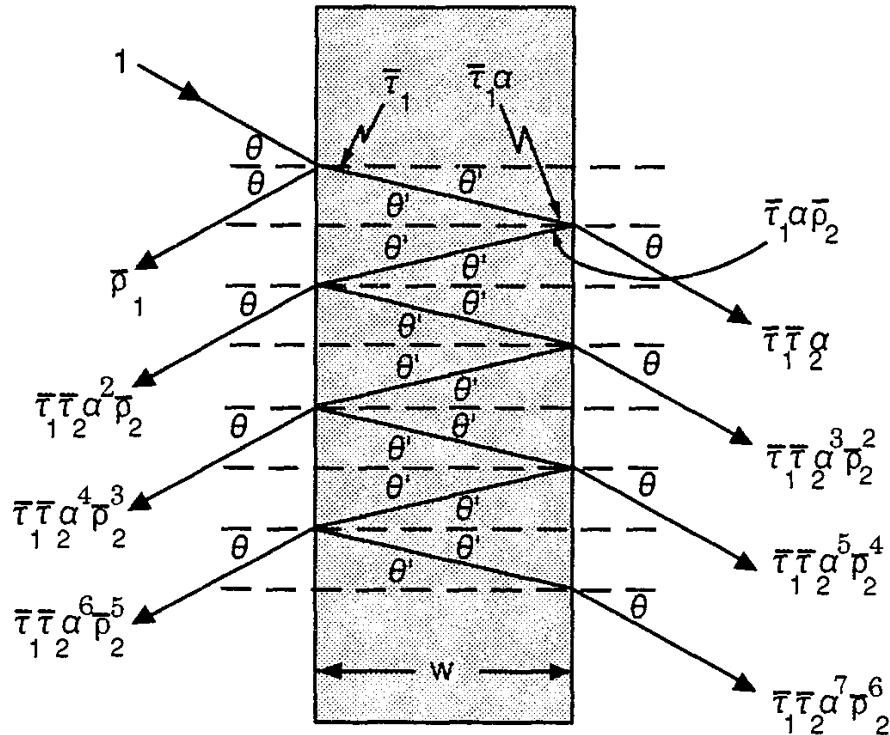


Figure 6.8 Illustration of multiple reflections between two interfaces.

the angle θ in air. The transmitted radiance is $\bar{\tau}_1$. (As shown in (6.14), the change in radiance upon entering the plate of refractive index n is included in the definition of $\bar{\tau}$ used here.) $\bar{\tau}_1$ is the transmittance from air to glass—from a refractive index of 1.0 to a refractive index of n . After the transmitted ray propagates across the medium having absorption coefficient a , its radiance will be reduced by a fraction called the transmittance due to absorption, designated α , to the value $\bar{\tau}_1 \alpha$, where

$$\alpha = \frac{L(x)}{L(0)} = e^{-ax} \quad (6.42)$$

Equation (6.42) comes from (6.27) with $c(\lambda)$ replaced by a , a being the absorption coefficient (6.29), and x being the ray distance through the medium of thickness w , given by

$$x = \frac{w}{\cos \theta'} \quad (6.43)$$

A portion of the ray will next be transmitted back into the air through the second interface with transmittance $\bar{\tau}_2$, given by (6.14) with the refractive indices reversed—going from a medium with refractive index n to one whose refractive index is 1.0. The reflected portion of the ray will reflect from the second interface,

with reflectance $\bar{\rho}_2$ given by (6.17), (6.19), and (6.21) with the refractive indices reversed—going from index n to index 1.0—and with θ_1 and θ_2 interchanged. After this reflection, the radiance of the ray will be $\bar{\tau}_1 \alpha \bar{\rho}_2$. It will then propagate back to the first face, where it will have radiance $\bar{\tau}_1 \alpha^2 \bar{\rho}_2$. Here it will be partially transmitted with transmittance $\bar{\tau}_2$, and partially reflected with reflectance $\bar{\rho}_2$. The ray emerging to the left will have radiance $\bar{\tau}_1 \bar{\tau}_2 \alpha^2 \bar{\rho}_2$ and the reflected ray will have radiance $\bar{\tau}_1 \alpha^2 \bar{\rho}_2^2$, as shown in Figure 6.8. This process of repeated reflections and transmissions continues on, as shown in the illustration, for a large number of times. The radiance L does not appear explicitly in these expressions because of our assumption of unity incident radiance.

The total reflectance R of the plate for rays incident at the angle θ will be the sum of all the radiances emerging to the left in Figure 6.8 and the transmittance T will be the sum of all the radiances emerging to the right of the plate. For a plate of infinite extent, there will be an infinite number of interreflections and the overall reflectance will be given by:

$$\begin{aligned} R &= \bar{\rho}_1 + \bar{\tau}_1 \bar{\tau}_2 \alpha^2 \bar{\rho}_2 + \bar{\tau}_1 \bar{\tau}_2 \alpha^4 \bar{\rho}_2^3 + \bar{\tau}_1 \bar{\tau}_2 \alpha^6 \bar{\rho}_2^5 + \dots \\ R &= \bar{\rho}_1 + \bar{\tau}_1 \bar{\tau}_2 \alpha^2 \bar{\rho}_2 (1 + \alpha^2 \bar{\rho}_2^2 + \alpha^4 \bar{\rho}_2^4 + \dots) \\ R &= \bar{\rho}_1 + \alpha^2 \bar{\tau}_1 \bar{\tau}_2 \bar{\rho}_2 \sum_{i=0}^{\infty} (\bar{\rho}_2 \alpha)^{2i} \end{aligned} \quad (6.44)$$

Similarly, the transmittance T of the plate will be given by

$$\begin{aligned} T &= \bar{\tau}_1 \bar{\tau}_2 \alpha + \bar{\tau}_1 \bar{\tau}_2 \alpha^3 \bar{\rho}_2^2 + \bar{\tau}_1 \bar{\tau}_2 \alpha^5 \bar{\rho}_2^4 + \bar{\tau}_1 \bar{\tau}_2 \alpha^7 \bar{\rho}_2^6 + \dots \\ T &= \bar{\tau}_1 \bar{\tau}_2 \alpha (1 + (\alpha \bar{\rho}_2)^2 + (\alpha \bar{\rho}_2)^4 + (\alpha \bar{\rho}_2)^6 + \dots) \\ T &= \bar{\tau}_1 \bar{\tau}_2 \alpha \sum_{i=0}^{\infty} (\alpha \bar{\rho}_2)^{2i} \end{aligned} \quad (6.45)$$

For $x \leq 1$, it is true that

$$\sum_{i=0}^{\infty} x^{2i} = \frac{1}{1 - x^2} \quad (6.46)$$

so that

$$\begin{aligned} R &= \bar{\rho}_1 + \frac{\bar{\tau}_1 \bar{\tau}_2 \bar{\rho}_2 \alpha^2}{1 - (\alpha \bar{\rho}_2)^2} \\ T &= \frac{\bar{\tau}_1 \bar{\tau}_2 \alpha}{1 - (\alpha \bar{\rho}_2)^2} \end{aligned} \quad (6.47)$$

and

$$A = 1 - R - T \quad (6.48)$$

It can be shown using (6.11) through (6.14) that for the same refractive index n on both sides of the plate, $\bar{\rho}_1 = \bar{\rho}_2$ and the product $\bar{\tau}_1 \bar{\tau}_2$ is equal to $(1 - \bar{\rho}_2)^2$, thereby simplifying (6.47) somewhat. Values for R and T in (6.47) are plotted for a medium of thickness $w = 0.003\text{m}$, absorption coefficient $a = 0.01\text{m}^{-1}$, and refractive index $n = 1.55$ in Figure 6.9. The equivalency of $\bar{\rho}_1(\theta_1)$ and $\bar{\rho}_2(\theta_2)$ was derived by G. G. Stokes [30] using a principle of optical reversibility.

6.5.2 Scattering Media

Consider the case, previously postponed, of an absorbing *and* scattering medium. Determining the bulk transmittance T , reflectance R , and absorptance A of a parallel plate of such material is not an easy problem in general because of the possibility of multiple scattering effects within the medium, in addition to the multiple reflections dealt with previously.

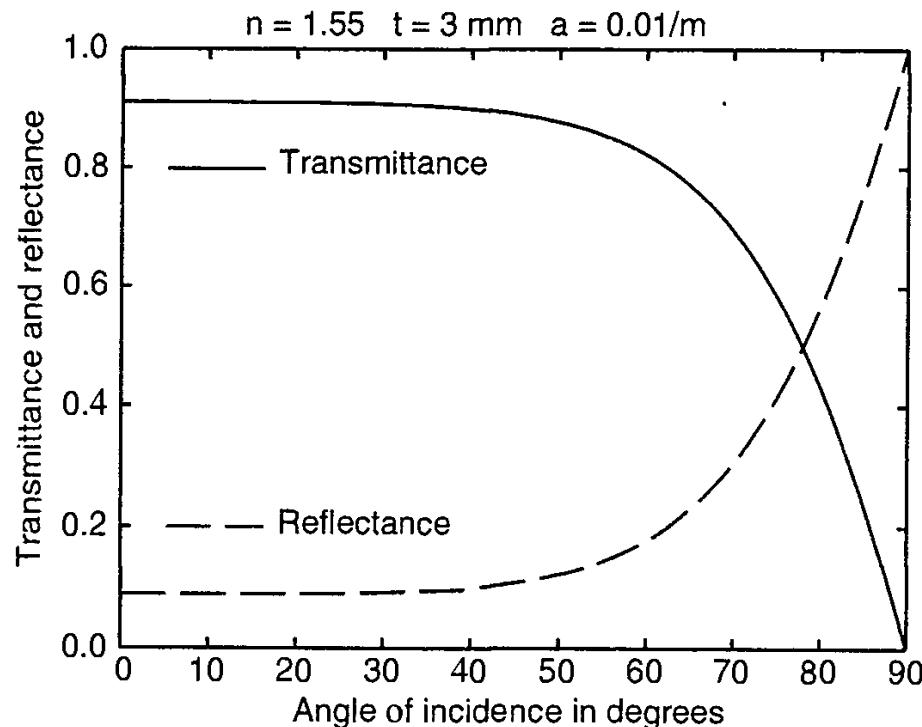


Figure 6.9 Transmittance and reflectance for a 3mm-thick plate of glass of refractive index 1.55 and absorption coefficient 0.01m^{-1} .

If single scattering only is assumed—that photons once scattered by a particle within the medium are never scattered again, but subsequently obey the laws of reflection, refraction, and attenuation described above—the problem becomes somewhat tractable. A start on this problem can be made by solving (6.35) for the element $dL_{\lambda,s}(\theta')$ of radiance scattered into the direction θ' by the element dx of distance along a ray in the medium (see Figure 6.10).

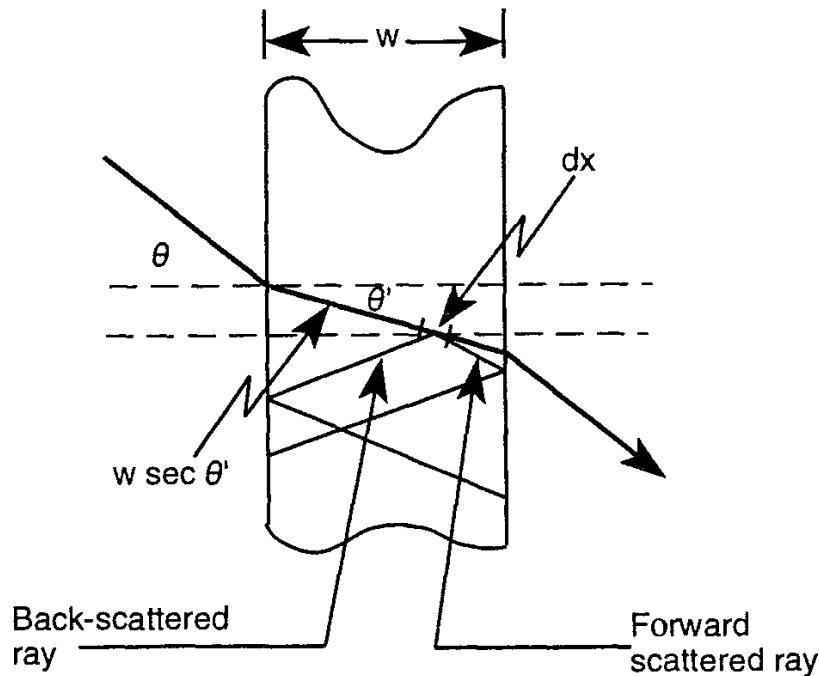


Figure 6.10 Geometry for scattering within a parallel plate of thickness w .

The resulting expression can then be integrated over the path (of length $w/\cos \theta$) of the ray through the medium from one face of the parallel plate to the other. This will give the scattered radiance of the plate, inside it, in terms of the irradiance E_λ from (6.35) in a plane perpendicular to the ray in the medium. With some care, this can be replaced with an expression depending on the irradiance $E_{0\lambda}$ incident on the outside of the plate. Similarly, the scattered radiance inside the plate can be replaced with an expression containing the radiance outside the plate. The resulting expression can finally be manipulated so that it gives the radiance of the plate for any direction of emergence from the plate. This final result will still represent only the singly scattered radiance in the direction specified by the angle θ' . The incident ray is reflected at the second interface of the medium and travels back across the plate, reflecting again at the incident interface, the process being repeated an infinite number of times. Each traversal of the medium produces scattered radiation at all angles of scattering. The scattered rays also are multiply reflected at the medium's interfaces. In the single-scattering approximation, one presumes that the scattered rays are not again scattered but propagate through the

medium, obeying only the laws of reflection, refraction, and attenuation presented above.

To complete this calculation of the reflectance and transmittance from a parallel plate of material exhibiting scattering effects requires a detailed treatment of the multiple reflections of the scattered rays, much as was done above for nonscattering media, and a summing up of all the scattered components of the emerging radiation. It should be assumed that the incident ray is incident on the parallel plate at some arbitrary angle of incidence θ_i and the scattered radiance emerges at an arbitrary angle θ_s to the normal to the plate.

This problem subsequently becomes quite involved, beyond the scope of this book, and is not completed here. Problems such as this are seldom solved by appealing to the scattering properties of the material directly, but by measurement of the bulk optical properties of a whole object or sample of it.

A full treatment of multiple scattering effects is beyond the scope of this book. The above single-scattering discussion was included only to illustrate the use of the radiometric definitions that have been introduced. The interested reader is referred to more comprehensive treatments of radiative transfer in homogeneous and inhomogeneous media exhibiting significant multiple scattering effects [31–33].

6.6 ANGULAR DEPENDENCE

The key to a full understanding of the angular or directional dependence of material optical properties lies in the concept of solid angle, discussed extensively in Chapter 1. Radiation incident upon a point in a surface can come to that point from many directions. The concept of pencils of rays was introduced in the previous chapter as a means of specifying solid angles of incidence on and emergence from a point. This concept is useful in describing the directional dependences of transmittance and reflectance, for both theoretical treatments and practical measurements.

In making a transmittance or reflectance measurement, for example, a sample to be tested is illuminated with radiation filling some solid angular range of directions—a range of incident directions. The reflected or transmitted flux is then collected over another range of directions within some solid angle, a range of emergent directions.

In order for the transmittance or reflectance value to have meaning, either theoretically or experimentally, it is essential that the directions and solid angles of incidence *and* exitance be specified. These tell the range of angles involved in the measurement and is equivalent to giving the size of the (usually conical) solid angles and their directions of incidence and emergence.

In reflectance and transmittance, there are three kinds or sizes of solid angles of interest and several different definitions of reflectance and transmittance using

these. The three solid angles are *directional*, *conical*, and *hemispherical*. They are illustrated in Figure 6.11.

The term “directional” refers to an infinitesimally small solid angle in a specified direction from the point P of interest. Infinitesimally small solid angles are used mainly in mathematical treatments of reflectance and transmittance—not in actual measurements, since an infinitesimally small solid angle will contain an infinitesimally small, and hence unmeasureable, quantity of flux.

The term “conical” refers to an intermediate-sized solid angle, usually in the shape of a right circular cone, but other shapes can also be used. The term “hemispherical” refers to a full 2π steradians hemispherical solid angle. The terms direction, conical, and hemispherical may be applied to the incident and the emergent directions of flux flow, yielding nine possible different definitions of reflectance and transmittance [34]. The geometrical basis for these definitions is illustrated in Figures 6.11 through 6.14 and the list of nine definitions is as follows:

- Bidirectional;
- Directional-conical;
- Directional-hemispherical;
- Conical-directional;
- Hemispherical-directional;
- Biconical;
- Conical-hemispherical;
- Hemispherical-conical;
- Bihemispherical.

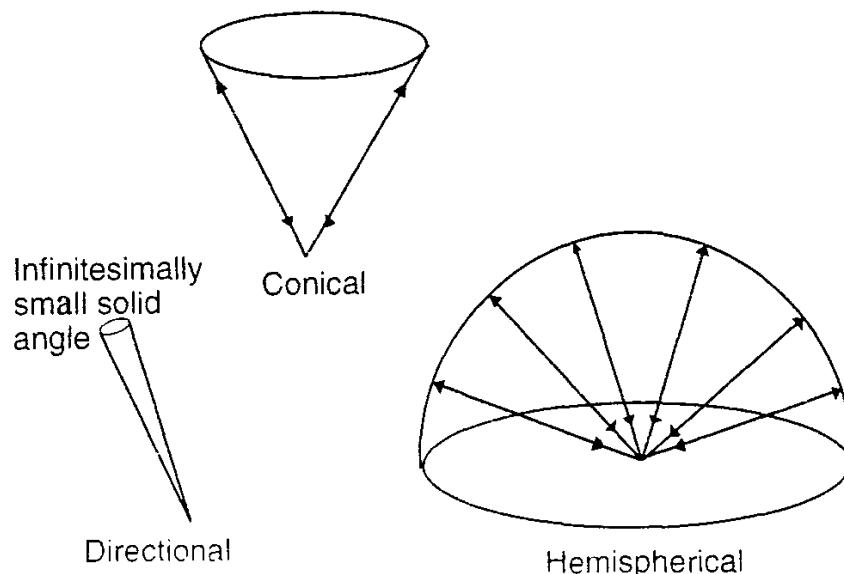


Figure 6.11 Geometry for directional, conical, and hemispherical solid angles.

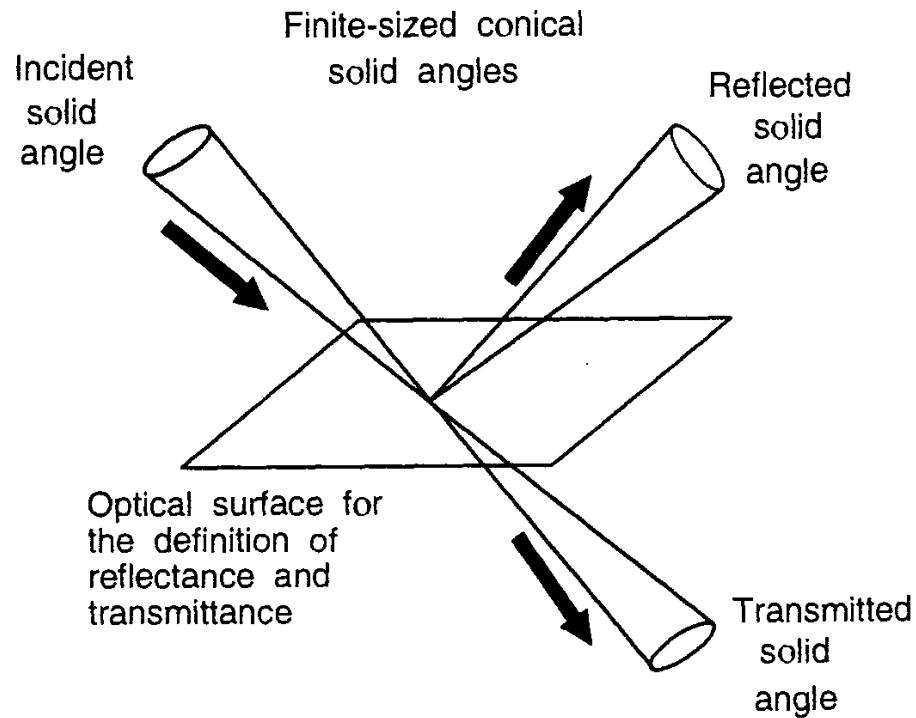


Figure 6.12 Geometry for the definition of biconical transmittance and reflectance.

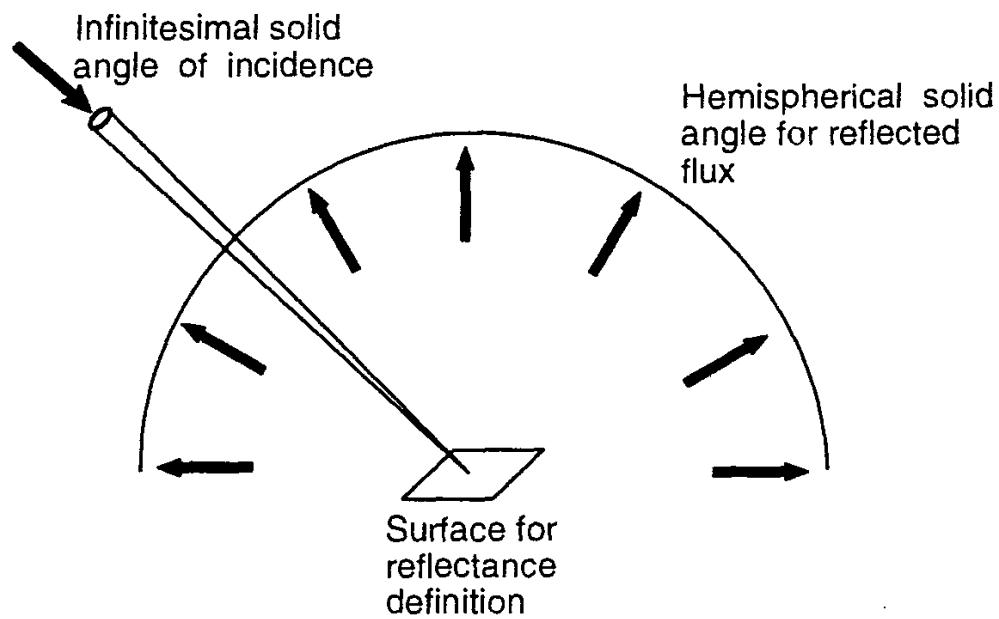


Figure 6.13 Geometry for the definition of directional-hemispherical reflectance.

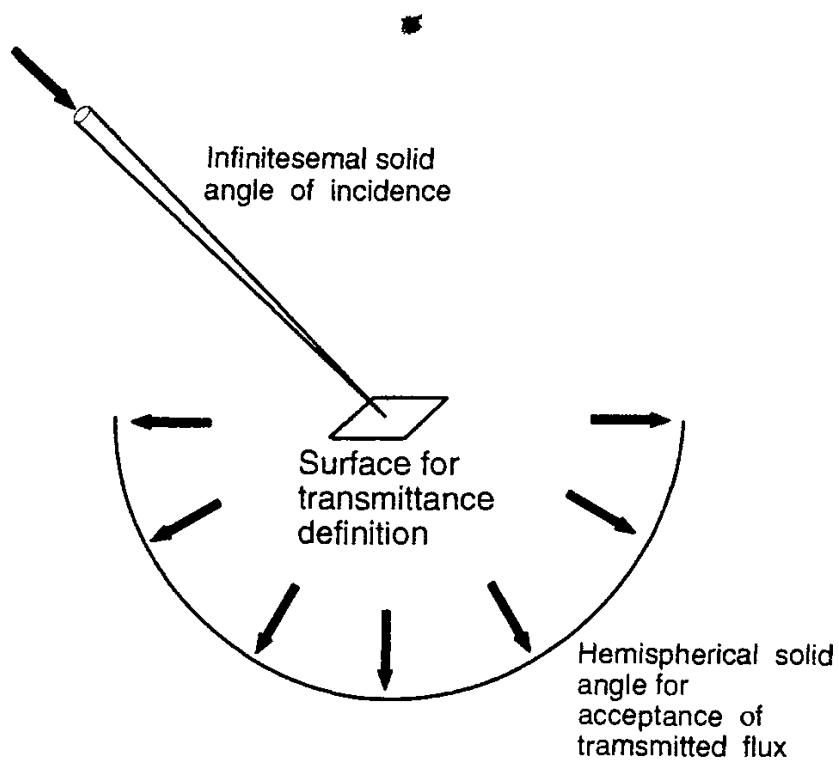


Figure 6.14 Geometry for the definition of directional-hemispherical transmittance.

Not all of these are used in practice. The most common definitions are:

- Bidirectional;
- Directional-conical;
- Directional-hemispherical;
- Biconical;
- Conical-hemispherical;
- Bihemispherical.

The first three of these are mainly found in theoretical discussions. The last three are used in reflectance and transmittance measurements. Solar optical property standards published by various organizations refer to conical-hemispherical measurements. This is because for most practical problems, it is only the total transmitted irradiance due to the directly incident beam alone that is of interest. For other applications and complex situations, the bi-conical definition is important to have as well. Other applications have need of other optical properties. Theoretical treatments of radiative transfer deal almost exclusively with the directional versions of the definitions. Sometimes the terminology "bidirectional" is used to refer to bi-conical measurements where the solid angles involved are relatively small.

In this discussion so far, and in the drawings of Figures 6.12 through 6.14, it is implied that transmittance and reflectance measurements are made by focusing

a pencil of rays onto a small point on the sample being measured. In reality, the incident beam is made to cover a finite area on the sample, hopefully large with respect to any spatial inhomogeneities across the sample surface. The concept of a conical pencil of rays incident on a point on the sample is still valid, however. The reason is that the actual flux incident on each illuminated point of the sample does arrive in some defined solid angle. It is simply that a lot of points on the sample are so illuminated. A desirable goal in the design of transmissometers and reflectometers is to have all portions of the illuminated area of a sample receive flux from the same range of directions, over the same solid angular ranges.

A similar argument holds with respect to measurements of transmitted or reflected flux. Although the reception and detection subsystems collect flux from a finite portion of the sample area, the actual flux detected from each point in the field of view on the sample will enter the reception subsystem over some solid angular range. A design goal is to keep the range of detected angles small, filling a relatively small solid angle from each point on the sample.

Commonly found in the literature is the term *albedo*, which relates to the reflectance of objects or surfaces. According to an encyclopedia of science and technology [35], there are two meanings to this term. *Bond albedo* is used when discussing the energy balance of planets. It is the fraction of incident solar radiation reflected back out into space. According to the above terminology, bond albedo refers to the conical-hemispherical reflectance of the planet. *Normal albedo* refers to the normal incidence reflectance of a surface, probably referring again to the conical-hemispherical reflectance. Other usages for the term albedo can be found in the literature.

6.7 BROADBAND ANGLE PROPERTIES

As indicated previously, materials generally have a directional dependence in their responses to incident radiation. Sources of radiation incident on these materials also have some angular dependence. The proper way to determine the response of a material to an incident distribution of flux is to multiply the incident flux by the optical property on a direction-by-direction (and wavelength-by-wavelength) basis. Nicodemus and others have written extensively on this subject [34].

6.7.1 Transmittance and Reflectance Equations

According to Nicodemus [34], the transmission process is embodied in the fundamental equation

$$dL_i(\theta_i, \phi_i, \theta_t, \phi_t, E_i) = f_i(\theta_i, \phi_i, \theta_t, \phi_t) dE_i(\theta_i, \phi_i) \quad (6.49)$$

where:

$dE_i(\theta_i, \phi_i)$ is the element of incident (spectral) irradiance (considered uniform over the transmitting surface and approaching from the direction (θ_i, ϕ_i) of incidence within the element of solid angle $d\omega_i$),

$dL_t(\theta_t, \phi_t, \theta_i, \phi_i, E_i)$ is the element of transmitted (spectral) radiance, and

$f_t(\theta_i, \phi_i, \theta_t, \phi_t)$, having units of steradian⁻¹, is called the bidirectional (spectral) transmittance distribution function (BTDF).

The BTDF transforms incident irradiance into transmitted radiance. If the BTDF (or its wavelength-integrated broadband version) is known, one can integrate the resulting angular variation in transmitted radiance over whatever solid angle is of interest. This same reasoning applies to reflectance, and there is a bidirectional reflectance distribution function (BRDF) that is analogous to the BTDF.

The wavelength subscript has again been dropped from the quantities in (6.49), since their spatial variations are of more concern here. Each quantity, however, can be thought of as a spectral quantity or as an integral of the spectral quantity over a defined wavelength interval.

The BTDF and the BRDF are useful concepts, but they cannot be measured, because truly infinitesimal elements of solid angle include infinitesimal and hence unmeasurable quantities of flux. Following the treatment of Nicodemus and others [34], a general expression is presented for transmittance in cases where the geometry (solid angles) of the incident (*i*) and transmitted (*t*) beams can be specified:

$$T(\omega_i, \omega_t, L_i) = \frac{\int_{\omega_i} \int_{\omega_t} f_t(\theta_i, \phi_i, \theta_t, \phi_t) L_i(\theta_i, \phi_i) \cos \theta_i \cos \theta_t d\omega_t d\omega_i}{\int_{\omega_i} L_i(\theta_i, \phi_i) \cos \theta_i d\omega_i} \quad (6.50)$$

Replacing T with R and subscript *t* with *r* converts (6.50) to an equation for the reflectance. $T(\omega_i, \omega_t, L_i)$ is the biconical transmittance. It is dependent upon the incident radiance. Changing the angular distribution of incident radiance could result in a change in the transmittance, so the transmittance in this definition depends upon the incident radiance distribution. One can think of this radiance distribution as a weighting function, weighting the BTDF and BRDF with the directional distribution of L_i over the solid angle ω_i , in determining the transmittance T and reflectance R . (This is similar to what is done below with spectral distributions. The spectral transmittance is weighted by the incident spectral irradiance distribution to determine the total transmittance over the whole spectrum.) If the incident solid angle ω_i is small, there may be little variation in L_i over it. If ω_i is large, there could be considerable variation in L_i over it.

If the incident radiance L_i is uniform over the solid angle ω_i of incidence, then it can be taken out of the integrals in the numerator and denominator of (6.50) and it cancels. The dependence of the BTDF and the BRDF on L_i disappears

in this case, leaving the biconical transmittance for incident radiance uniform over ω_i :

$$T(\omega_i, \omega_t) = \frac{\int_{\omega_i} \int_{\omega_t} f_t(\theta_i, \phi_i, \theta_t, \phi_t) \cos \theta_i \cos \theta_t d\omega_i d\omega_t}{\int_{\omega_i} \cos \theta_i d\omega_i} \quad (6.51)$$

which is dimensionless. The condition that L_i be constant over the incident beam is closely met in well-designed and well-built transmissometers. It is always important when making measurements to specify the geometry of both the incident and transmitted beams associated with a given value of transmittance.

Note that removing the integrals from numerator and denominator in (6.51) converts it to a definition of bidirectional transmittance. Making ω_i and ω_t small but finite in (6.51) converts it to a definition of biconical transmittance. Making ω_i small but finite and $\omega_t = 2\pi$ sr converts (6.51) to a definition of conical-hemispherical transmittance.

The above discussion is equally valid for reflectance, by replacing T and t in the above equations with R and r , respectively. For measurements of conical and hemispherical incidence versions of the transmittance and reflectance where the incident radiance is not constant over the solid angle of incidence, (6.50) (or its reflectance version) must be used and the resulting transmittance or reflectance will be weighted by the angular distribution of incident radiance. Readers interested in exploring the topic of reflectance and transmittance more thoroughly are referred to the important publication by Nicodemus and others on this subject [34].

It was mentioned in Section 6.4.1 that Kirchhoff's law can be extended to finite wavelength ranges under special conditions. This applies as well to finite angular ranges. Table 6.5, taken from Grum and Becherer [29], specifies these conditions.

6.7.2 Specular and Diffuse Optical Properties

It is useful to speak of *specular* and *diffuse* reflectance and transmittance. These terms refer to the degree of random scattering effects within the material or on its surface(s). Specular reflectance and transmittance occur when all parallel incident rays are reflected and transmitted as parallel rays, with their angles of reflection all equaling their angles of incidence and the angles of transmission also equaling the angles of incidence or the angle of refraction specified by Snell's law, depending upon the medium in which they propagate, i.e., according to the laws of geometrical

optics. There are no scattering effects in specular reflectance and transmittance. Specular reflectance is also called *regular reflectance*.

With diffuse reflectance and transmittance, the emerging rays have angles of reflection or transmission that are different from each other. A small beam of parallel rays incident on a diffusely reflecting and transmitting surface or medium will be converted to a spread of rays distributed over the hemispherical solid angle of emergence. There is great variety in how these rays (and the corresponding fluxes) are distributed over the emerging solid angle. Some materials produce strong forward scattering effects while others scatter almost equally in all directions. If the emerging radiance is constant for all directions in a hemispherical solid angle, then the surface is said to be a Lambertian reflector or transmitter. The term "diffuse" generally conveys semi-Lambertian properties, but it is also used for surfaces that are non-Lambertian but with substantial nonspecular components.

Most real surfaces can be thought of as having a mixture of specular and diffuse optical properties. In real situations a specularly reflecting and transmitting material is one whose diffuse transmittance and reflectance are very much less than the specular versions of these quantities. A diffusely reflecting and transmitting material is one whose specular transmittance and reflectance are comparable to or much less than the diffuse properties. Perfectly specular reflection and transmission have no diffuse component, but perfectly diffuse reflection and transmission can of course have some flux reflected or transmitted in the specular direction.

A standard terminology has been developed to deal with specular, mixed, and diffuse reflection and transmission properties [1]. In addition to the usual definitions for reflection and transmission, the remaining terminology from Ref. [1] is summarized below:

Diffusion; scattering: The process by which the spatial distribution of a beam of radiation is changed when it is deviated in many directions by a surface or by a medium, without change of frequency or of its monochromatic components. *Non-selective diffusion* refers to diffusion that does not vary with the wavelength of the incident radiation.

Regular reflection/transmission; specular reflection/direct transmission: Reflection or transmission in accordance with the laws of geometrical optics, without diffusion.

Diffuse reflection/transmission: Diffusion by reflection/transmission in which, on the macroscopic scale, there is no regular reflection/transmission.

Mixed reflection/transmission: Partly regular and partly diffuse reflection/transmission.

Isotropic diffuse reflection/transmission: Diffuse reflection/transmission in which the spatial distribution of the reflected/transmitted radiation is such that the radi-

ance or luminance is the same in all directions in the hemisphere into which the radiation is reflected/transmitted.

Perfect reflecting/transmitting diffuser: Ideal isotropic diffuser with reflectance/transmittance equal to 1.

Lambertian surface: Ideal surface for which the radiation coming from that surface is distributed angularly according to Lambert's cosine law (see Section 1.5). Note that the definition refers to both reflecting and transmitting surfaces, as well as self-luminous ones, regardless how the former may be illuminated.

Retroreflection: Reflection in which radiation is returned in directions nearly opposite to the direction from which it came, this property being maintained over wide variations of the direction of the incident rays.

Regular (specular) reflectance/transmittance (ρ_s , R_s/τ_s , T_s): Ratio of the regularly reflected/transmitted part of the reflected flux to the incident flux.

Diffuse reflectance/transmittance (ρ_d , R_d/τ_d , T_d): Ratio of the diffusely reflected/transmitted part of the reflected/transmitted flux, to the incident flux.

Reflectance factor (R) (at a surface element, for the part of the reflected radiation contained in a given cone with apex at the surface element, and for incident radiation of given spectral composition, polarization and geometrical distribution): Ratio of the radiant or luminous flux reflected in the directions delimited by the given cone to that reflected in the same direction by a perfect reflecting diffuser identically irradiated or illuminated.¹

Radiance factor is defined by the CIE as the ratio of the radiance of a surface element of a non-self-radiating medium in a given direction to that of a Lambertian surface identically irradiated. A similar definition is used for *luminance factor*.

The discussion of spectral dependence in the optical properties of materials and media has been suspended to this point, for the sake of simplifying the discussion. The next section deals with this important subject.

6.8 SPECTRAL DEPENDENCE

In general, the optical properties of materials are all functions of the wavelength (and thus the frequency) of the radiation incident upon them. This means that any spectral distribution of radiation incident upon materials will be altered in mag-

¹For regularly reflecting surfaces that are irradiated or illuminated by a beam of small solid angle, the reflectance factor may be much larger than 1 if the cone includes the mirror image of the source. If the solid angle of the cone approaches 2π sr, the reflectance factor approaches the reflectance for the same conditions of irradiation. If the solid angle of the cone approaches zero, the reflectance factor approaches the radiance or luminance factor for the same conditions of irradiation.

nitude and shape by reflection from or transmission through them. A central problem in radiometry is the determination of the emergent spectral distribution given the optical properties of the components of an optical system.

It is important to transform the incident spectral radiance, irradiance, or radiant flux by the relevant spectral optical property of the material, on a wavelength-by-wavelength basis, *before* the summing up or integration process is pursued. The process can be specified symbolically by the equation

$$Q_e(\lambda) = T[Q_i(\lambda)] \quad (6.52)$$

where $Q_i(\lambda)$ is the incident spectral distribution of radiant flux, irradiance, or radiance, $Q_e(\lambda)$ is the emergent spectral distribution, and T is a generalized transformation operator which takes an incident spectral (and angular) distribution of flux and converts it to an emergent spectral (and angular) distribution of flux.

The transformation T usually consists of a multiplication of the incident spectral radiant flux irradiance, or radiance by the spectral transmittance or reflectance of the material at each wavelength for which data is available (and for each different direction of incidence). (For more complex processes such as fluorescence, the transformation is not so simple.) The multiplication must be performed on a wavelength-by-wavelength basis to determine the spectrum of emerging radiance or irradiance, as indicated by the following equations:

$$\begin{aligned} Q_i(\lambda) &= T(\lambda)Q_i(\lambda) \\ Q_r(\lambda) &= R(\lambda)Q_i(\lambda) \end{aligned} \quad (6.53)$$

The process is illustrated schematically for transmittance in Figure 6.15. The total emerging flux, radiant intensity, irradiance, or radiance can then be determined through the integrations given in (2.8) through (2.11) in Chapter 2. The emerging photopic quantities can be determined through the application of (2.1) to those given in (6.53).

6.9 BROADBAND SPECTRAL PROPERTIES

Early in Chapter 1, it was pointed out that mistakes can be made when trying to mix spectra having different spectral distributions without being careful to do so on a wavelength-by-wavelength basis. For example, when one knows an incident spectral irradiance distribution, $E_{\lambda i}$, and the spectral transmittance $T(\lambda)$ of an optical filter or a pane of glass or a window in an analytical instrument, the proper way to determine the transmitted irradiance distribution $E_{\lambda t}$ is by using the equation

$$E_{\lambda t} = E_{\lambda i}T(\lambda) \quad (6.54)$$

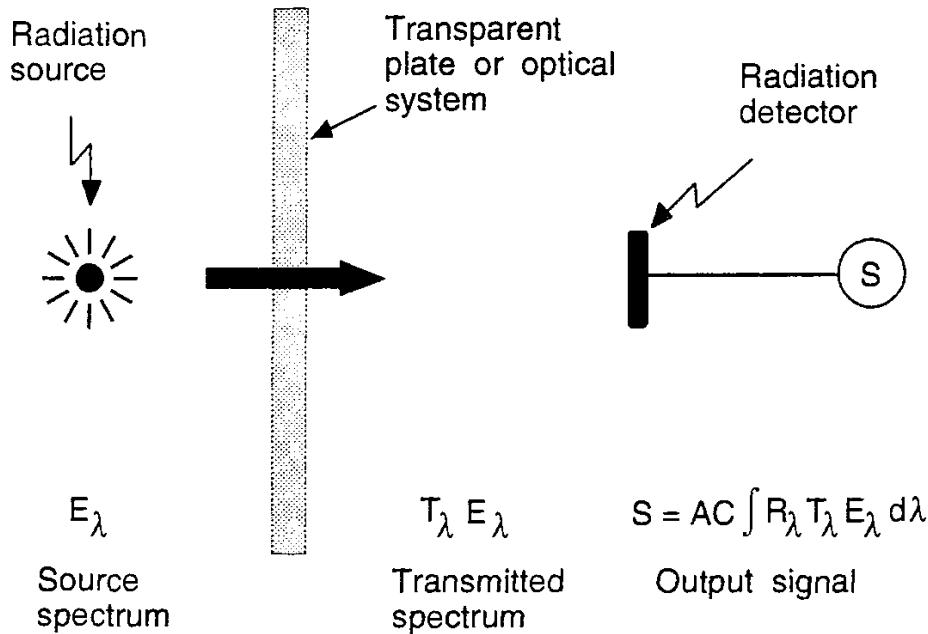


Figure 6.15 Schematic diagram for determining a transmitted spectral distribution. A is the area of the radiation detector and R_λ is its C is its calibration constant spectral response.

for each wavelength that $E_{\lambda i}$ is nonzero. The total integrated transmitted irradiance E_t is determined from

$$E_t = \int_0^\infty E_{\lambda i} d\lambda = \int_0^\infty E_{\lambda i} T(\lambda) d\lambda. \quad (6.55)$$

Note that the single integral on the right of (6.55) is not generally the same as the product of the integrals of $E_{\lambda i}$ and $T(\lambda)$ separately. However, by analogy with the definitions of transmittance and reflectance given above, if $E_{\lambda i}$ has the constant value E_i over the range of nonzero values of $T(\lambda)$, then it can be removed from the integral and we can write $E_t = E_i T$, with T being the integral of $T(\lambda)$ over its nonzero range. Similarly, if $T(\lambda)$ has the constant value T over the range of nonzero values for $E_{\lambda i}$, then it can be removed from the integral and we again have $E_t = E_i T$, where now E_i is the integral of $E_{\lambda i}$.

Seldom in real situations are either $T(\lambda)$ or $E_{\lambda i}$ constant over the spectral range of interest. There are situations, however, where at least the shapes of both of these spectra are constant. The magnitude of $E_{\lambda i}$ might vary with time, while its relative spectral distribution remains essentially constant. In such a case, we can write $E_{\lambda i}$ as a product of the variable portion, E_{oi} , and the constant spectral shape portion, $w(\lambda)$. In this case, (6.55) becomes

$$E_t = \int_0^\infty E_{\lambda i} d\lambda = E_{oi} \int_0^\infty w(\lambda) T(\lambda) d\lambda \quad (6.56)$$

showing the role of $w(\lambda)$ as a weighting function. We call the integral on the right side of (6.56) the weighted transmittance

$$T_w = \int_0^{\infty} w(\lambda) T(\lambda) d\lambda \quad (6.57)$$

The operation shown in this equation forms the basis for a number of definitions in several fields of study of broadband or integrated transmittance and reflectance. For example, in the field of illumination engineering there are several light source spectral distributions that have been standardized for use in (6.57) so that everyone talking about a single-number transmittance value for a sheet of transparent material used in a light fixture or window will be talking about essentially the same thing. These procedures involve multiplying a standardized incident irradiance distribution ϵ_{λ} (thought to be typical of the intended application) with the spectral transmittance (or reflectance) of the material to produce the transmitted spectral irradiance distribution. The result is then divided by the integral of the incident spectral irradiance. This normalizes the weighting function $w(\lambda)$ in (6.57) as follows:

$$w(\lambda) = \frac{E_{\lambda}}{\int_0^{\infty} E_{\lambda} d\lambda} \quad (6.58)$$

The integral of $w(\lambda)$ over all wavelengths is 1.0. Table 6.6 lists examples of some commonly used spectrally weighted optical properties. These include what are called the *photopic* optical properties, where the weighting function is $V(\lambda)$, the human photopic spectral luminous efficiency function given in Table 2.1.

Some of the concepts discussed here can also be applied to spectral emittance, and one can speak of the long-wavelength emittance, with $\epsilon(\lambda)$ integrated over all wavelengths, say, above $4 \mu\text{m}$. Table 6.5 in Section 6.5.1 gives conditions for which the broadband versions ϵ and α of the spectral emittance $\epsilon(\lambda)$ and the spectral absorptance $\alpha(\lambda)$ can be considered equal.

6.10 SPECTRAL SELECTIVITY

The term *spectral selectivity* means that one or more of the optical properties of a surface or transparent object varies with wavelength, so that the object behaves differently in one part of the spectrum than in another. There is much variety in spectral selectivity, depending upon which optical property most exhibits the effect and which wavelength regions are involved.

An example can be offered. Figure 6.16 shows the hemispherical irradiance spectrum of a 23.9°C (75°F) blackbody, the V -lambda curve, and a solar spectrum, plotted together on the same wavelength scale to show the spectral separation

Table 6.6
Spectrally Weighted Optical Property Definitions

Symbol	Name	Defining Equation
T_v, R_v	Photopic or luminous transmittance or reflectance	$T_v = \frac{\int \tau_\lambda V_\lambda d\lambda}{\int V_\lambda d\lambda}$ $R_v = \frac{\int \rho(\lambda) V(\lambda) d\lambda}{\int V(\lambda) d\lambda}$ <p>$V(\lambda)$ is the CIE photopic spectral luminous efficiency function, <i>IES Lighting Handbook</i>, 1981, Ref. Vol., pp. 3-5.</p>
T_v, R_v	Solar photopic transmittance or reflectance	$T_v = \frac{\int \tau(\lambda) E_\lambda V(\lambda) d\lambda}{\int E_\lambda V(\lambda) d\lambda}$ $R_v = \frac{\int \rho(\lambda) E_\lambda V(\lambda) d\lambda}{\int E_\lambda V(\lambda) d\lambda}$ <p>$V(\lambda)$ is as above. E_λ is the terrestrial direct normal solar irradiance for air mass 1.5, ASTM E891-87. Also ASTM E971-88.</p>
$\epsilon(T)$	Emittance	$\epsilon(T) = \frac{\int_0^\infty E_{bb\lambda}(T) \epsilon(\lambda) d\lambda}{\int_0^\infty E_{bb\lambda}(T) d\lambda}$
T_s, R_s	Solar transmittance or reflectance	$T_s = \frac{\int \tau(\lambda) E_\lambda d\lambda}{\int E_\lambda d\lambda}$ $R_s = \frac{\int \rho_\lambda E_\lambda d\lambda}{\int E_\lambda d\lambda}$ <p>E_λ is the terrestrial direct normal solar irradiance for air mass 1.5, ASTM E891-87.</p>
T_v, R_v	Light transmittance or reflectance	$T_v = \frac{\int \tau(\lambda) D(\lambda) V(\lambda) d\lambda}{\int D(\lambda) V(\lambda) d\lambda}$ $R_v = \frac{\int \rho(\lambda) D(\lambda) V(\lambda) d\lambda}{\int D(\lambda) V(\lambda) d\lambda}$

Table 6.6 (Continued)

<i>Symbol</i>	<i>Name</i>	<i>Defining Equation</i>
T_{uv}	UV - transmittance	$V(\lambda) \text{ is as above. } D(\lambda) \text{ is the relative spectral power distribution of illuminant } D_{65}, \text{ CIE Publ. No. 15.2 Colorimetry (2nd ed.), Vienna, 1986.}$ $T_{uv} = \frac{\int_{280 \text{ nm}}^{380 \text{ nm}} \tau(\lambda) U(\lambda) d\lambda}{\int_{280 \text{ nm}}^{380 \text{ nm}} U(\lambda) d\lambda}$ $U(\lambda) \text{ is the relative spectral distribution of the UV part of the global solar radiation, ISO 9050 : 1990 (1st ed.).}$

between these spectra. There is by definition one air mass going from the surface of the earth straight upward through the atmosphere to outer space. Thus a 1.5 air mass refers to a slant path at an angle of about 48 deg to the zenith. The solar spectrum shown in Figure 6.16 is for this slant path.

One can take advantage of the spectral separation between the incident solar flux and the long-wavelength flux emitted by surfaces inside buildings to design

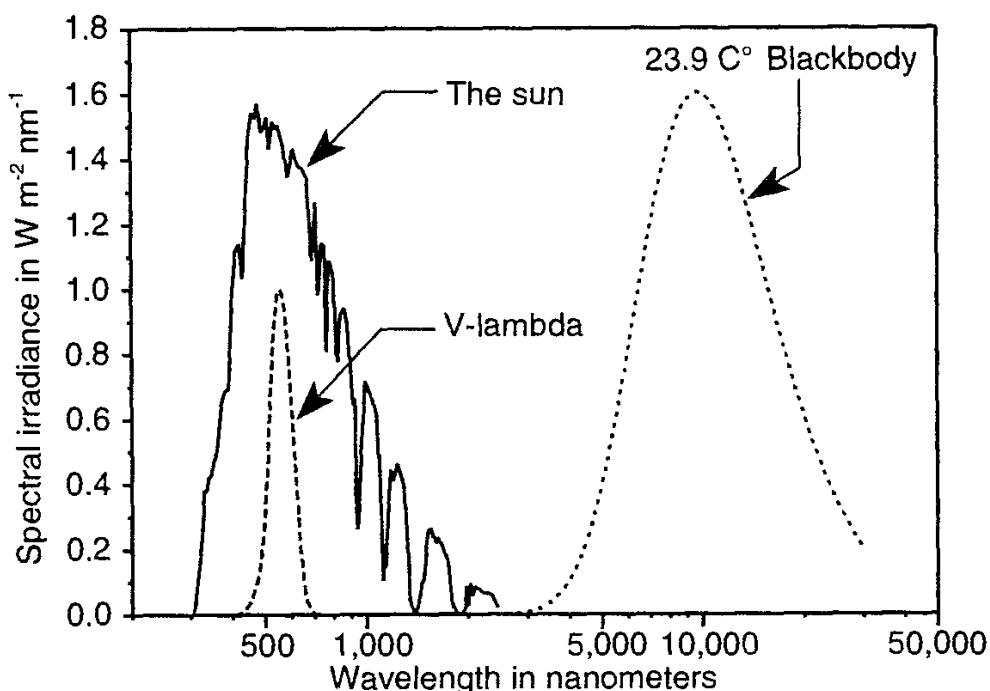


Figure 6.16 Sample spectral plots, showing the terrestrial solar spectrum, the human eye response spectrum, and the spectrum of emission from a 23.9°C (75°F) blackbody.

improved windows for buildings. The solar spectrum on the left in Figure 6.16 produces heat when it is absorbed inside the building. The spectrum on the right represents (Fig. 6.17) what is emitted by the room surfaces and furnishings and is then incident on the windows from the inside of the building.

Sometimes it is desired to admit and trap solar heat inside a building, preventing its escape. At other times, such as in hot climates, one wants to prevent the solar radiation from coming in at all. In both cases, it is usually desired to have a reasonably high light transmittance to admit daylight illumination and permit views of the outdoors. Either objective can be approximated with suitably chosen spectrally selective window glazings. An ideal window intended for a *cold* climate would have high spectral transmittance over the solar spectrum, such as the dashed curve in Figure 6.17, permitting maximum solar heat entry, but also high reflectance (low emittance) over the long wavelength thermal infrared portion of the spectrum. The high reflectance and low emittance insure that long-wavelength thermal radiation from the interior (or the innermost and hence warmest glazing of a multipane system) will be largely prevented from escaping through the window.

Most U.S. window manufacturers, and many in other developed countries, now offer what are called "low-e" coatings on one of the interior surfaces of their multiple-pane insulated windows intended for residences in cold climates. From Table 6.2, it is seen that the emittance of uncoated window glass is 0.84. Low-e coatings applied to this glass can lower the emittance to 0.10, nearly a tenfold

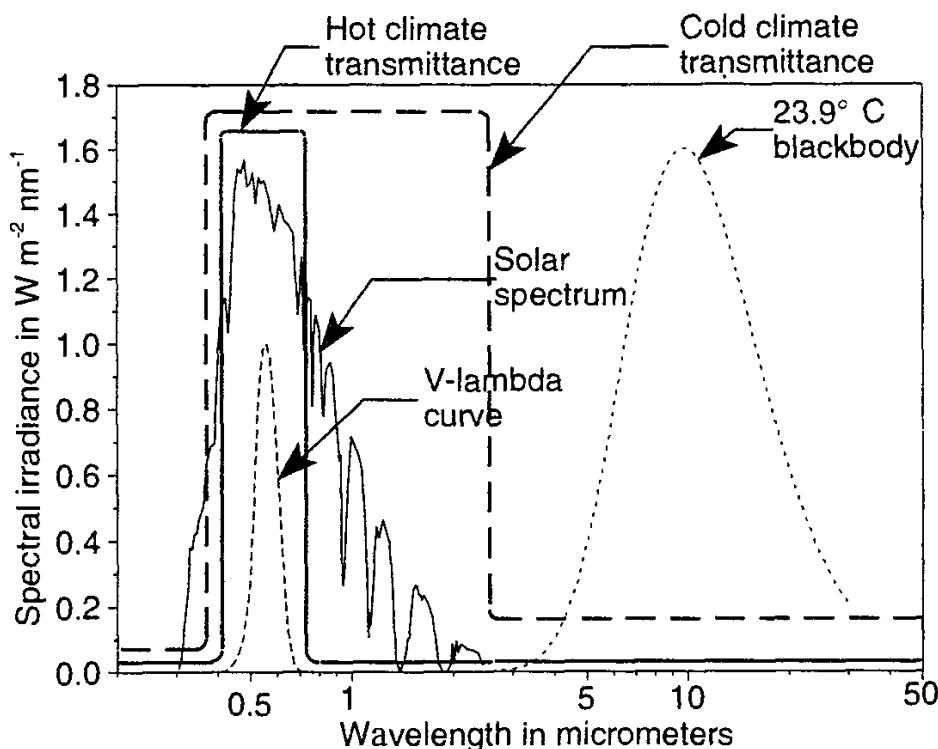


Figure 6.17 Ideal window transmittance spectra for cold and warm climates.

reduction [38]. The emittance referred to is the infrared emittance of the coated glass at wavelengths greater than around $3\text{--}4 \mu\text{m}$ where the absorptivity of glass is quite high. If the absorptivity of the glass is high, then its transmittance for reasonably thick panes will be low. According to (6.39) and (6.40), this means that the reflectance of the coated pane of glass is quite high, approaching 90%. If this were true at all wavelengths, the glass would not be very appropriate for use in a window, since little light would transmit through it.

Fortunately, the coatings for this application are made to be strongly spectrally selective, so that the high reflectance over the long-wavelength infrared switches to low reflectance (and high transmittance) over the solar spectrum (which includes the visible). (See Figures 6.17 and 6.18.) The result is, in effect, a “heat mirror” combined with a “visible and solar window,” an accentuation of the greenhouse effect. The long-wavelength radiant heat emitted from the warm interior pane towards the cold outer pane on a cold winter night is reflected back to the warm pane by the low-e coating, trapping this valuable heat inside the building. Interestingly enough, the low-e coating works just about as well on the outward-facing surface of the interior pane of a two-pane window. The low-e coating ensures that nearly ten times less flux is radiated from the warm inner pane to the cold outer pane than would be the case for an uncoated pane. The result is better window insulation without sacrificing the visible and solar transmittance. Modern multiple-pane insulated windows of this sort can make up in solar heat gain on cold winter

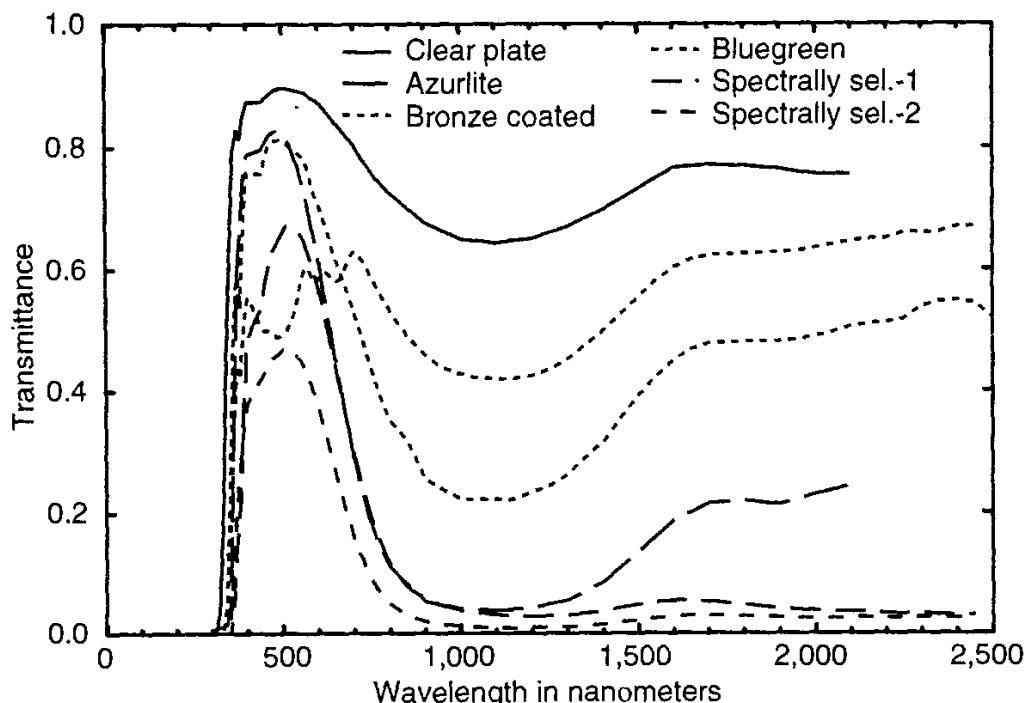


Figure 6.18 Spectral transmittances of several commercially available window glazing systems.

days what they lose by conduction on cold winter nights, thereby out-performing insulated walls in terms of net energy performance.

An ideal window intended for a *hot* climate would be transparent to only the *visible* portion of the solar spectrum and reflective to the remaining portions of it, like the solid curve in Fig. 6.17, thereby reducing the radiant heat gain into the building through the window while permitting visible daylight inside for illumination and for a good view of the outdoors.

The conical-hemispherical transmittance of a window system can be used to determine the fraction of direct beam solar irradiance incident on the window which is transmitted in all directions inside.

The conical-hemispherical spectral transmittances at normal incidence of several different coated and uncoated glazing systems that can be used to partially accomplish the desired spectral selective objectives are shown in Figure 6.18. The “spectrally selective-1” curve is for a two-pane window with good visible transmittance and strong absorption in the outer glazing (over the near infrared portion of the solar spectrum). This high absorptance prevents the heat-producing invisible IR radiation from entering the building directly. By virtue of its high absorptance, however, this outer glazing gets fairly hot. If this were the only glazing in the window, a sizable portion of this absorbed heat would enter the building (by conduction, convection, and radiation), defeating the purpose of the spectral selectivity. With the addition of the second pane, however, the trapped air space along with the second pane of glass effectively insulates the building interior from the absorbed and reemitted heat of the first pane. Blockage of this inward-flowing fraction of the absorbed solar radiation is further enhanced in this glazing system by the use of a low-e coating on the inward side of the outer glazing. Because of this coating, with an emittance below 0.15, a small fraction of the heat that would be radiated by bare window glass, with its emittance of about 0.84, is radiated to the second pane. The result is a window glazing system that exhibits good illumination and energy performance in warm and hot climates.

Similar spectral selectivity is used in the projection of photographic slide images and motion pictures. The emission spectrum of a representative incandescent lamp is shown in Figure 6.19. The shape of this spectrum is similar to that of a blackbody at 3,000K. Very little of the radiation from such a source lies in the visible portion of the spectrum. There is substantial absorption of the incident radiation by the film in a projector. Most of this radiation is infrared and is not used in projecting the visible image from film to screen. If the heat from this absorbed radiation is excessive, it can damage the film. On the other hand, for good luminance of the projected image, it is important to pass as much visible light through the film as possible. So-called “hot mirrors” are spectrally selective plates made to reflect the infrared portion of the source spectrum while transmitting the visible portion without altering the color of the projected image. The transmittance spectrum of a representative hot mirror is shown in Figure 6.20.

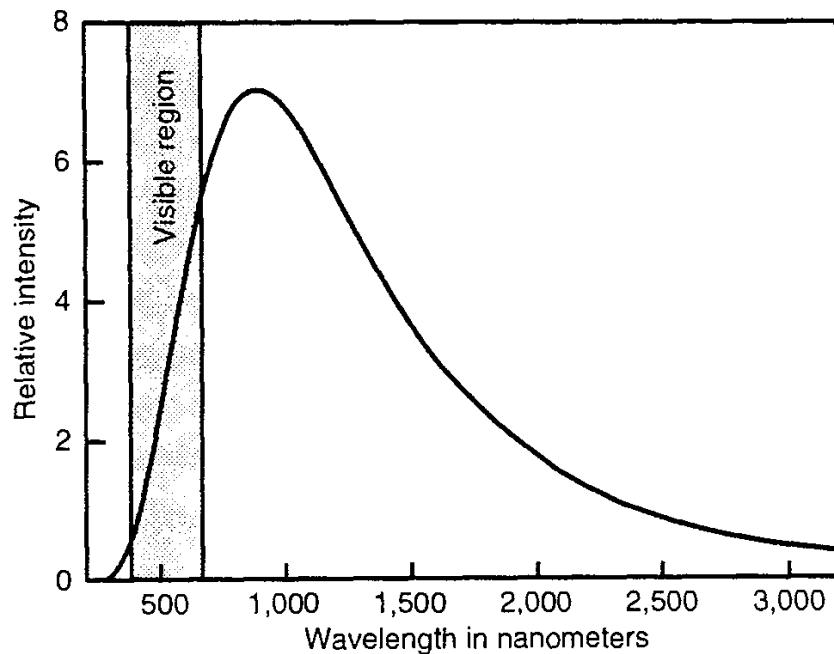


Figure 6.19 Tungsten filament incandescent lamp spectral output distribution.

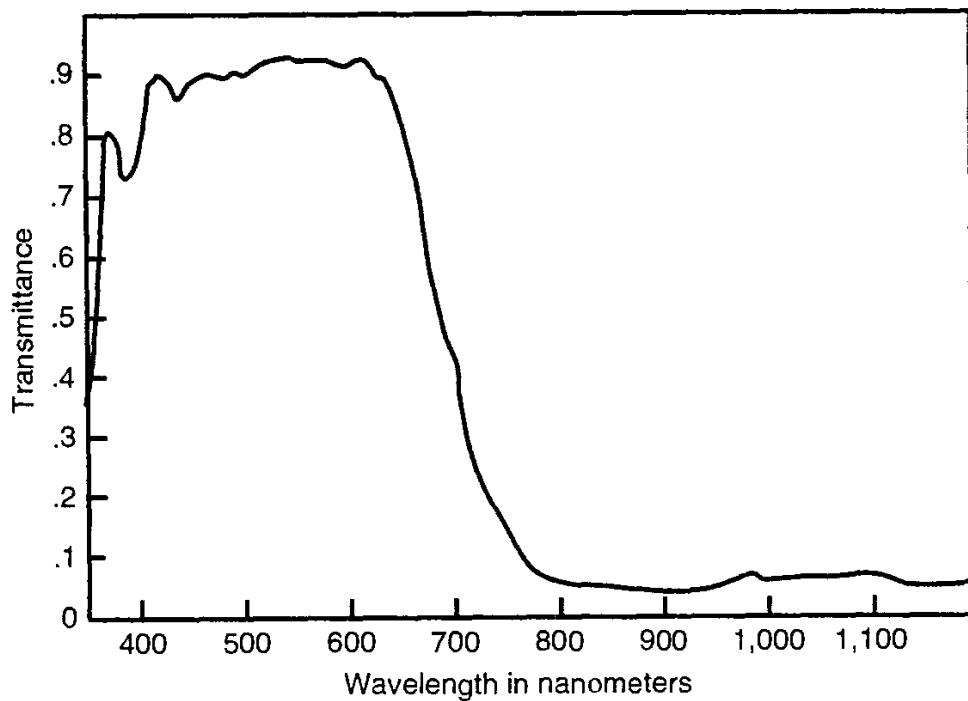


Figure 6.20 Typical commercially available hot mirror transmittance spectrum.

Another example of the use of spectral selectivity is in the design of solar collectors. Most solar thermal collectors involve a black absorbing surface with an air space and a clear glass or plastic sheet to serve as a window or cover plate. The incident solar radiation travels through this cover plate and is absorbed by the black plate, from which the heat is carried away for use by a fluid traveling through pipes or channels in the black plate itself.

There are two opportunities in such a system for the use of spectrally selective optical properties. First, with the cover plate, the spectral transmittance curve for cold climates illustrated in Figure 6.17 can be used. Such a cover glass would transmit the incident solar spectrum with as high a transmittance as possible, but would block the long-wavelength infrared radiation emitted by the heated black absorber plate. The reflectance of the cover plate would be low over the solar spectrum (and the emittance would be high) but the reflectance would be high over the thermal IR spectrum (and the emittance would be low, making this “low-e glass”). This arrangement would trap heat inside the space behind the glazing.

The second opportunity for spectral selectivity is in the control of the shape of the absorber plate’s emittance spectrum $\epsilon(\lambda)$. This emittance can be high (making the absorptance high) over the solar spectrum and low over the emitted IR spectrum. This arrangement enhances the absorbed solar radiation while trapping the resulting heat in the black plate.

EXAMPLE PROBLEM 6.1

Problem: The signal from one *pixel* (picture element) in an imaging infrared radiometer, corresponding to a point on the insulated surface of a large industrial boiler, a metal tank to which heat is applied for the purpose of making steam, indicates a received spectral radiance at a wavelength of $8.23 \mu\text{m}$ of $153 \text{ W}\cdot\text{m}^{-2}\cdot\text{sr}^{-1}\cdot\mu\text{m}^{-1}$. The normal emittance at this wavelength for the insulation surface coating is known to be $\epsilon(\lambda, 0, 0) = 0.756$. Is there a problem with the boiler?

Solution: Solving (6.5) and (3.5) for the spectral radiance of a nonblackbody for the temperature T of the outer surface of the insulation yields

$$T = \frac{C_2}{\lambda} \left[\ln \left(1 + \frac{\epsilon(\lambda, 0, 0) C_{1L}}{\lambda^5 L_{nb\lambda}} \right) \right]^{-1} \quad (6.59)$$

Substituting values for C_{1L} , C_2 , λ , $\epsilon(\lambda, 0, 0)$, and $L_{nb\lambda}$ yields the result that $T = 635\text{K}$ or 362°C , quite a bit hotter than the boiling point of water. We conclude from this that there is indeed a problem. If the insulation is this hot, then either the boiler is very overheated, possibly on the verge of exploding, or the insulation

around the boiler at this point is very weak, allowing a lot of heat to escape, or both. Imaging infrared radiometers are commonly used to spot overheated surfaces or faulty insulation remotely, without need for direct contact temperature measurements.

EXAMPLE PROBLEM 6.2

Problem: A helium-neon laser is used as a light source to measure the transmittance T (at normal incidence) of a 1 meter long rod of clear plastic, whose refractive index at the 632.8 nm laser wavelength is known to be 1.4326. Although there is some scattering of light from the laser beam by the plastic, the magnitude is observed to be very small, at all angles of scattering. What is the spectral absorption coefficient $a(\lambda)$ at this wavelength if the measured spectral transmittance is 0.734?

Solution: Assuming air as the medium surrounding the rod, the reflectance $\bar{\rho}$ of the air/plastic interface, given by (6.10), is 0.03163. Replacing the quantity $\bar{\tau}_1\bar{\tau}_2$ with $(1 - \bar{\rho})^2$ in (6.47), and substituting the above values for $\bar{\rho}$ and T into the resulting equation yields the quadratic equation

$$\begin{aligned} \bar{\rho}^2 T \alpha^2 + (1 - \bar{\rho})^2 \alpha - T &= 0 \\ 0.00073 \alpha^2 + 0.937 \alpha - 0.734 &= 0 \end{aligned} \quad (6.60)$$

Recall that α is the internal transmittance of the medium due to absorption over length x in that medium. Substitution of the relevant coefficients of α in (6.60) into the well-known quadratic formula gives that our solution of (6.60) for α is either 0.7615 or -1278 . Negative transmittance makes no sense here, so we have $\alpha = 0.7615$. Solving (6.42) for a in terms of α and the length x of the plastic rod gives

$$a = -\frac{\ln(\alpha)}{x} \quad (6.61)$$

for the absorption coefficient of the rod. Since $x = 1\text{m}$ in our experiment, we can substitute our value for α into this equation to obtain the solution, $a = 2.272\text{m}^{-1}$.

REFERENCES

- [1] *CIE International Lighting Vocabulary*, CIE Publ No. 17.4 IEC Publ. 50(845), Central Bureau of the *Commission Internationale de l'Eclairage*, Kegelgasse 27, A-1030 Vienna, P. O. Box 169, Austria, and Bureau Central de la Commission Eletrotechnique Internationale, 2, rue de Varembe, Geneva, Switzerland, 1987. Also available from the national committees of the CIE. In the U.S.,

-
- this document can be obtained from TLA-Lighting Consultants, Inc., 7 Pond Street, Salem, MA 01970-4893.
- [2] Siegel, R., and J. R. Howell. *Thermal Radiation Heat Transfer*, 2nd ed., Hemisphere Publishing/McGraw-Hill Book Co.: New York, NY, 1981.
 - [3] Born, M., and E. Wolf. *Principles of Optics*, 3rd (revised) ed., Pergamon Press: New York, NY, 1965, p. 615.
 - [4] *Ibid.*, pp. 38–42.
 - [5] ASHRAE. *1993 ASHRAE Handbook Fundamentals*, Amer. Soc. Heat., Refrig., and Air-Cond. Eng., Atlanta, 1993, p. 3.8.
 - [6] ASTM, "Standard Test Method for Emittance of Specular Surfaces Using Spectrometric Methods," *Standard E-1585*, Amer. Soc. Test. Materials, 1916 Race St., Philadelphia, PA, 19103, 1994.
 - [7] Weast, R. C., ed. *CRC Handbook of Chemistry and Physics*, CRC Press: Boca Raton, FL, 1985.
 - [8] Lieberman, R. A., ed., "Chemical, Biochemical, and Environmental Fiber Sensors III," *Proc. of the SPIE*, Sept. 4–5 1991, Boston, MA, Vol. 1587, 1992.
 - [9] Udd, E., and R. P. DePaula, eds. "Fiber Optical and Laser Sensors VII," *Proc. of the SPIE*, Sept. 5–7 1989, Boston, MA, Vol. 1169, 1990.
 - [10] Lieberman, R. A., and M. T. Wlodarczyk, eds. "Chemical, Biochemical, and Environmental Fiber Sensors," *Proc. of the SPIE*, Sept. 6–7 1989, Boston, MA, Vol. 1172, 1990.
 - [11] Shurcliff, W. A. *Polarized Light-Production and Use*, Harvard University Press: Cambridge, MA, 1966, p. 80.
 - [12] Bennett, J. M., and L. Mattsson. *Introduction to Surface Roughness and Scattering*, Optical Society of America: Washington, D. C., 1989.
 - [13] McCluney, W. R. "The Radiometry of Water Turbidity Measurements," *J. Water Pollution Control Federation*, Vol. 47, Feb. 2 1975, pp. 252–266.
 - [14] *Standard Methods for the Examination of Water and Wastewater*, 12 ed., Amer. Pub. Health Assoc.: New York, NY, 1965.
 - [15] Hardenbergh, W. A. *Water Supply and Purification*, International Textbook Co.: Scranton, PA, 1938.
 - [16] Fischer, R. B. "Turbidity and Nephelometry," in *Standard Methods of Chemical Analysis*, Vol. III, F. J. Welcher, ed., D. Van Nostrand: Princeton, NJ, 1966.
 - [17] Eichner, D. "The Absolute Turbidity of Pure Water," *Tech. Info. Booklet No. 2*, Hach Chemical Co., Ames, IA, 1971.
 - [18] Pickard, G. L., and L. F. Giovando. "Some Observations of Turbidity in British Columbia Inlets," *Limn. & Ocean.*, Vol. 5, 1960, p. 162.
 - [19] *Manual on Industrial Water and Industrial Waste Water*, 2nd ed., Amer. Soc. for Testing and Materials: Philadelphia, PA, 1965.
 - [20] Jerlov, N. G. *Marine Optics*, Elsevier Scientific Publishing: New York, NY, 1976.
 - [21] Jerlov, N. G., and E. S. Nielsen. *Optical Aspects of Oceanography*, Academic Press: New York, NY, 1974.
 - [22] McCartney, E. J. *Optics of the Atmosphere-Scattering by Molecules and Particles*, John Wiley & Sons: New York, NY, 1976.
 - [23] Liou, K. N., *Radiation and Cloud Processes in the Atmosphere*, Oxford University Press: New York, NY, 1992.
 - [24] Kondratyev, K. Ya. *Radiation in the Atmosphere*, (International Geophysics Series, Vol. 12), Academic Press: New York, NY, 1969.
 - [25] McCluney, R. "Color-Rendering of Daylight from Water-Filled Light Pipes," *Solar Energy Materials*, Vol. 21, 1990, p. 191.
 - [26] Morel, A. "Optical Properties of Pure Water and Pure Sea Water," in *Optical Aspects of Oceanography*, ed. by N. G. Jerlov and E. S. Nielson, Academic Press: New York, NY, 1974.

-
- [27] Logan, N. A. "Survey of Some Early Studies of the Scattering of Plane Waves by a Sphere," *Proc. IEEE*, Vol. 53, 1965, pp. 773–785. Reprinted in Kerker, M., ed., "Selected Papers on Light Scattering," *SPIE Milestone Series*, Vol. 951, Part One, SPIE, Bellingham, WA, 1988.
 - [28] Duffie, J. A., and W. Beckman. *Solar Energy Thermal Processes*, John Wiley & Sons: New York, NY, 1974, p. 89.
 - [29] Grum, F., and R. J. Becherer. *Optical Radiation Measurements Vol. 1 Radiometry*, Academic Press: New York, NY, 1979, p. 114.
 - [30] Halliday, D. and R. Resnick, *Physics*, John Wiley & Sons, New York, 1966, p. 1089.
 - [31] Preisendorfer, R. W. *Radiative Transfer on Discrete Spaces*, Pergamon Press: New York, NY, 1965.
 - [32] Chandrasekhar, S. *Radiative Transfer*, Dover Publications: New York, NY, 1960.
 - [33] Plass, G. N., and G. W. Kattawar. "Radiative Transfer in an Atmospheric-Ocean System," *Appl. Opt.*, Vol. 8, 1969, p. 455.
 - [34] Nicodemus, F. E., et al. "Geometrical Considerations and Nomenclature for Reflectance," NBS Monograph 160, NBS (Now NIST), Oct. 1977, p. 52.
 - [35] Parker, S. P., ed. *McGraw-Hill Concise Encyclopedia of Science & Technology*, McGraw-Hill Book Co.: New York, NY, 1982.
 - [36] ASHRAE, *1993 ASHRAE Handbook-Fundamentals*, Amer. Soc. Heat. Refrig., & Air Cond. Eng., Atlanta, 30329, Ch. 27.

Chapter 7

The Detection of Radiation

7.1 INTRODUCTION

There is considerable variety in the kinds of devices available for the detection and measurement of optical radiation. Some respond to the heat produced when radiant energy is absorbed by a surface. Some convert this heat into mechanical movement, and some convert it into electricity. Photographic emulsions convert the incident light into chemical changes made visible by the development process. Other detection devices (known as *detectors* or *sensors*) convert the electromagnetic radiant energy directly into electrical energy. Many electrical circuits have been devised to amplify the typically small electrical signals produced by detectors to levels that are easier to measure. There are unavoidable small fluctuations combined with the output signals of all detectors that mask or obscure the signal that results from the incident radiation. This is called noise, and is discussed briefly in this chapter.

It is not the purpose here to describe radiation detectors in great detail. Several excellent references are available that cover the subject well [1–3]. The goal is to show how detectors work, to describe their responses generally, to compare the relative characteristics of different devices, and to describe some optical systems used in conjunction with detectors. The use of detectors in radiometers and photometers is discussed in Chapter 9.

Most sensitive detectors of optical radiation have nonuniform spectral responses. Frequently, the inherent spectral response is not the one desired for an application. In many such cases, it is possible to add a spectrally selective filter to a detector to produce a spectral response of the combination that is closer to what is really needed. Due to the variety of detectors and filters available, this can be a challenging task. However, matching filters with detectors to produce a desired result is one of the most important problems in the field of radiometry and photometry. Accordingly, some effort has been put into discussions of this problem near the end of this chapter, as well as in Chapters 8 and 9.

7.2 BASIC CONCEPTS

In describing the response of detectors to incident radiation, we are interested in the ratio of the output signal to the input flux. For good detection devices, this ratio, called the *responsivity* or *sensitivity* of the detector, is constant over a wide range of incident flux levels. Such detectors are said to have good linearity, meaning they have a linear response to the incident flux. In such cases, we can use the following relationship between the input radiation Φ and output signal S :

$$S = R\Phi \quad (7.1)$$

where R is the calibration constant, or *responsivity*, of the detector. If the device outputs a voltage proportional to the incident flux and the flux is measured in watts, then R will have units of volts/watt. If it outputs current, then R will have units of amperes/watt. Some detectors output a pulse of current characterized by the total charge delivered. In such cases, R might have units of coulombs/joule.

The word *signal* is reserved for the portion of the output response of the detector that is directly attributable (that varies monotonically) with respect to the incident radiant flux on the detector. Thus, it is incorrect to speak of a “noise signal” coming from a detector. Noise is not the signal desired, but a contaminant of it.

Many detectors have a nonzero output when the incident flux is zero. This output is called the *dark output*. For current output devices such as multiplier phototubes and silicon photovoltaic detectors connected to current-measuring circuits, the dark output is called *dark current*. The dark output (which can either match or oppose the signal polarity) is here given the symbol S_o and (7.1) is rewritten as

$$S - S_o = R\Phi \quad (7.2)$$

or

$$S = R\Phi + S_o \quad (7.3)$$

Note that care was taken not to call the dark output a “dark signal.” The dark output is not part of the desired signal but a spurious output that is by definition not related to the true signal response to incident radiation.

Although the ideal is to have a signal output that is strictly linearly related to the incident flux on the detector, most detectors exhibit some response nonlinearity depending upon parameters such as temperature, wavelength of incident radiation, and the magnitude of the incident flux. Accordingly, one should more

properly speak of the response *function* of a detector, a mathematical expression (or table of values) that relates the output signal to the input flux and all the other parameters upon which the signal depends.

Let the response function of a detector be designated by the symbol $R(\Phi, x, y, z, \dots)$, shown here as dependent upon the incident flux as well as some additional variables x, y, z , and so forth, and having units of volts, amperes, coulombs, or other unit of output signal, per watt (or joule) of incident radiant flux (or energy) for radiometric detectors and per lumen of luminous flux for photometric ones. The detection equation is

$$S = R(\Phi, x, y, z, \dots) \Phi + S_o \quad (7.4)$$

Detectors generally have variable responses over their sensitive surfaces, some more so than others, and this must be considered in the design of complete radiation sensors. (One method used to overcome this variability is to precede the detector with a diffusing plate, integrating sphere, or other similar optical component intended to spread a fixed fraction of the incident flux over the whole sensitive area of the detector.)

In some cases, such as when the incident flux varies over a very wide range and expensive and/or slow automatic gain ranging circuits are not wanted, a non-linear response may be desired. For example, the global horizontal visible radiation from the sky during the twice-daily twilight transitions between day and night varies over some seven to eleven orders of magnitude, depending upon how one defines twilight. For automated measurements of the illuminance from the sky during these periods, it might be useful if the detector responded logarithmically to the incident flux, thereby not requiring gain changing mechanisms and other complex circuits.

In all cases, care must be taken to determine any nonlinearity in the response of a detector over the intended range of its use. An alternative for accommodating a wide range of flux levels with a detector having a limited response range could be the use of neutral density filters (filters with flat spectral transmittance over the spectral range of the measurement) or some other means to attenuate by precisely known amounts high levels of flux incident upon the detector so that the flux actually reaching it varies over a restricted range, the range where the detector is known to have linear response.

Most detectors also have variable responses over the angles of incidence of the radiation. Letting θ be the angle of incidence, the functional dependence on this angle can be shown explicitly in the response function as $R(\Phi, \theta, x, y, z, \dots)$. Section 1.5 in Chapter 1 showed that the irradiance received by a flat surface in a collimated incident beam decreases with the cosine of the angle of incidence. For this reason, the flux received by the sensitive surface of an ideal detector placed

in such a beam should also show a $\cos(\theta)$ dependence. Departures from this dependence can be indicated in the detection equation as follows:

$$S = R(\Phi, \cos(\theta), x, y, z, \dots) \Phi + S_o \quad (7.5)$$

The wavelength dependency of the response function is of special interest. This response is shown generally as

$$S(\lambda) = R(\lambda, \Phi_\lambda, \cos(\theta), x, y, z, \dots) \Phi_\lambda + S_o \quad (7.6)$$

For the moment, consider the function R to depend only on the wavelength of the incident radiation. It is therefore constant over the detector's sensitive surface, independent of incident spectral flux levels (the detector has linear response over a wide range of flux levels), and independent of all other variables such as pressure or temperature. For now, consider that only normally incident radiation at $\theta = 0$ is involved. In this case, R is a function of wavelength only and is written as $R(\lambda)$. If $E_\lambda(\lambda)$ is the spectral irradiance incident upon the sensitive area A of such a detector, then the output signal S will be given by

$$S = A \int_0^\infty E_\lambda(\lambda) R(\lambda) d\lambda + S_o \quad (7.7)$$

If either the incident spectral flux $E_\lambda(\lambda)$ or the detector spectral response $R(\lambda)$ is zero outside some spectral range, then the limits of the integral in (7.7) can be restricted to this range. To evaluate the integral in (7.7), one needs to know the spectral distributions of both the detector's response $R(\lambda)$ and the incident irradiance $E_\lambda(\lambda)$. If there is a window or an optical system interposed between the source of the radiation and the detector, then one needs to know also the spectral transmittance $T(\lambda)$ of this system and to include it in the integrand of (7.7). Thus,

$$S = A \int_0^\infty E_\lambda(\lambda) T(\lambda) R(\lambda) d\lambda + S_o \quad (7.8)$$

This measurement situation is diagrammed schematically in Figure 7.1.

There are four special cases of interest. First is the case where the detector response is "flat" over the spectral range $[\lambda_1, \lambda_2]$ of flux incident on it; that is, the detector spectral response function $R(\lambda)$ is constant, $R(\lambda) = R_o$ from λ_1 to λ_2 and 0 elsewhere, as illustrated in Figure 7.2. In this case, (7.8), with $T(\lambda) = 1$, becomes

$$S = A R_o \int_{\lambda_1}^{\lambda_2} E_\lambda d\lambda + S_o = A R_o E_e + S_o \quad (7.9)$$

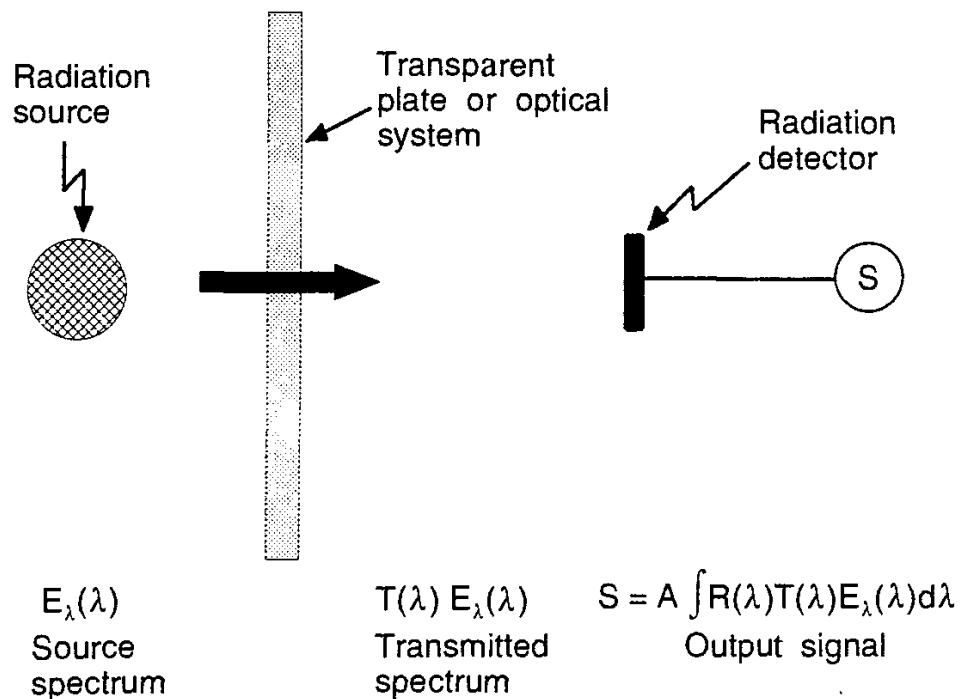


Figure 7.1 Determining the signal from a detector.

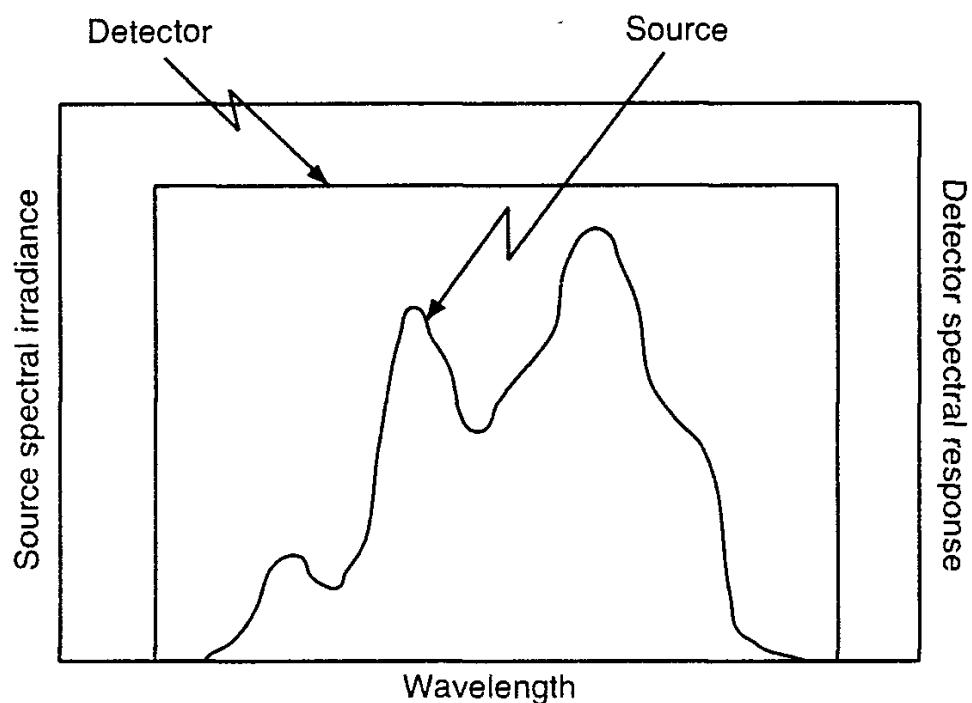


Figure 7.2 Detectors with flat spectral responses are easy to calibrate.

and the equation can be solved for the incident irradiance E_e :

$$E_e = \frac{1}{AR_o} (S - S_o) = K(S - S_o) \quad (7.10)$$

where $K = 1/(AR_o)$ is the calibration constant for using this detector as a normal incidence irradiance meter. Multiplying the net signal $S - S_o$ by K gives the incident irradiance on the detector, but only over the wavelength range of constant, nonzero detector spectral response.

Some instrument manufacturers offer radiometers with nearly constant spectral response over $[\lambda_1, \lambda_2]$ and essentially zero response outside this range. The output of these meters is valid only for the portion of the incident radiation within this range. Any radiation outside this range will not be measured by such instruments.

Note that if there is a significant quantity of radiant flux outside the range over which the detector response is constant, the use of (7.10) to measure irradiance will introduce an error in the measurement. This is illustrated in Figure 7.3. The shaded portion of the incident spectral irradiance is not measured.

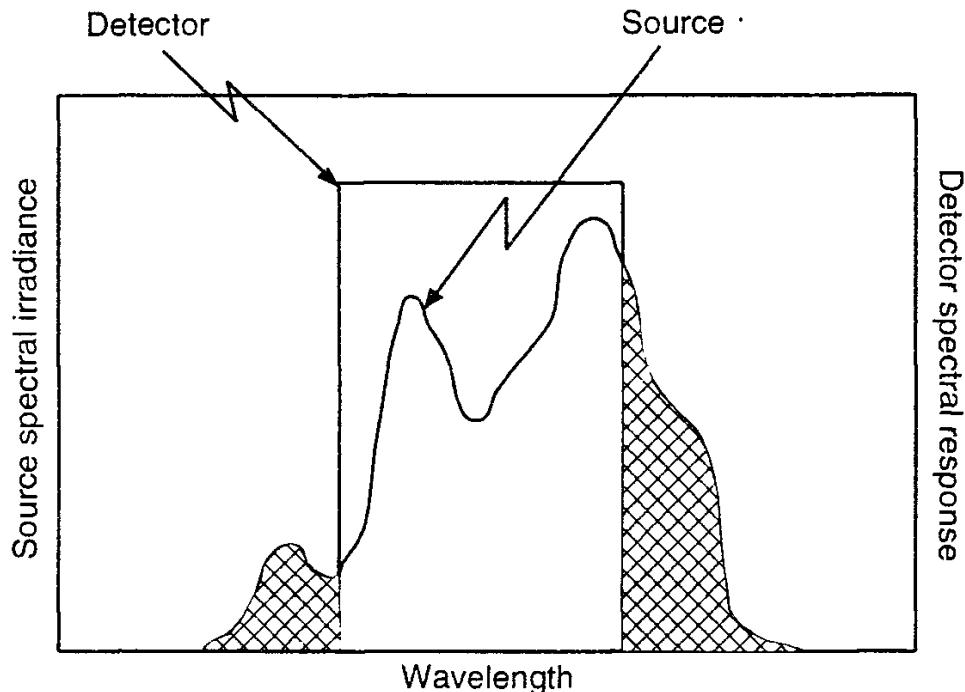


Figure 7.3 A detector with a limited spectral range can only sample a portion of a source's spectrum.

The second special case is that for which the detector response is not flat, not constant, but rather the incident spectral irradiance *is* constant over the spectral range $[\lambda_1, \lambda_2]$ of interest, as illustrated in Figure 7.4.

In this case, we can write

$$S = AE_e \int_{\lambda_1}^{\lambda_2} R_\lambda d\lambda + S_o = AE_e R + S_o \quad (7.11)$$

and solve for

$$E_e = \frac{S - S_o}{AR} \quad (7.12)$$

Few sources of radiant flux have constant outputs over significant wavelength regions. Thus, (7.12) is included more for the sake of completeness than for its practical value.

The third case occurs if the detector response is made to be nonzero only over a very narrow spectral range, one over which the incident spectral flux can be said to be nearly constant, as illustrated in Figure 7.5. This is the case, for example, when the detector is preceded by a narrow bandpass spectral filter or a

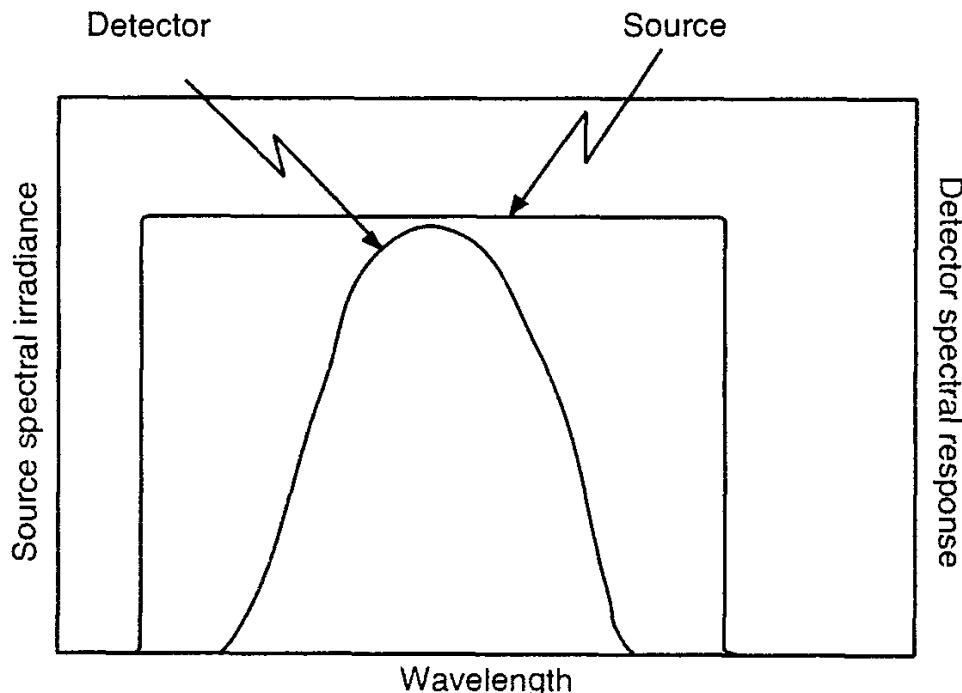


Figure 7.4 Flat source spectrum with non-constant detector response.

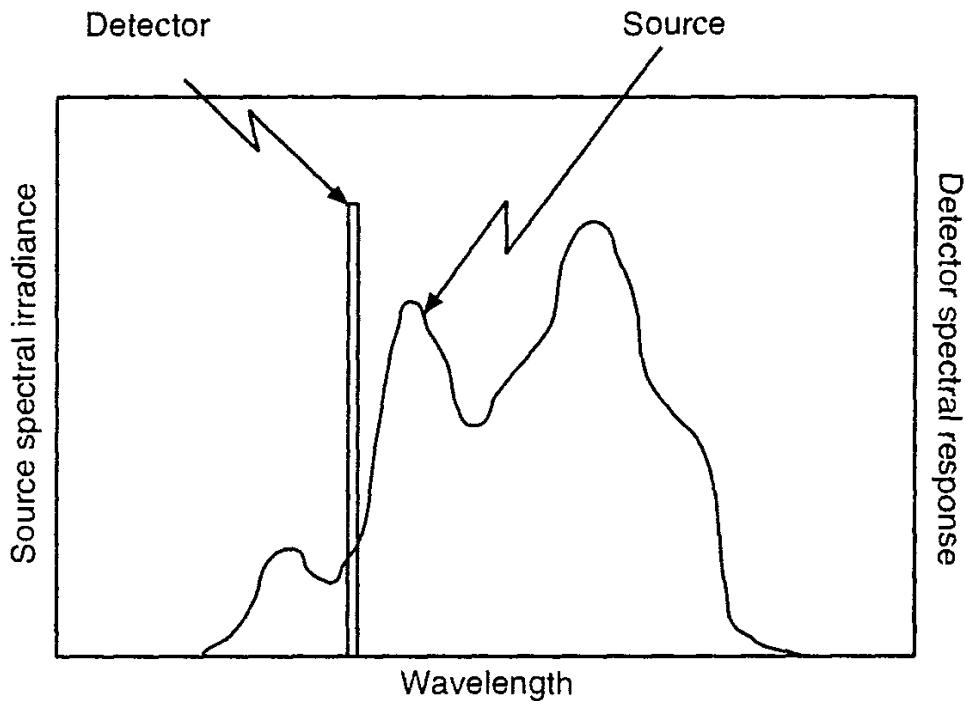


Figure 7.5 A detector with a narrow spectral band response samples only a tiny portion of a source spectral distribution.

monochromator. The wavelength range is so narrow that the source spectrum can be said to be (or is treated as if it were) constant within that range. This case is discussed more fully in Sections 8.12 and 9.6.

The fourth case occurs when the spectral distribution of flux falling on a detector is nonuniform in shape but the shape never varies. In this case, a non-uniform detector spectral response will output a signal that is a spectrally weighted integral or average of the flux from the source. It is possible to calibrate such detectors to determine from their output signals the correct integrated total radiant flux incident on them even with varying source flux magnitudes, provided that the relative spectral distribution of the incident flux remains constant. If this shape is not constant, then calibration errors will result from the shape variations and a detector with flat (constant spectral) response over the spectral range of interest should be used instead.

Another parameter that generally affects detector responsivity is temperature. With some devices this is a wavelength-dependent effect. In the case of a silicon photodiode, to be described in Section 7.3.3, the change in responsivity with temperature is modest from 500 to 900 nm, increases below 500 nm to about $-0.4\%/\text{C}$ at 300 nm and increases rapidly above 900 nm to over $1.5\%/\text{C}$ above 990 nm.

Budde provides an extensive discussion of the characteristics of many types of optical radiation detectors [1].

7.3 CLASSIFICATION OF DETECTORS

Detectors of electromagnetic radiation can be classified according to the physical mechanisms or principles involved in the conversion of incident radiant flux into a measurable signal.

7.3.1 Thermal Detectors

Thermal detectors deliver a signal proportional to the increase in temperature resulting from the absorption of radiation incident upon them. To increase the absorption, they are generally made of materials or are given coatings with high absorptance over the spectral region of use. Because of their essentially thermal nature, the coatings or materials can have relatively constant spectral absorptance over a significant wavelength range, resulting in flat (nonvarying) spectral response detectors, detectors with spectral response nearly constant over a wide range of wavelengths. This flatness is generally desirable because it makes the calibration of these devices much easier. On the other hand, thermal detectors are characterized by relatively low sensitivities when compared with detectors operating on other principles. They therefore work best when measuring high flux levels.

Thermal detectors are either *metal junction* devices or *bulk* devices. The junction devices, such as the thermocouple and thermopile, rely upon the Seebeck (or thermoelectric) effect [4], whereby two separate junctions of two dissimilar metals at different temperatures generate a voltage proportional to the difference in temperature between them. If one junction is kept at some known reference temperature, then the series output voltage will be proportional to the temperature of the other junction, which is made to absorb some of the electromagnetic radiation incident upon it. By careful design of the device, the increase in temperature will be proportional to the incident radiant flux. The active junction is made to have low mass so that it will respond as quickly as possible to changes in the incident flux level. In spite of this, these devices are characterized by relatively slow response to changing radiation levels.

The reverse of this effect, called the Peltier effect, is of some interest. In this case, a voltage is applied to the circuit containing the two junctions of two dissimilar metals. The result is an increase in the temperature of one junction and a decrease in the temperature of the other one. It is not a particularly efficient process for producing either heating or cooling. However, when efficiency is not that important and when the quantity of heat generation or removal is not great, this process can be used to provide or remove heat. The Peltier effect can be used, for example, to cool detectors that operate better at low temperatures, without the need for complicated, bulky, or expensive refrigeration apparatus.

The Peltier effect is also important in the design and characterization of

thermocouple detectors. The reason is that the current generated by the incident radiation (Seebeck effect), flowing through the thermocouples, can induce a small cooling of the irradiated junction through the Peltier effect. This cooling introduces a systematic error into the measurement by lowering the hot junction temperature and reducing responsivity. This effect can be eliminated by minimizing the current flow through the thermocouple, by increasing the resistance of the external circuit. The cost of this strategy is increased noise. In practice some compromise must be found that provides an adequate signal-to-noise ratio (*SNR*) without substantial responsivity changes due to Peltier cooling. Budde discusses this in more detail [5].

Individual thermocouples are effective as temperature sensors—electrical thermometers that produce a small electrical signal proportional to their temperatures. They are not very practical as detectors of electromagnetic radiation, however, because of their low sensitivities. Placing a number of thermocouples in series, however, can substantially increase the sensitivity. This configuration is called a *thermopile*. Two mechanical configurations of thermopiles are common. One is composed of a chain or series of individual thermocouples connected by wires. The other is made by the vacuum deposition of a thin film onto a suitably chosen substrate. The latter typically uses evaporated bismuth and antimony. The active junction is made large (compared to the diameter of the wires connected to it), coated with a black material with little spectral selectivity, and sealed inside a protective container with an insulating transparent window made of glass or other material with high (and constant) spectral transmittance over the useful range of the device.

One set of thermocouples (called the active set) is in close thermal contact with the black coating receiving the radiation and the other (called the reference set) is shielded from this radiation. Because of the way the reference thermocouples are thermally isolated, performance is not greatly affected by modest variations in the ambient temperature. When extra care is given to prevent the reference junction(s) from responding to changes in ambient temperature, the thermocouple or thermopile detectors are said to be *compensated*. Due to its relatively higher mass, the thermopile typically has an even slower response than an individual thermocouple. A schematic diagram of a chain of thermocouples connected by wires is shown in Figure 7.6. This diagram is intended to be illustrative only. The top thermocouples are in thermal contact with the black absorbing plate, shown shaded in the diagram. The bottom (cold) thermocouples are thermally insulated from the top ones.

Thin film thermopiles have been developed in recent years to increase sensitivity and decrease response times. They are typically made of a strip of blackened gold foil welded at each end to a different semiconductor material, chosen to have different thermoelectric potentials relative to the gold foil. The semiconductors are mounted on supporting metal pins, which are in turn mounted in an electrically

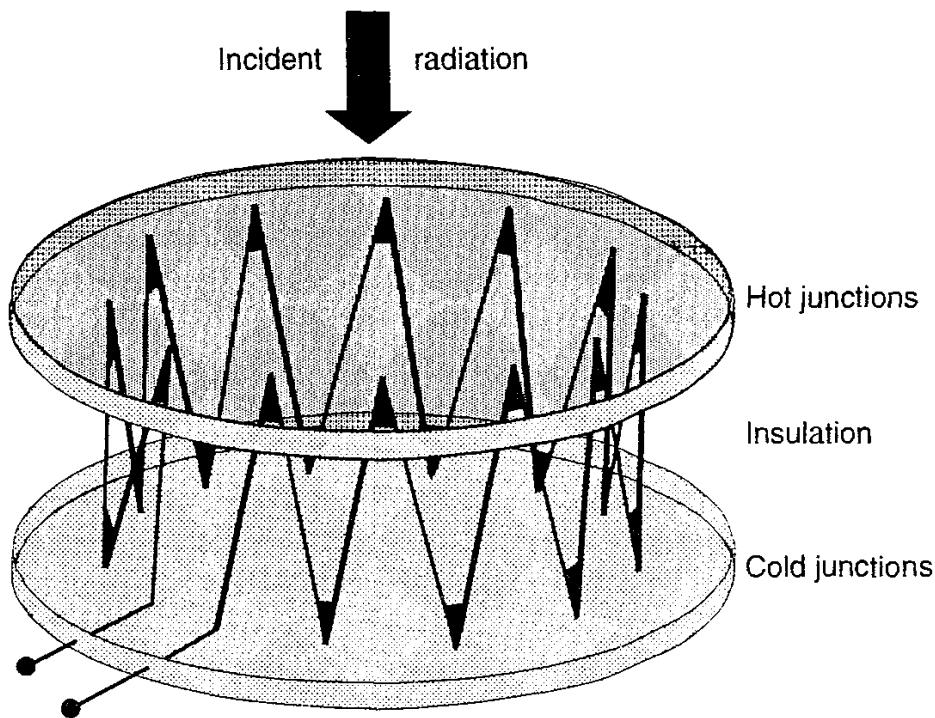


Figure 7.6 Schematic diagram of thermopile connections.

insulating but thermally conducting block of relatively large thermal mass. These thermopiles are available in evacuated, gas-filled, or air-filled enclosures.

The *pyranometer* is a special type of thermopile intended for the measurement of solar radiation. Pyranometers measure the *global* solar radiation (direct beam plus diffuse sky). These detectors are given coatings with flat response over the solar spectrum from around 350 nm to beyond 3.5 μm . They are mounted in fairly massive and durable housings with a shade ring surrounding the entrance window. The shade ring protects the housing and the reference junctions inside from excessive heating from the solar radiation. Responsivities are typically 8 to 11 $\mu\text{V} \cdot \text{W}^{-1} \cdot \text{m}^{-2}$. Typical responsivities of other detectors are given in Section 7.10. *Pyrheliometers* are thermopiles specially constructed to measure the direct beam radiation from the sun.

Thermopile coatings are often made of Parsons Black (whose reflectance drops smoothly from 0.025 at about 300 nm to 0.015 at 700 nm and then increases to about 0.05 at 2.5 μm), Carbon black (with a reflectance dropping from around 0.018 over the range from 400 to 900 nm down to 0.008 at 2.5 μm), or gold black (with reflectances below 0.007 from 400 nm to 2.5 μm). Fused silica windows are made for these detectors with spectral transmittances that range from 0.90 at 200 nm to 0.93 at 3,000 nm, and are slowly varying in between. Windows are discussed in Section 8.13. Spectral transmittance curves for a variety of optical

materials can be found in the Oriel Volume III catalog for optics and filters [6] and in the catalogs of other optical component manufacturers.

Bulk thermal detectors utilize materials whose physical properties change with temperature. The changing property most used in the design of this class of detectors is resistance (or equivalently conductance), however other properties, such as a change in the electrical polarization (permanent internal charge separation) of the material, are also used. The Golay cell uses the thermal expansion of a gas in thermal contact with an absorbing surface to deflect a light beam, the deflection being proportional to the incident radiant flux.

One device in this category may seem at first to lie in the junction category. It is really a bulk device. This is the bimetallic strip. It consists of flat strips of two different metals with different coefficients of thermal expansion bonded together. Due to their different expansion coefficients, the metals lengthen by different amounts when heated. Since they are bonded together, the result of the differential expansion is a bending of the strip. For small deflections, the bending is proportional to the increase in temperature. If the bonded strip is shaped into a coil or spiral with many turns, then the result is a rotation of one end if the other is held fixed and the temperature changes. This same principle is used in many thermostats used to turn on and off building heating and cooling systems.

Molecular polarizability is utilized in another class of bulk detectors called *pyroelectric detectors*. Certain kinds of dielectric (electrically insulating) materials have a permanent electrical polarization (charge separation) within them, even in the absence of any applied electric field. The degree of this polarization is affected by the temperature of the material, and a change in the surface charge results from a change in temperature. Heating of the material results in an expansion of the crystal lattice and a change in the charge on the surfaces. Electrical contacts can carry some of this charge through an attached circuit. Since the charge cannot flow from these materials in one direction indefinitely, these devices can detect only *changes* in the incident radiant flux. A varying flux incident on the exposed surface produces a varying current signal in the measuring circuit whose amplitude is proportional to the amplitude of oscillating flux incident on the device.

Typical pyroelectric materials are triglycerine sulfide (TGS), strontium barium niobate, polyvinylidene fluoride, and lithium tantalate. All such detectors share one common limitation. If their temperature is elevated above a specific temperature, called the Curie temperature, they lose their electrical polarization. Fortunately, the polarization can be restored by the application of an external electric field. Introducing α -analine into TGS as an impurity produces a pyroelectric that returns automatically to its original polarized state after its temperature drops back below the Curie temperature, which is 49°C for this material. Due to their construction as electrical conductors on opposite sides of a dielectric, these materials behave electrically very much like capacitors.

There are two very important advantages to pyroelectric detectors. First is that, being thermal detectors, they can be given coatings with flat responses over very wide ranges. The second one results from the fact that because the signal is generated by a change in polarization at the molecular level, the response to incident radiation for small detectors with very small mass can be fast, with measurable signals being produced at radiation modulation frequencies as high as a gigahertz [2]. However, when coatings are added having a flat absorptance spectrum over a wide spectral range, the mass of the detector necessarily increases and the speed of response decreases. Normal pyroelectric detectors have a coating that is three to four times as thick as the pyroelectric material and these are seldom operated at high frequencies. Additionally, the temporal response is also influenced by the shunt resistance and capacitance of the detector/amplifier circuit. These various characteristics are adjusted to optimize pyroelectric detectors for operation as either detectors of pulsed radiation, the pulses being generally of short duration, or of relatively steady chopped radiation at a fixed frequency. In the latter case, chopping frequencies in the range from tens to hundreds of Hz are often used, but these can be extended up to 1 kHz for certain applications.

It is possible, through increasing the size and thermal mass of a pyroelectric and careful design of the detection circuit and other characteristics of the detector housing, to reduce drastically the frequency of peak sensitivity. In this case, the detector can be used without chopping, to sense slowly varying radiation, such as the long-wavelength infrared radiation emitted by the human body. In an example of this type of application, presented in Section 7.10, the frequency of peak sensitivity is below 0.1 Hz.

The disadvantages of pyroelectric detectors include a sensitivity to mechanical vibration (producing what is called microphonics) and lower sensitivity than some other classes of detectors.

7.3.2 Photemissive Detectors

Photoemissivity refers to the process generally known as the photoelectric effect. Photons incident upon a metal surface inside a transparent vacuum tube cause the ejection of electrons from the surface if the frequency—and hence energy, according to (1.37)—of the photons is high enough to overcome what is called the *work function* of the surface. There is therefore a wavelength threshold for this process. Radiant flux with wavelengths above the threshold wavelength for the material does not produce photoemission of electrons. Below the critical wavelength, the number of electrons emitted per unit time is proportional to the incident flux level.

The photoelectric effect was a puzzle to physicists when first discovered. For frequencies shorter (wavelengths longer) than a critical value, no amount of flux

was sufficient to free electrons from their atoms. Furthermore, increasing the irradiance at frequencies above the critical threshold didn't produce more energetic electrons, it only increased the number of them ejected from the metal. This effect provided direct confirmation of the quantum nature of light and forced a new view of physics onto a reluctant scientific community. Its explanation led Albert Einstein to receive the Nobel Prize in physics.

Optical radiation detectors using the photoelectric effect are categorized by the manner in which they take advantage of the emitted electrons to produce an output signal. The simplest of these is the *vacuum phototube*. The special surface receiving the radiant flux and emitting the electrons is called the *photocathode*. The *cathode* is negatively biased. Another electrode, called the *anode*, is positively biased and is placed in proximity to the cathode, as illustrated in Figure 7.7.

If Φ_p is the photon flux (defined in Chapter 1, as the number of photons per second) incident on the photocathode, η is the photoelectric efficiency—the fraction of absorbed photons producing electrons (with units of electrons per photon)—and q_e is the charge of an electron (1.6×10^{-19} coulomb), then the current i generated by the flux Φ_p will be given by

$$i = q_e \eta \Phi_p \quad (7.13)$$

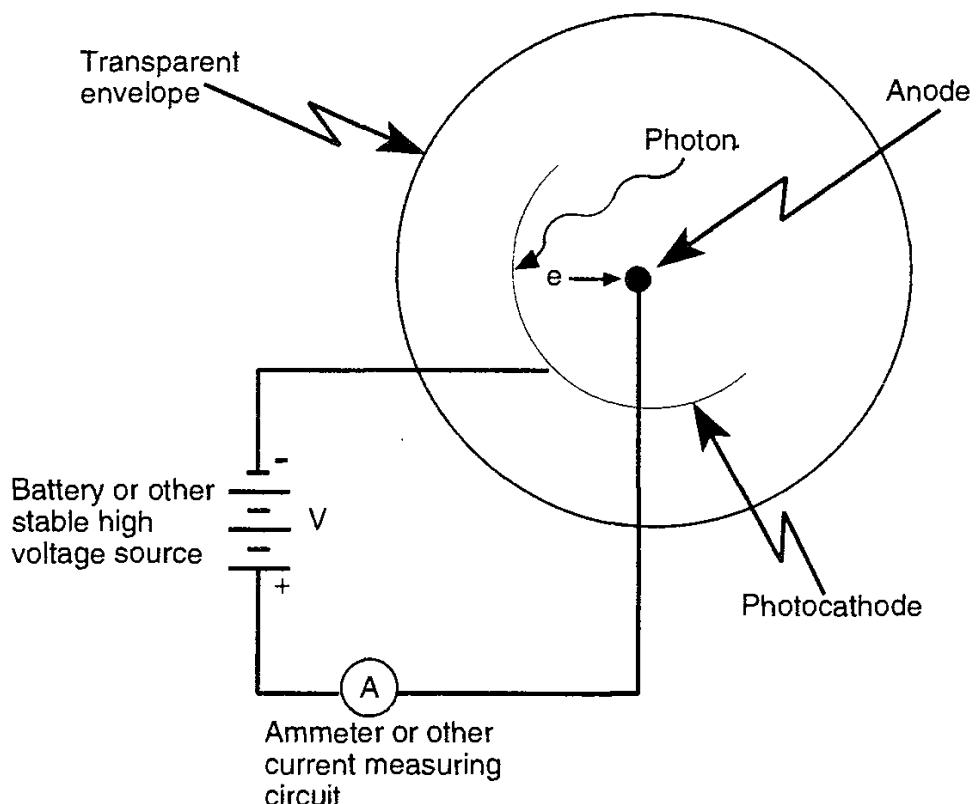


Figure 7.7 Schematic diagram of a vacuum phototube and its associated electrical circuit.

assuming that all the photons have energy $E = h\nu$ greater than the work function ϕ_o of the photocathode, i.e.,

$$\nu > \frac{\phi_o}{h} \quad (7.14)$$

with ν being the frequency and h being Planck's constant.

Any excess photon energy is converted into kinetic energy (given by $\frac{1}{2}mv^2$) of the photoelectron

$$\frac{1}{2}mv^2 = h\nu - \phi_o \quad (7.15)$$

with m being the mass of the electron and v being its velocity. (Although they appear similar, the Greek symbol ν (nu) is for frequency and the italic Roman symbol v (vee) is for velocity.)

With no radiation incident on the cathode, no electrons are released and no current flows through the circuit. Vacuum is a good insulator and it takes a very high voltage to force electrons out of the cathode and over to the anode in the absence of radiation. When this happens, it is called *field emission*. When radiation below the critical wavelength is incident on the cathode, however, electrons are emitted and drawn to the positively charged anode, causing a current to flow even when modest voltages, below the field emission threshold, are applied to the device.

With this detector, for each electron ejected from the cathode by a photon, called a *photoelectron*, one electron is contributed to the current in the circuit. The current output of a phototube therefore is linearly proportional to the flux received by the cathode. If the voltage is reversed, no current flows. The device therefore is a diode, permitting current flow in one direction (proportional to the incident radiant flux) and preventing it in the opposite direction in the circuit. Devices that behave in this manner are called *photodiodes*. For weak radiant fluxes, the current in a photodiode can be very small and difficult to measure or "buried in the noise," meaning that the *SNR* is well below unity. In such cases specialized signal recovery electronic systems can be used to enhance the SNR. At very low irradiation levels, individual photons may arrive at the photocathode at random intervals, producing pulses of current or charge packets. If the detection electronics are sensitive enough, the result is what is called a "photon-counting" system.

The second category of photoemissive detector improves on the weak signal response of the photodiode by increasing the number of electrons contributing to the output current for each photoelectron produced at the photocathode. This feat is accomplished by the addition of several intermediate anodes to the vacuum tube. Each new anode, called a *dynode*, is given a voltage higher than the previous one

and they are arranged in a special geometrical configuration, shown schematically in Figure 7.8. If the voltage is high enough, each photoelectron drawn from the photocathode to the first dynode by the potential difference between them is made to arrive with sufficient energy to cause the ejection of multiple electrons from the dynode. These electrons are then drawn to the second dynode. Because this dynode has a still more positive (higher) electrical potential, the secondary electrons are drawn to it with enough energy to cause another increase in the number of electrons flowing between dynodes in the vacuum tube.

This process is repeated a number of times, producing a cascade of secondary electrons flowing from the dynodes. Because the device multiplies the electrons emitted by the photocathode, it is called a *multiplier phototube* (MPT). (In the past and in some current writings, this device is called a photomultiplier tube. This terminology can be considered a misnomer, because it is not the photons that are being multiplied but the electrons. On the other hand, one could say that this device multiplies the *effect* of the incident photons and can therefore reasonably be called a *photomultiplier tube* or PMT. In any event, the photomultiplier tube terminology is so widespread that it is unlikely to be changed in the near future.) Since the greater the voltage, the greater the multiplication of electrons in the dynode chain, very high voltages are desired for maximum overall response. Thou-

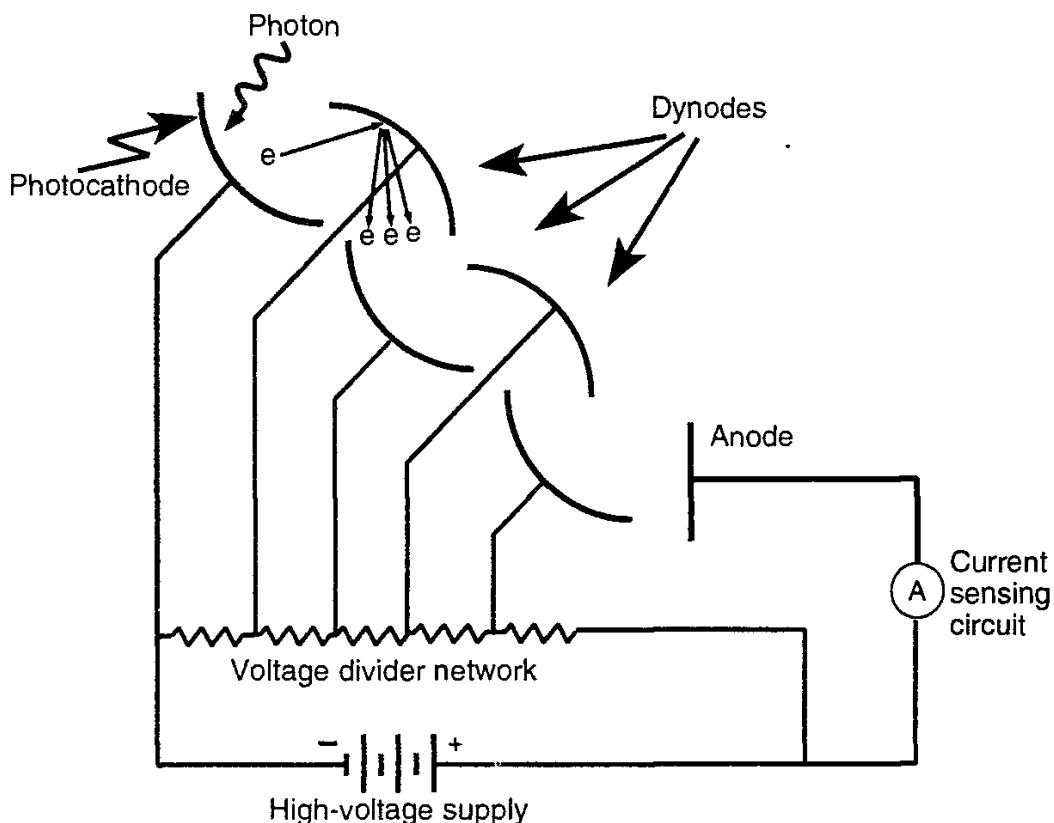


Figure 7.8 Schematic diagram of a multiplier phototube.

sands of volts are not uncommon for many PMTs. There are limits to the maximum voltage that can be applied to a PMT. One is determined by the breakdown voltages of the insulating materials and spacings used to electrically isolate the dynodes from each other. Another is the maximum voltages that can be applied to the resistors in the chain of resistors frequently used in a voltage divider network, as illustrated in Figure 7.8, to provide each dynode with the proper voltage. The current flowing through the last dynode in the chain contains the highest number of electrons and is the one from which this device's output is drawn.

For still greater multiplication of the secondary electrons, a gas can be placed in the PMT. If the right gas is chosen, with the correct density, the secondary electrons will ionize some of the neutral atoms in the gas. In this case, the free electrons that are stripped from the atoms are accelerated by the electrical potential differences between the dynodes and added to those emitted from the previous dynodes, causing an even higher multiplication of the electrons cascading from dynode to dynode [1].

Modern PMTs have a thin flat semitransparent photocathode deposited in the inside of an evacuated glass envelope containing the dynode chain. The photocathode is illuminated through the glass and the photoelectrons are emitted into the vacuum on the other side of the coating. This straight-through arrangement produces a well-defined and homogeneous photocathode surface area. The dynode chain is specially constructed as well, with the total arrangement being somewhat like the illustration in Figure 7.9.

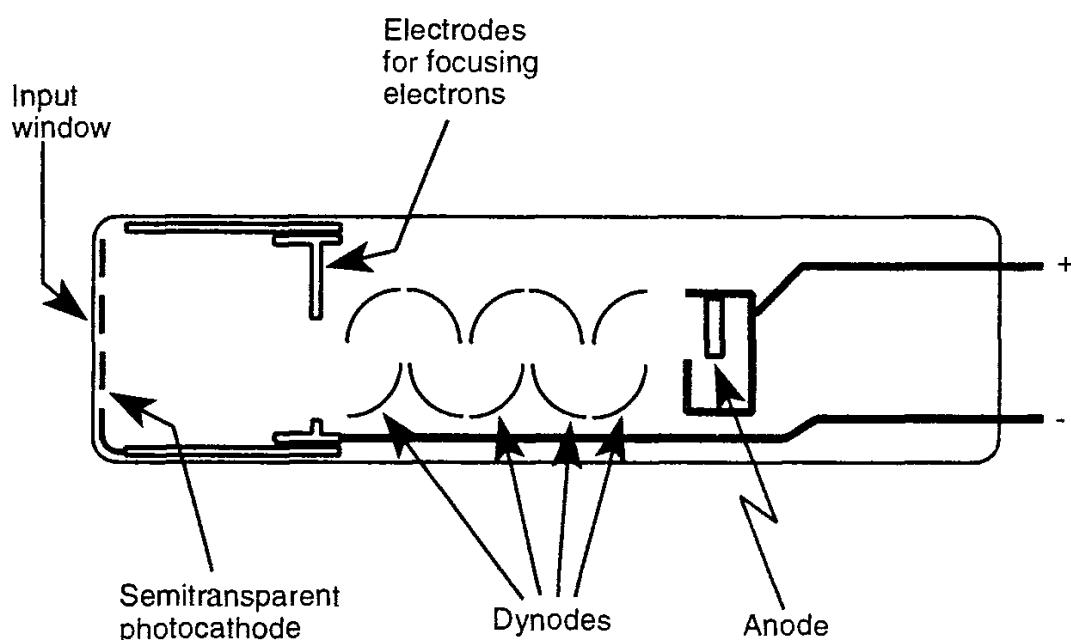


Figure 7.9 Representative geometric placement of the photocathode, multiplying dynodes, and anode of a modern photomultiplier tube. (Courtesy Oriel Corporation.)

There are several variations on the multiplier phototube. One very interesting one is the microchannel, a small tube of evacuated glass that operates as an electron multiplier. The glass of the tube is made of oxides of silicon, lead, and alkali compounds so as to provide just enough conductivity to give a gain of two electrons ejected per single electron collision against the wall of the glass tube. A large voltage is applied to the ends of the tube, creating in effect a continuous dynode chain with a strong electric field down the length of the tube. Ejected electrons are drawn down the tube, which is slightly bent, until they collide with the wall of the tube at a higher potential and produce secondary electrons, which are themselves drawn further down the tube to produce still more wall collisions further along.

Single microchannels have been constructed with lengths of 5 cm and having electron gains as great as 10^8 when the potential difference applied to the tube is 3,000V. The photocathode is placed near to the negative end of the tube and is given a slightly lower potential than the end of the tube, so that the photoelectrons are drawn into the tube. An anode is placed just beyond the positive end of the tube to collect the many electrons emerging from the tube. An advantage of this device is that it can be made quite small and therefore can be used in two dimensional arrays for imaging applications [1].

The gain of a photomultiplier tube depends upon the voltage applied to it. Because of this fact, very slight changes in the high voltage powering the PMT can produce significant changes in the output current signal, even with no change in input radiant flux levels. In consequence, it is very important that the output of the high-voltage power supply be very well regulated, thermally compensated, and in other ways made to produce an extremely constant output voltage. This task is aided somewhat by the low currents drawn by most PMT designs.

All photoemissive devices are characterized by very fast time response. The time between the absorption of a photon and the collection of an electron at the anode of a photodiode can be less than 10^{-10} sec. These devices also have large dynamic ranges and low noise levels. Photocathode spectral responses however, generally vary strongly with wavelength, and the commonly used ones have negligible responses in the infrared portions of the spectrum. Quantum efficiencies η are generally in the range from .01 to .30 electrons per photon (i.e., it takes from about 3 to 100 photons to produce one electron). An exception to this is a special kind of cathode made of a semiconducting material specially devised to have what is called a negative electron affinity. This means that any electron in the conduction band of the semiconductor will be expelled from the material [1]. A negative electron affinity photocathode made of GaAs-P can have $\eta > 0.3$ at 700 nm.

7.3.3 Semiconductor Devices

For a thorough understanding of the operation of semiconductor detectors, knowledge of condensed matter (solid state or solid phase) physics is needed. A number of good introductory texts on solid-state physics have been published and succinct descriptions can be found in college level textbooks [7–10]. The important concepts are described here briefly.

In gases, at low absolute temperatures, atoms and molecules are widely spaced and do not interact with each other strongly. The electrons in each atomic and molecular species have their own set of allowable discrete energy levels. In general, these are relatively widely spaced. According to the quantum theory, electrons bound to the neutral atoms in such a gas can change energy only between these levels, either by the absorption of discrete energy increments or by the emission of discrete quanta of energy equal to the difference between the initial and subsequent energy levels.

As a gas is pressurized or condensed into a liquid or solid, however, the interatomic spacing decreases. As two identical atoms are drawn closer together, their energy levels are perturbed. According to *The Encyclopedia of Chemical Technology* [11]: “An electron situated between the nuclei is attracted to both and, therefore, experiences a lower potential energy. The energy level associated with the unperturbed electron undergoes splitting. The lowered energy level corresponds to an electron having a high probability of residing between the nuclei. Two electrons of opposite spin may occupy this bonding level, which results in a new lowering of the electron energy of the system. . . . As more nuclei are added to this system, additional energy-level splitting occurs. For a large atomic array, many closely spaced levels form bands of allowed electron energies.”

An example of this process is illustrated with the energy diagram of Figure 7.10 for silicon. Silicon has two electrons in a 3s energy level and two in a 3p level, the numbers referring to energy states and the letters to spin states. According to the encyclopedia cited above: “As the silicon atoms are brought together into the diamond configuration, the 3s and 3p levels begin to broaden and the gap between them decreases. Eventually the gap disappears, and electron states take on a mixed s and p character; that is, hybridization. Further decreases in atomic spacing result in the appearance of two new energy bands separated by a gap E_g . ”

At a temperature of absolute zero, the electrons attached to atoms in a pure crystalline solid such as silicon fill all the lower energy levels allowed to them by the rules of solid-state physics. These energy levels are numerous and narrowly spaced in energy, resulting in what is called a *band* of energies. The highest or topmost filled energy level of a conductor at absolute zero is called the Fermi level, E_F by many authors. The range of energies below this level that is occupied by the valence electrons attached to atoms in the solid is called the *valence band*. Valence

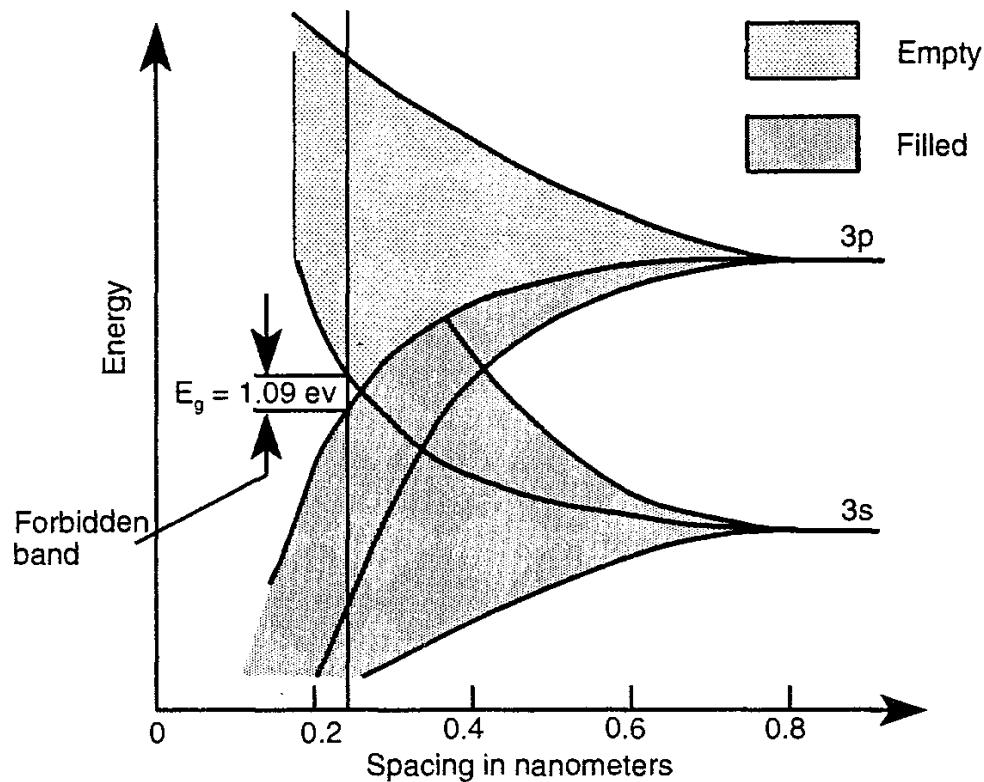


Figure 7.10 As the interatomic spacing between silicon atoms in a crystal lattice decreases, the energy levels broaden into energy bands. (After *Encyclopedia of Chemical Technology*) [11].

electrons are those in the topmost energy states that participate in chemical bonding of atoms to one another in the formation of molecules.

With increasing temperature, thermal motions of the atoms in the lattice of certain crystals impart energy sufficient to elevate a number of the electrons to higher energy levels. The distribution of electrons amongst the energy levels available to them is governed by statistical distribution functions that arise from various theories of the behavior of electrons in solids. The materials of interest in this discussion, those used for optical radiation detection, are generally operated at temperatures significantly above absolute zero. A number of electrons occupy the higher energy levels within such solids. The range of energy levels occupied by these electrons is called the *conduction band*. Also above the valence band lie a number of energy levels, called the *excitation levels*, which are populated by electrons following their acceptance of energy from electrical or radiant stimuli—the flow of electrical current through the material or the absorption of radiation incident on the material. According to the terminology of Hunter [12], the excitation levels include the energy levels of the conduction band. If one can find a solid material for which at normal temperatures few electrons occupy the conduction band, but

which can be elevated to the conduction band by the absorption of radiant flux, this material can be used as a photoconductive detector, a device whose conductance increases with increasing absorbed flux. This is one reason why knowledge of the band structure of solids is important in designing and using detectors based on these principles.

The *ionization energy* E_i is the energy needed to strip an electron completely from its bonding to the atoms in the lattice. At absolute zero, therefore, the work function ϕ_o encountered in the earlier discussion of the photoelectric effect is equal to the difference between the ionization energy level and the Fermi energy level:

$$\phi_o = E_i - E_F \quad (7.16)$$

The relative positions of these energy bands and levels are illustrated generically in Figure 7.11. Filled energy levels in the valence band are indicated by black dots. The difference in energy between the bottom of the conduction band and the top of the valence band is called the *band gap* energy, E_g .

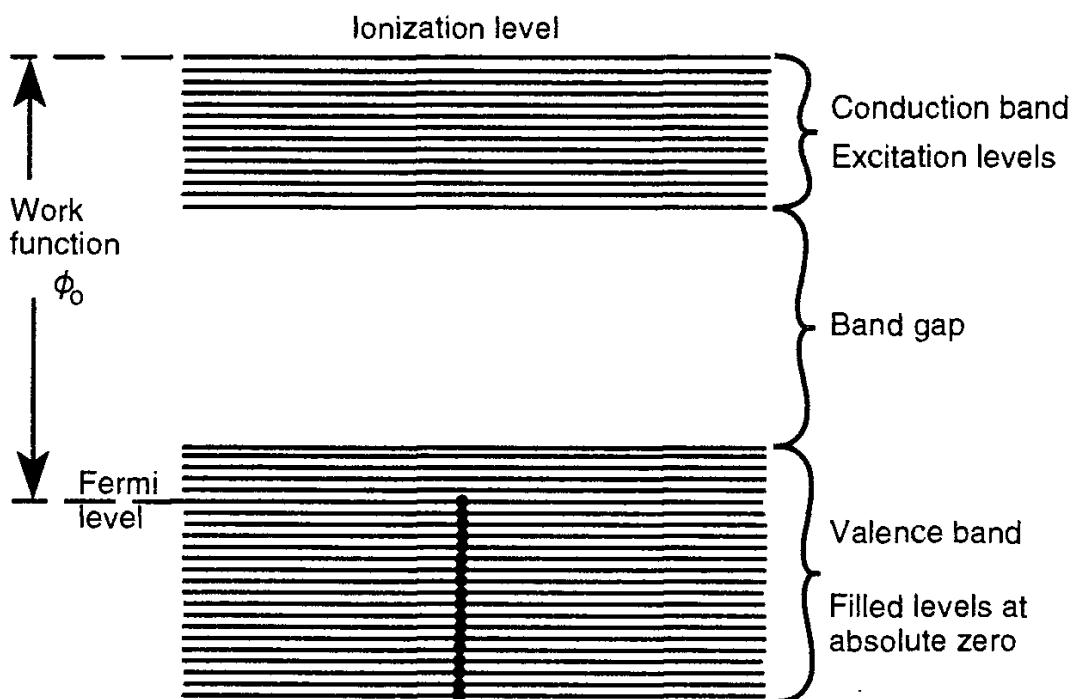


Figure 7.11 Illustration of the relative positions of conduction and valence energy bands in solids. The work function is the energy needed to free electrons bound to the atoms of the crystal lattice from the crystal. Filled energy levels at $T = 0\text{K}$ are indicated by black dots. The Fermi level, E_F , is described in the text.

7.3.3.1 Insulators

The Pauli exclusion principle says that each energy state in a finite solid can be occupied by one electron of positive spin and one electron of negative spin. If all these states are filled by electrons and no additional ones are available for the electrons to occupy, the electrons are bound to their atoms and cannot move around in the solid. This is the case with an electrical insulator.

An *insulator* receives this designation by virtue of its valence band being completely filled with electrons at absolute zero and because of a separation between its valence and conduction levels large enough that thermal motion at normal temperatures is insufficient to put electrons into higher conduction band energy states where they can become mobile and participate in an electrical current. This situation is illustrated in Figure 7.12(a). The electrons are bound to their atoms and an applied electric field cannot dislodge them (unless it is high enough to produce field emission of electrons from the solid). The Fermi level lies between the valence and conduction bands in an insulator. Solids in which the highest filled band is completely occupied by electrons and no other unfilled bands are available are insulators [1].

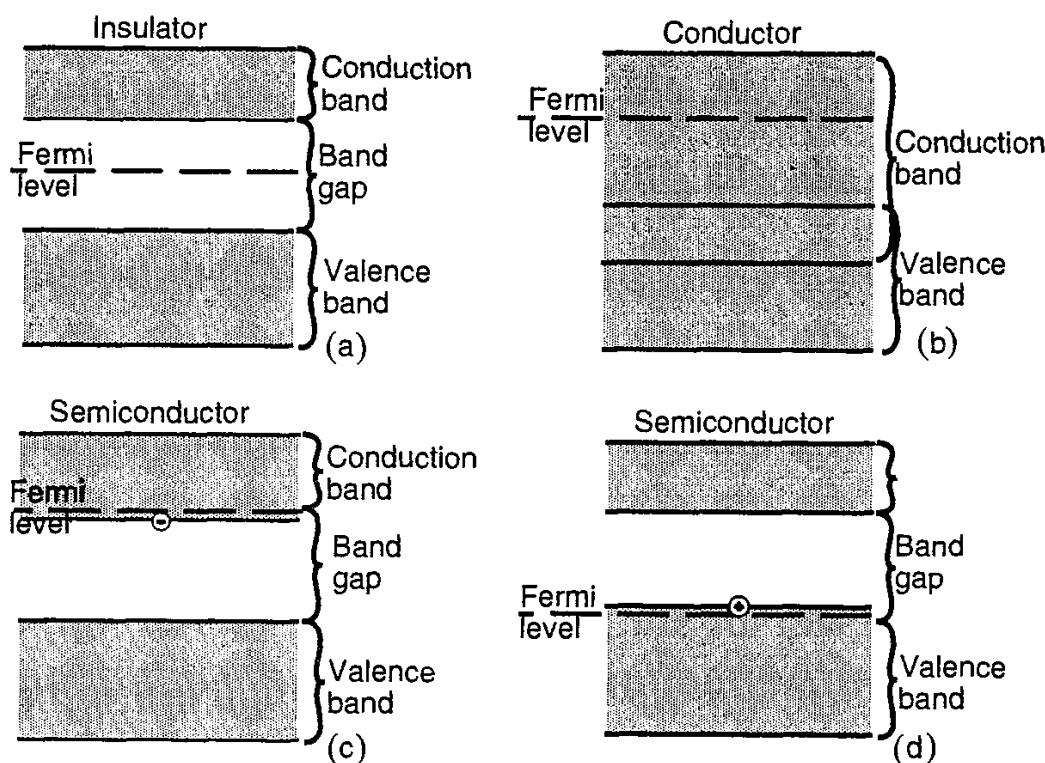


Figure 7.12 Illustration of the relative energy band positions for (a) an insulator, (b) a conductor, (c) an N-type semiconductor, and (d) a P-type semiconductor.

7.3.3.2 Conductors

In some crystals the valence levels and the lowest excitation levels overlap in energy and the bandgap shown in Figures 7.11 and 7.12(a) disappears. This produces a number of unoccupied energy levels near to those occupied by the valence electrons, as shown in Figure 7.12(b). The imposition of an electric field across such a crystal (as is produced when a voltage is applied to opposite sides of the material) can move an electron from its valence level on one atom to an excitation level (or conduction band) on another, without the need for it to change its energy level appreciably.

Valence electrons in the lower portion of the next higher energy band are free to flow in the solid. The repulsive force of an electron moved from one atom to another atom can then move the adjacent electron to another atom, and so on, producing a domino effect of electron motion, manifested on a macroscopic scale as an electric current. Crystals having these properties are classed as *conductors*. The Fermi level lies at the top of the filled levels at absolute zero, wherever in the overlapping energy bands this happens to occur.

A special case is that of the alkali metals, solids with one valence electron per atom, where the valence band is inherently only half-filled, leaving the valence electrons free to move about the crystal and permitting current flow upon the application of a voltage, even without the need for overlapping energy bands. The Fermi level here lies in the middle of the half-filled band, on top of the highest filled energy level at absolute zero.

7.3.3.3 Semiconductors

A *semiconductor* is a material in which the energy gap between a normally filled valence band and a normally empty conduction band is large enough for it not to be classed as a conductor but small enough that a few electrons can be elevated to the conduction band at normal temperatures, just by the thermal agitation present in materials above a temperature of absolute zero. The few electrons in the mostly empty conduction band are free to respond to applied electric fields and carry small currents. The current is limited by the relatively few numbers of electrons available to participate in the process. The electrical conductivities of such materials at room temperature is modest.

For each electron momentarily in the conduction band, there is a void in the valence band which it left. This void can be filled by an adjacent electron in the lattice, producing another void at the lattice position previously occupied by this electron. These voids are free to move about in the crystal and they behave electrically like positive charges. They also can participate in current flow through the

crystal, but their motions are in opposite directions to the electrons by virtue of their opposite charges. The voids are known as *holes*. They serve as positive *charge carriers*. (The *mobility* of the holes is generally less than that of the electrons.) The population of conduction band electrons can be calculated using the appropriate statistical distribution function for the temperature of the solid. For absolutely pure silicon at 300K, the total population of charge carriers is about $4 \times 10^{20} \text{ cm}^{-3}$ [13].

As might be expected, the conductivity of a semiconductor generally increases with temperature. At high enough temperatures, semiconductors become conductors. The conductivity can also be increased by the irradiation of the semiconductor with light of sufficient frequency to elevate valence electrons into the conduction band. Keeping the temperature constant, the changing conductivity of a semiconductor will be approximately linearly proportional to the absorbed light level. The conductance of a length of semiconductor material gives a good measure of the radiant flux received by that material.

The previous description of semiconductors applies to what is termed the *intrinsic* semiconductor effect, exhibited by very pure crystals of semiconductor material. At normal temperatures relatively few electrons will be in the conduction band for intrinsic semiconductors and their conductances are relatively low in consequence. Real materials are seldom pure enough to show intrinsic semiconductivity at room temperature. The impurities generally found in semiconductors can be used to advantage if they are introduced intentionally in controlled concentrations for a specific purpose. Properly chosen impurities, introduced into the semiconductor by a process called *doping*, can place available electron energy states in the previously forbidden gap between the valence and conduction bands. Such semiconductors are termed *extrinsic*.

In an *N-type extrinsic semiconductor*, the impurity atoms introduce, or *donate*, electron energy states just below the conduction band, effectively lowering the conduction band boundary and raising the Fermi level. This is illustrated in Figure 7.12(c). The Fermi level is raised because at absolute zero all the valence band states are filled with electrons, as are the new donor level states just below the conduction band. With increasing temperature, however, the Fermi energy drops, approaching the value $E_g/2$ it has for an intrinsic semiconductor at temperatures of several hundred degrees K. N-type semiconductors conduct electrical current mainly by electron transport.

In a *P-type extrinsic semiconductor*, illustrated in Figure 7.12(d), the impurity atoms accept electrons and create a surplus of hole states, called *acceptor energy levels*, that lie just above the valence band. The Fermi level in this case is lowered to just below the acceptor states and just above the valence band. With increasing temperature, the Fermi level rises, approaching $E_g/2$ at high temperatures. P-type semiconductors conduct electrical current mainly by hole transport.

A very interesting device is formed when part of an intrinsic semiconductor is doped with N-type impurities and an adjacent area with P-type impurities. The

two regions are forced to be directly adjacent, in electrical contact in the material. In this case, the electrons migrate from the N-region to P-region acceptors and holes migrate from the P-region to N-region donors. Thermal excitation of the hole-electron pairs also produces charge carrier transport across the boundary region. The net result is the establishment of a permanent equilibrium voltage across the boundary. The device becomes a nonlinear circuit element, a diode, in which an external voltage greater than and opposite to this inherent potential difference causes a large forward current through the device while an opposite voltage reinforces the potential barrier and very little current flows.

There is another interesting characteristic of some P-N junction devices that is very useful for both the design and construction of high-sensitivity detectors and the conversion of optical radiation directly into electrical power. When photons of energies greater than the forbidden gap energy E_g are absorbed by a P-N junction of such devices, electron-hole pairs are formed and the built-in potential difference pulls electrons from the P region into the N region and holes in the opposite direction. If a good conductor is connected to the two regions, a current will flow through what is effectively a short circuit. Since the current is linearly proportional to the incident radiant flux, such devices make excellent detectors. This phenomenon is called the *photovoltaic effect* and devices utilizing it are called photovoltaic detectors, photovoltaic diodes, or photovoltaic cells. (The latter term is generally applied to solar cells intended to produce electrical power from solar radiation for other uses, not just for the measurement of radiant flux.) These devices can be operated with a reverse bias voltage or without as illustrated in the circuit diagrams shown in Figure 7.13, to enhance certain performance characteristics. The photovoltaic effect can be used to convert solar radiation into electrical power with an efficiency on the order of 10–15%. Many research efforts are underway to increase the efficiency of these devices as well as to lower their cost of manufacture.

With a reverse bias voltage, a change in conductivity is produced when photons of energy greater than the bandgap energy are incident, exciting additional electrons into the conduction band. These new conduction electrons and their hole counterparts in the valence band are available as charge carriers, thereby increasing the conductivity and passing a current proportional to the incident radiant flux of the proper frequencies. In this mode, the photodiode is referred to as a *photoconductive detector*.

In practice, the photodiode is most commonly connected as illustrated in Figure 7.13(a). The external circuit, the inverting input of an operational amplifier, appears to the photodiode as a virtual short circuit. This configuration converts the very small photocurrent into a voltage signal at the operational amplifier output, which is linearly proportional to the input current. The current through the photodiode, rather than the voltage across it, is the measured quantity, since it is linearly proportional to the incident flux over a very wide range. A very important advantage of such detectors for radiometric and photometric applications is the

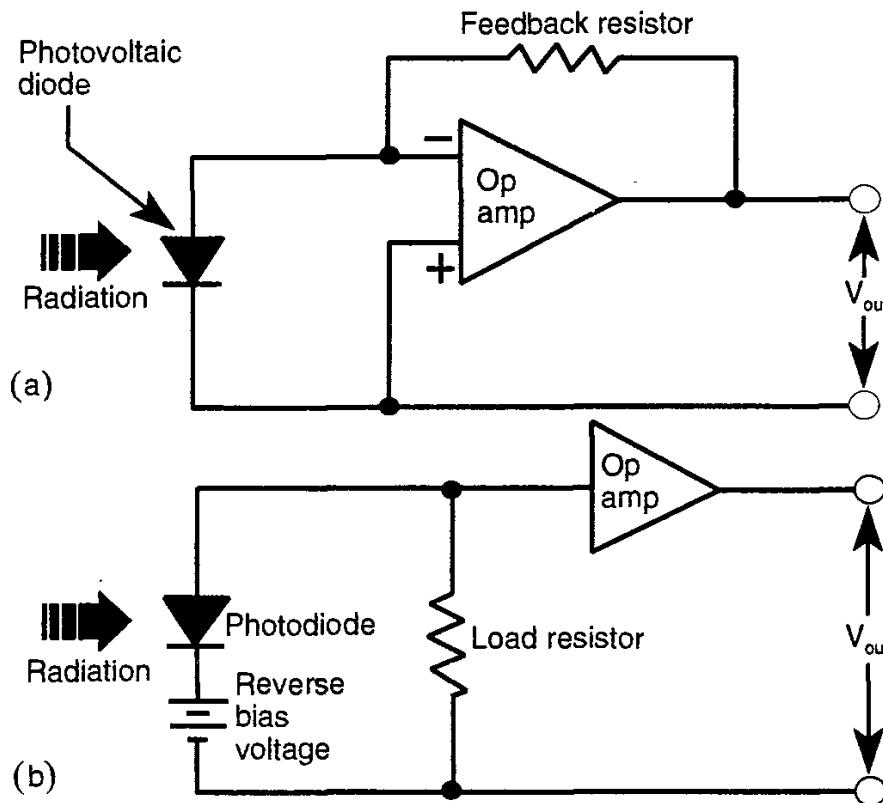


Figure 7.13 Illustration of two photodiode circuit connection schemes. (a) Unbiased method, (b) biased method.

huge range of linearity that most of them exhibit, commonly covering more than five orders of magnitude.

To avoid amplifier noise or saturation, the gain of the amplifier can be altered to change the range of input photocurrent, which produces the full range of output voltages available from the amplifier. A multiranging radiometer or photometer is one for which the operator, or an internal electronic circuit, can change the gain of the operational amplifier so that the range of useful output signals from the amplifier better matches the range of flux levels incident on the detector.

The *silicon photodiode* is made by diffusing N-type material into a P-doped substrate material, or the reverse of this. The result is a movement of charge carriers away from the junction region, creating a *depletion zone* across the junction. The resulting photodiode has many desirable characteristics as a detector of radiation having wavelengths from about 350 nm to over 1,000 nm.

If a narrow region of undoped silicon separates P and N-type regions, the resulting detector is called a PIN photodiode, the “I” standing for the intrinsic semiconductor. The resulting photodiode structure responds to incident radiant flux having wavelengths from 200 to 1,100 nanometers, a spectral range that covers a useful part of the ultraviolet spectrum, the entire visible region, and the near infrared spectrum. (Special techniques [14] are needed to extend the range of

photodiode response below about 350 nm, producing what is called a “UV-enhanced” photodiode.) The response to absorption of an incident photon is the creation of an electron-hole pair charge carrier within the silicon. The minority carrier is pushed across the junction due to the internal electric field to a region where it can be drawn through an external circuit. Both photoconductive and photovoltaic modes are possible with PIN photodiodes. In general, photoconductive devices are faster in response due to the external reverse bias voltage applied (generally at the expense of increased noise), while photovoltaic devices are the simplest and least expensive to employ and less noisy. PIN photodiodes are characterized by their high-speed responses to changing input flux levels.

For optimal performance, photodiodes should be designed for operation in one or the other of the operating modes. It is therefore important to consult manufacturer product literature for advice on operating circuits and other application information [15].

During the Second World War, lead sulfide (PbS) photoconductive detectors were widely used for infrared radiometers or imaging devices. Although there has been much expectation over the years that they would be replaced by better devices, and while it is true that a wide variety of infrared-sensing detectors has become available, the versatility of the PbS detector has kept it in great demand as sensors for major military programs as well as for industrial, commercial, and medical applications [16]. Useful spectral sensitivities of various formulations of the PbS detector cover ranges from somewhat below 1 μm to around 4 μm .

An alternate IR detector that has found widespread application in the field of pollution monitoring, is the thin film lead selenide photoconductor. It is formed by chemical deposition onto a suitable substrate, using the techniques of photolithography. Gold electrodes provide noise-free contact between the PbSe film and the lead-out wires. Detector response is strongly sensitive to temperature. The cooler the detector can be operated, the greater the magnitude of the normalized detectivity D^* (defined in Section 7.6.4) and the longer the wavelengths at which the detector can be used [17]. The latter varies from around 4.8 μm at room temperature to 7 μm at -196°C .

The properties and comparisons of commercially available detectors are presented in Section 7.10.

7.3.4 Multi-element Detectors, Charge Transfer Devices, and Imagers

So far we have discussed only single-element detectors, detectors having only a single output signal. Numerous side-by-side arrangements of detectors, even in the same housing, are possible. A useful one is the so-called quadrant detector, illustrated in Figure 7.14. It is useful for detecting the center of a symmetrical beam of radiation, such as that produced by a laser. If the beam is centered on the

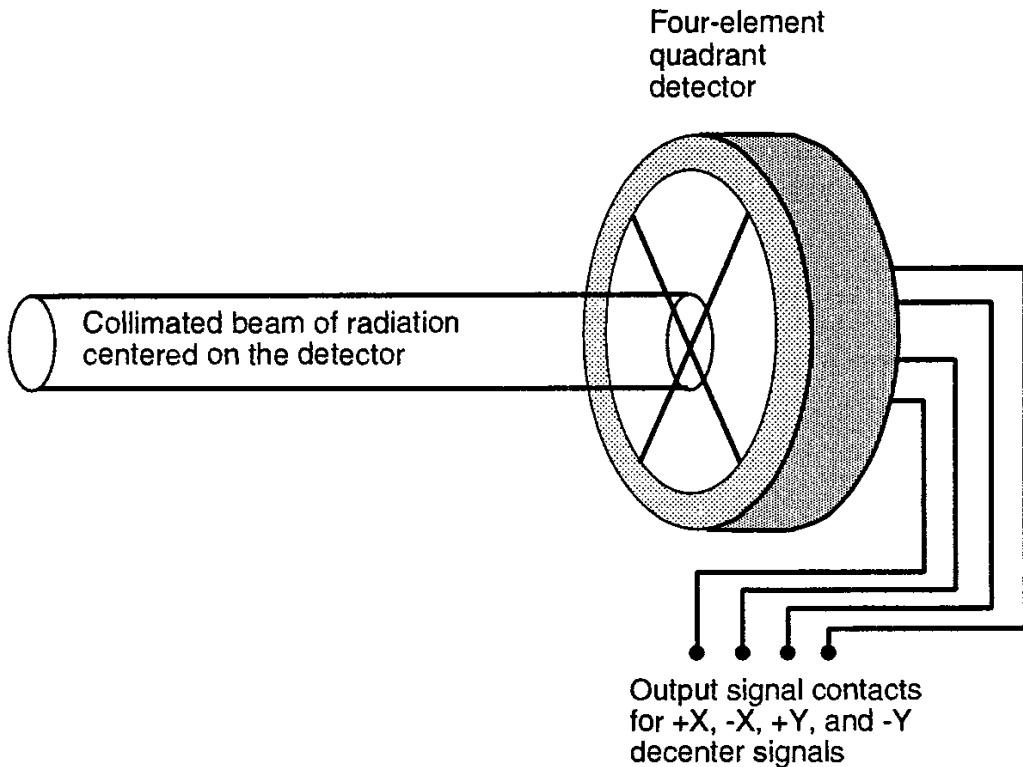


Figure 7.14 Quadrant detector. A centered beam produces equal signals from the four detectors.

intersection of the boundaries between the four detectors, each detector will output the same signal. Any decentering of the beam will produce a drop in the signal from one or more detectors and an increase in the signal from one or more of the others, indicating the direction and the approximate magnitude of the decentering.

The Sidewinder heat-seeking missile's guidance system is based on the use of a radiometric system similar in principle to that of a PbS quadrant detector. The device provides direction signals to the missile's steering mechanisms, aiming it to an essentially point-source of infrared radiation whose radiance is much greater than the background radiance over the wavelengths of interest. For more sophisticated guidance, additional detector areas can be added to the quadrant detector at increasing radii from the center of the detector. Such an arrangement can produce a steering signal that depends upon both the direction of misalignment and its magnitude.

It becomes impractical to increase the number of detectors in a two-dimensional array beyond about 100 by the simple addition of components. One reason for this is the need for at least two conductors attached to each detector and a preamplifier close to each element in the array for noise mitigation purposes. If the number of detectors in a two-dimensional array can be increased significantly, a whole new class of imaging radiometers is made possible.

The *charge-coupled device* (CCD) was developed by Boyle, Smith, Amelio, and Tompsett [18] to overcome such limitations. The CCD array is a closely spaced array of metal-insulator-semiconductor (MIS) capacitors. The arrangement is shown schematically in Figure 7.15. The insulator is frequently an oxide, in which case the capacitor receives the designation MOS. The semiconductor material (typically silicon) is spatially continuous, as is the insulator (typically silicon dioxide). The metal is discontinuous and is provided with a positive voltage, which lowers the energy levels in the semiconductor adjacent to the insulator.

Electromagnetic radiation can enter the semiconductor from the “front” (if the conductor is of a material or is thin enough to transmit radiation of the appropriate wavelengths) or from the “back” (if the relevant components are semitransparent).

When a photon of energy greater than the bandgap energy is absorbed by the semiconductor material under any one of these metal electrodes, it produces an electron-hole pair. The hole is mobile and moves out of the active region of the semiconductor adjacent to the insulator (again called the *depletion region*) and toward the electrical ground on the other side of the semiconductor. The electron

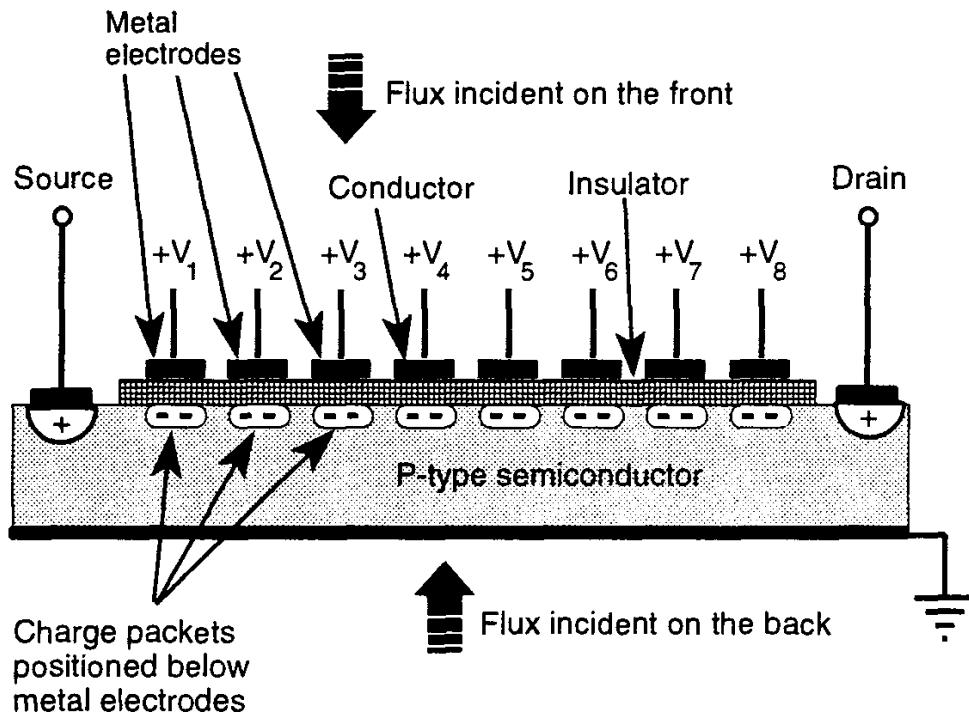


Figure 7.15 Schematic diagram of the principle of operation of a CCD array detector. Metal electrodes are given positive voltages relative to the semiconductor on the other side of an insulator. Conduction electrons created by absorbed photons in the semiconductor are drawn by the positive voltages on the metallic electrodes to create charge packets in the semiconductor just below each electrode. If $V_1 > V_2 > \dots > V_8$ momentarily, then the charge packet beneath each metal electrode can be drawn over to the next electrode to the right. The process is repeated and each charge packet is measured when it passes into the drain electrode.

remains fixed in the semiconductor, adjacent to the insulator, being attracted there by the positive charge of the metal electrode on the other side of the insulator. A quantity of charge is therefore held on opposite sides of the insulator that is proportional (within limits) to the number of photons absorbed by the semiconductor over a controlled period of time T . The total amount of electrical charge that can be collected in this way is proportional to the size of the applied voltage and other parameters.

If the spacing between the two metal surfaces is correct and if the voltage of one surface is V_1 , and that of an adjacent one is V_2 , and if V_2 is sufficiently greater than V_1 , the negative charge beneath metal electrode 1 will be pulled over to reside below electrode 2. The application of a regular sequence of voltage pulses to the array of metal electrodes can cause sequential movement of the charge on each capacitor from the source electrode on the semiconductor to the next adjacent capacitor and thence on across the series of capacitors to the final electrode, called the drain, electrically connected to the semiconductor. The arrangement is shown schematically in Figure 7.15. For charge movement, one must momentarily pulse $V_1 > V_2 > \dots$, and so forth.

The transfer of packets of charge from one capacitor to another continues until the charge packet reaches the drain at the edge of the array where it can be collected and measured in an external circuit. By knowing the timing of the charge transfers, it can be determined from which location in the array a given packet of charge came. The amount of charge in each packet is proportional to the radiant flux incident on the capacitor during the collection time.

There is a variety of ingenious ways of sending voltage pulses to the metal electrodes to cause the charge beneath them to move sideways so that it lies beneath the next electrode over. In one scheme described by Dereniak and Crowe [2], only four different electrical connections to a linear array are needed to cause a charge to move across a number of capacitors. With a sequence of voltage pulse maneuvers, the charges are pushed along through the array in a controlled manner until they can be collected at the end and measured. By electronically controlling the phasing of the pulse sequences, one can know which capacitor accumulated any given measured packet of charge. Thus one can determine quantitatively the radiant energy absorbed by the semiconductor beneath each of the metal electrodes in the two-dimensional array.

Such a two-dimensional array can be used to record and store an optical image formed by a lens or other imaging system placed before the array. It can, for example, be used as the basis for a television camera, capturing images and delivering the electronic version of the image to a VCR tape or a transmitter for broadcast. Each element of a discrete two-dimensional array of detectors such as this makes up a point in the recorded image—called a pixel—a single picture element. Linear CCD arrays can be used to collect and measure radiant flux dispersed by a spectrometer. The result is a useful variation from the traditional

design of scanning spectroradiometers that has the advantage of completing a spectral "scan" in a time period much shorter than the previous approaches. Spectroradiometry is discussed in Section 9.6.

For extended spectral response, hybrid CCD arrays have been developed in which the detector material and the CCD are separate components bonded together to form a detector array. Thus the spectral response can be set by the detector material, independently from the charge transfer mechanism. These hybrid devices are most important for systems needing significant infrared sensitivity where typical infrared photoconductors such as mercury cadmium telluride are not compatible with CCD operation. This arrangement allows the use of the same CCD system with a variety of different detector materials having different spectral responses. Since the main connection between the detector and the CCD is used only for transferring photon-induced charge carriers from the detector to the CCD capacitor, even nonphotoconductive detectors, such as the pyroelectric one, can be used. As a result, CCD devices can be made with spectral responses covering the range from 1 to 30 μm ! On the other hand, cooling requirements (for noise reduction) are stringent for systems operating over long-wavelength regions. The reason is described in Section 7.6.5. Cooling requirements are slightly simplified for CCD arrays in that a relatively small number of heat-conducting wires is needed to transfer the image signal from the cold CCD array to the external, room-temperature electronics. It is therefore somewhat easier to keep these detectors cool.

CCD arrays and related semiconductor array devices have become widespread in commercially available imaging radiometers and photometers and in fast-response spectroradiometers.

7.4 DETECTOR NOISE

Before describing detector performance, it is important to discuss electrical noise. An electrical output signal from a detector that is proportional to the flux incident upon a detector is of little value if it is exceeded by the random electrical fluctuations, known collectively as *noise*, produced by a variety of sources.

The terminology originates in acoustics, where extraneous sounds of one or more frequencies degrade a desirable acoustic signal, whether it be music, the spoken voice, or an undersea sonar echo. If the output of a radiation detector is amplified and sent to a loudspeaker, then the presence of noise in the circuit can be heard as a superimposed hissing sound with a broad range of sound frequency components present. Some noise generated in the postdetector electronics, or in the detector itself, can be reduced by special techniques. Methods have been devised both to reduce noise levels and to avoid some of their limitations on detectability.

Random noise is distinguished from interference, or unwanted "pickup." The latter constitutes a systematic error output from the system. If the source of this

error is known, it is frequently possible to eliminate it either electronically or by subsequent numerical processing. An example is the unwanted intrusion of radio frequency transmissions into electronic circuits. Sometimes this can be eliminated by adequate electrical shielding of the susceptible circuits.

Radiant flux and detection systems generally suffer from two different kinds of noise. Radiation noise results from the fact that sources of electromagnetic radiation themselves exhibit an inherent temporal variability by virtue of the statistical nature of the photon emission process. As an example, the rate of photon emission from incandescent sources exhibits some randomness. The second kind of noise experienced in radiation detection systems is electrical noise produced in the detector and its associated electronics. Noise in this case comes from a variety of different physical processes, each having its own character.

Dereniak and Crowe categorize noise sources as follows [2]:

1. Photon noise
 - a. Noise contained in the signal radiation itself.
 - b. Noise due to background radiation from the environment of the signal source.
2. Detector-generated noise
 - a. Johnson noise (also called Nyquist noise)—caused by thermal motion of charged particles, thermal current fluctuations, in a resistive element of the detection circuit.
 - b. Shot noise—noise generated in photon detectors containing or utilizing a potential barrier as is present with P-N junction detectors such as the photovoltaic diode, resulting from variance in the rate of photoelectron generation.
 - c. Generation-recombination noise—produced in a photoconductor by fluctuations in the generation and recombination of current carriers.
 - d. 1/f noise—though not well understood, this noise is thought to be due to a lack of perfect conductive contact at the detector electrodes. It receives its name because the magnitude of the noise current falls off approximately with the inverse of the frequency. It is always present in photoconductors or bolometers requiring a bias current or voltage to operate. Photovoltaic detectors operating at zero bias do not exhibit this noise.
3. Post-detector electronic noise
 - a. Johnson noise
 - b. Shot noise
 - c. Generation-recombination noise
 - d. 1/f noise
 - e. Temperature fluctuations—output variations resulting from temperature changes not caused by the flux being detected. Sometimes called phonon noise.

Dereniak and Crowe list one more source for categories 2 and 3 above. The

phenomenon called *microphonics* is caused by changes in the electrical characteristics of the detector/measurement circuit, resulting from mechanical displacements of circuit elements when the circuit elements are subjected to vibration or mechanical shock. Pyroelectric detectors are particularly sensitive to this effect. Microphonics is not really a noise source, however, since it originates from a relatively known, not necessarily random, process. It is more of a systematic error or interference that can be eliminated if its manner of origination is known.

It is beyond the scope of this book to deal with noise sources in detail. The reader is referred to the treatments by Dereniak and Crowe [2] and by van der Ziel [19].

In assessing the impact of noise on radiant or luminous flux detection systems, we begin by assuming that the signal to be measured is essentially constant over some time period T . Another way of putting it is that we wish to begin by looking at time intervals T that are short compared with the variations in the signal that are of interest. Thus, any departures in the measured signal S from its (constant) value S_0 over this time period are by definition attributed to the noise sources. Of course, there could be a spurious source of current or voltage that is constant over T not produced by the flux on the detector. This would not normally be called noise. It would more properly be called a spurious output, a systematic error. The dark responses described previously fall into this category. What we are calling noise is generally characterized by a spread of frequencies—oscillations substantially random in nature.

“White noise” is a term given to noise sources whose frequency spectrum is essentially constant over the frequency range of interest—they contain a range of frequencies covering the spectrum. The frequency distribution of noise sources is important in the characterization of their effects on the detection of optical radiation.

When optical radiation measurements are noise limited, steps can be taken, such as cooling the detector to liquid nitrogen temperatures, to reduce the magnitude of the noise. One can also build electrical detection systems to filter out noise frequencies, leaving only fluctuations due to the signal being detected. This last idea leads to the important topic of signal modulation.

7.5 SIGNAL MODULATION AND RADIATION CHOPPING

Suppose a spectrum analyzer is attached to the amplified output of a photoconductive detector receiving a nearly constant flux (flux variations never exceeding some low frequency f_o in the range of one or two hertz). This device performs an analysis of the electrical power associated with each temporal frequency over the analyzer’s range of operation. The plot of Figure 7.16 illustrates what the spectrum analysis might look like. The signal to be detected, labeled S_o , is shown at the left edge of the plot, at frequency f_o . The problem is that the desired signal is less than

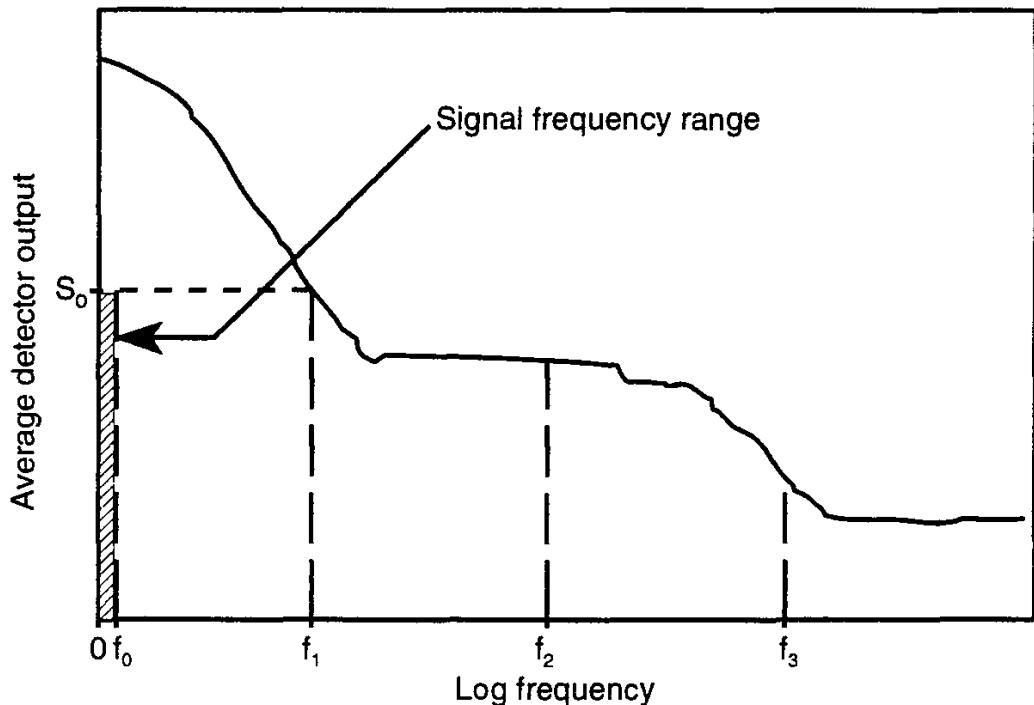


Figure 7.16 Sample noise spectrum for a photoconductive detector.

the noise at frequencies below f_1 . Higher frequency noise is also present in the detector output and will contribute to errors in the measurement. If the output is not filtered in some way, all noise frequencies will be contained in the output voltage or current and the detector output will fluctuate over a wide range, randomly, with both fast and slow components, covering up the real signal to be measured.

One obvious solution would be to attach a low-pass filter to the detector output, passing only the low-frequency signal components having frequencies near zero, and filtering out the higher frequency noise. According to the plot in Figure 7.16, however, the signal would still be lost in the low-frequency noise. If some way could be found to "move" the signal (with its narrow range of frequencies) we want to measure to a higher frequency, say f_2 , where the noise is less, it should be possible to reduce the noise contribution to less than the signal, producing a *SNR* greater than unity. (One could try and move the signal into the still lower noise range above frequency f_3 in Figure 7.16, but, as a practical matter, the detection electronics, or the detector itself, might not be capable of responding to such high frequencies. The thermal detectors discussed in Section 7.3.1, for example, generally are unable to respond to high-frequency fluctuations in the radiant flux incident upon them.)

Moving the flux signal to a higher frequency (a few hundred hertz or so is not atypical) is not difficult to do. Instead of trying to measure nearly constant flux, the radiant flux can be modulated, made to vary with a fixed and known frequency substantially higher than the frequencies of variation contained in the signal. In this case the radiation detector will output an amplitude-modulated, oscillating signal of fixed frequency, the envelope of the signal over time being a measure of the incident flux. The output electronics can be made to sense only oscillations at the signal's modulation frequency, excluding all others, including noise at other frequencies than the modulation frequency.

One method of modulating the flux reaching the detector is shown in Figure 7.17. Since the source in this case is in the laboratory, the radiation coming from it can be chopped at the source with a rotating sectored wheel, called a chopping wheel—a flat disk on a rotating shaft with open and opaque segments, as illustrated in Figure 7.17. As the wheel rotates, it alternatively blocks completely and passes completely the radiation emanating from the source through the sample area to the detector. This produces a modulation of the light, the rate or frequency of modulation depending upon the number of open segments and the rate of rotation. The waveform will depend on the shape and size of the beam passing through the chopping wheel in relation to the blades or segments of the wheel. Square, rectangular, triangle, and trapezoidal signal waveform shapes are common. The advan-

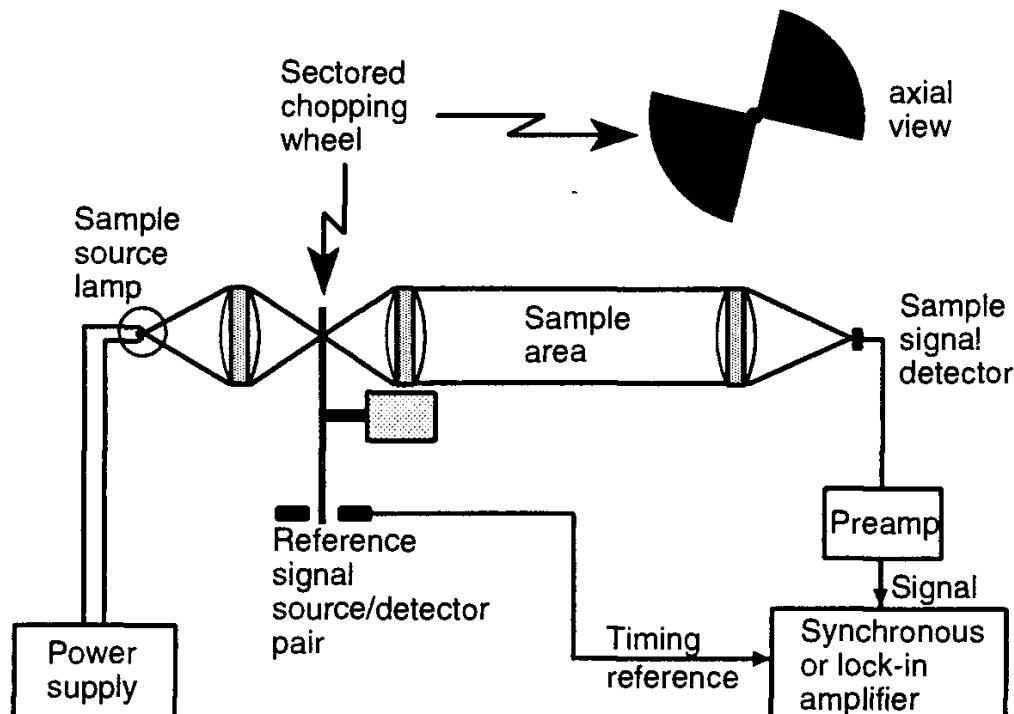


Figure 7.17 Schematic diagram of a hypothetical transmissometer showing the placement and role of a chopping wheel in modulating the sample beam for noise reduction.

tage of this system is that not only does it exclude noise at frequencies other than the chopping frequency f_c , but it also excludes any unchopped stray light that might find its way to the detector from the room or other sources. This stray-light rejection capability is an important attribute of chopped radiation source/detection systems.

The need for modulation of relatively slowly varying incoming radiant flux is not necessarily a disadvantage. Many radiometric and photometric detectors work best when the incoming flux is modulated at a fixed and known frequency selected to be within the speed range of the detector. Chopping wheels are commonly used in systems to minimize noise and stray-light effects or to detect signals at the chopping frequency that are much smaller than the accumulated noise over all other frequencies.

The synchronous or lock-in amplifier shown in Figure 7.17 is worth additional description. The light source and detector combination at the chopping wheel is provided to sense when the wheel is blocking and letting through the beam passing through it. This signal, which is in the form of a square wave having one voltage level when the beam is blocked and another when it is not, is fed into the "reference" input of the synchronous amplifier. (Other reference signal waveforms are possible.) The synchronous amplifier uses this reference signal to select a small range of signal frequencies centered on the reference frequency to amplify and deliver to the signal measuring circuit. Input frequencies outside the selected range are rejected. The process is synchronous because any changes in the speed of the chopping wheel (resulting in changes in the frequency of the desired signal) are automatically sensed by the amplifier and compensated for.

The width of the passband of the synchronous amplifier is an important parameter. The narrower it is, the greater is its exclusion of noise and the better the resulting *SNR*. The wider it is, the lower will be the *SNR*. One would normally desire a very narrow passband. However, narrowing the passband is equivalent to increasing the averaging or integrating time needed for the electronics to average out the noise. The longer the integration time needed, the slower will be the response of the amplifier to changing signal levels. For example, suppose that a 0.1 Hz passband is used to achieve a high *SNR*. This is equivalent to saying that one should average over the inverse of this frequency, or 10 seconds, in order to properly average out the unwanted noise signals at frequencies outside the 0.1 Hz passband. In that case we should not expect this system to be capable of measuring fluctuations in the signal being measured that happen in a period shorter than 10 seconds.

Even with unchopped systems, noise can still be removed from the signal through the use of low-pass filters that block the high-frequency components of the noise present in the unconditioned detector output. One must be careful that the low-pass filtering used has its transition frequency high enough to permit the valid signal to pass through. This is a trade-off found in the design and operation of most radiometers and photometers. For applications involving weak radiant flux,

a long integration or smoothing time is needed to average out the random noise superimposed on the signal being measured. This slows the response of the detection system. Flux signals varying faster than the smoothing circuit can handle will not properly be measured.

If one is designing a scanning spectroradiometer, for example, where the signal from the source through a scanning monochromator varies strongly with wavelength, the wavelength scanning speed must be limited in low-signal regions to give the signal handling electronics a chance to smooth out the noisy signal from the detector while still responding to the valid changes in flux resulting from the changing wavelength selected by the monochromator. As signal strength increases, it is possible to correspondingly increase wavelength scanning speed.

In applications where there is no choice but to attempt measurements of rapidly varying signals, the signal level must be increased or the detector characteristics chosen so that an adequate signal-to-noise level is achieved with the correspondingly short integration or averaging times needed to restrict noise levels. If choosing a better detector is not an option, then one is left with a need to gather more flux, through larger aperture flux-gathering optics or stronger sources of radiation.

The temporal frequency response and noise output over a specified frequency range are important characteristics of radiation detectors. They are described in the next section. In preparation for that discussion, let us suppose that the total noise produced by a source/detector/electronics system can be represented precisely by the function $N(t)$ over a measurement interval of length T . Noise can be quantified by the variance σ^2 of the function $N(t)$ over time T associated with the standard deviation σ , defined to be

$$\sigma = \sqrt{\frac{1}{T} \int_0^T [N(t) - \langle N \rangle]^2 dt} \quad (7.17)$$

Here $\langle N \rangle$ is the average or mean value of $N(t)$ over the complete time period T . The quantity σ is also known as the root mean square noise or rms noise. If $N(t)$ is a voltage, then σ will have units of volts. If $N(t)$ is current, then the units of σ will be amperes. It is important to note that the determination of σ requires a finite time period, it is not an instantaneous variable.

7.6 CHARACTERIZATION OF DETECTOR PERFORMANCE

The performances of detectors are generally described by means of various figures of merit, which manufacturers of detection devices determine and quote in their product literature.

7.6.1 Responsivity, R

The first and most important of these is the *responsivity* R of the detector to input flux. In general, responsivity is a function of the wavelength of incident radiation and the frequency f of modulation of the received signal: $R(\lambda, f)$. If the responsivity is stated in terms of the incident spectral flux $\Phi_\lambda(f)$, chopped at the frequency f , and the output voltage is $V(\lambda, f)$ at this frequency and wavelength, then the spectral responsivity is defined by

$$R(\lambda, f) \equiv \frac{V(\lambda, f)}{\Phi_\lambda(f)} \quad (7.18)$$

The total responsivity of a detector is defined by presuming that the detector is irradiated with a standard spectral irradiance distribution $E_\lambda(f)$, chopped and detected at the frequency f :

$$R(f) = \frac{\int E_\lambda(f) R(\lambda, f) d\lambda}{\int E_\lambda(f) d\lambda} \quad (7.19)$$

A common standard spectral irradiance distribution is that of a blackbody at a given temperature $E_{bb\lambda}(f, T)$, where T is the temperature. The responsivity can then be designated $R_{bb}(f, T)$.

For detectors whose outputs are not linearly proportional to the incident flux, i.e., nonlinear response detectors, the spectral responsivity is a function of the incident flux, $R(\lambda, f, \Phi_\lambda)$, as is the total responsivity $R(f, \Phi)$. Budde lists and describes a number of additional factors affecting responsivity, including bias effects, load resistor effects, temperature effects, polarization effects, magnetic effects, nonuniformity over the detector surface, and fatigue and aging [20].

7.6.2 Quantum Efficiency, η

The responsivity of a detector to discrete photon flux is called the quantum efficiency of the detector. The CIE vocabulary defines quantum efficiency as “the ratio of the number of elementary events (such as the release of an electron) contributing to the detector output to the number of incident photons.” The output events are either the emission of a photoelectron from a photocathode or the

generation of an electron-hole pair in a semiconductor. For a complete photoelectric detector, the quantum efficiency can be expressed as

$$\eta(\lambda) = \frac{R(\lambda)}{\lambda} \frac{hc}{q} \quad (7.20)$$

Here h is Planck's constant, c is the speed of the electromagnetic radiation in a vacuum, and q is the charge of the electron. Two versions of this definition are noted: *external* quantum efficiency refers to the case where the detector input is an *incident* photon and *internal* refers to the case where the input is an *absorbed* photon. Substituting values for h , c , and q in (7.20) yields

$$\eta(\lambda) = 1.2398 \frac{R(\lambda)}{\lambda} \quad (7.21)$$

if λ is in μm and $R(\lambda)$ is flux responsivity in amperes per watt.

7.6.3 Noise Equivalent Power, NEP

The second most important figure of merit for a detector is a measure of the noise level it produces in the absence of incident flux. This so-called *noise equivalent power* is defined as the equivalent flux incident on the detector that would produce an output signal equal to the rms noise (only) output σ for operation at a stated frequency and over some stated bandwidth Δf . Putting it another way, it is the incident signal flux level that would produce an *SNR* of 1.0 when signal oscillations at frequency f are compared with noise oscillations at that same frequency. NEP is therefore given by

$$NEP = \frac{\sigma}{R} \quad (7.22)$$

where σ is the rms noise current or voltage and R is the current or voltage responsivity. A spectral version $NEP(\lambda, f)$ of the NEP can also be defined as the monochromatic radiant flux Φ_λ at wavelength λ needed to produce an *SNR* of 1.0 at the electronic (chopping) frequency f .

There can be a different NEP for different radiation spectral distributions, for different frequency bandwidths, and for different modulation frequencies. Usually some standard conditions for these variables are quoted or implied by manufacturers in product literature. However, it is important to know the specific conditions used to obtain a given value of R or NEP .

The NEP is as important in the design of most measurement systems using detectors as is the responsivity because it indicates a lower limit on the flux level that can be measured (with an $SNR \geq 1.0$) at a given modulation frequency and passband under stated conditions.

If one doubles the area of a given detector, twice as much flux can be intercepted by it, all other things being equal. Thus, the responsivity of most detectors increases with their area. The NEP does not in general remain constant, however, and the responsivity increase may not be linear since other factors in the practical design of detectors are involved. It is best to determine the specific operating characteristics of different detectors being considered for use in a specific system being designed.

7.6.4 Detectivity, D

Detectivity D is the reciprocal of the NEP . The higher the detectivity, the smaller the signal the detector can measure, and the better the overall performance.

$$D = \frac{1}{NEP} \quad (7.23)$$

More useful is the quantity called *normalized detectivity*, D^* , or “D-star,” which is the detectivity normalized for detector area A_d and bandwidth Δf :

$$D^* = \frac{\sqrt{A_d \Delta f}}{NEP} \quad (7.24)$$

This definition normalizes the detectivity to a 1-Hz bandwidth and a unit detector area, usually 1 cm^2 . The units of D^* are $\text{cm} \cdot \text{Hz}^{1/2} \cdot \text{W}^{-1}$. One can think of D^* as the SNR at the output of a unit area detector when receiving 1W of radiant power.

The blackbody D_{bb}^* is defined in terms of the spectrally normalized spectral detectivity D_λ^* as follows:

$$D_{bb}^*(T, f) = \frac{\int_0^\infty D_\lambda^*(f) E_{bb\lambda}(T) d\lambda}{\int_0^\infty E_{bb\lambda}(T) d\lambda} \quad (7.25)$$

It provides the weighted average of the spectral detectivity when blackbody radiation at temperature T is used as the weighting function.

7.6.5 Photon Noise-Limited Performance

The ideal detector would have zero noise. However, even if the noise generating processes associated with the detector itself were zero, there would still be the radiation noise, described previously, that is associated with the random generation of photons from the source producing the radiation being detected. It is possible in some cases to reduce the detector-generated and the post-detector electronic noise to levels below the photon noise. Such cases are termed *photon noise-limited* detection, since it is the incident photon noise that limits the smallest signal that can be detected with an adequate *SNR*. Actually, the random arrival of photons at the detector can be considered the signal to be measured, not noise. In this case, it might be better to speak of *photon fluctuation-limited* detection. The previous terminology is so widespread, however, that it is likely to be continued.

PMTs and microchannel detectors can achieve photon noise-limited performance over the spectral range from 0.4 to 1.1 μm . Photoconductive, photovoltaic, and bolometric detectors can achieve it in the thermal infrared, i.e., where λ is greater than 3 μm . In many cases, especially with infrared detectors, cooling of the detector is required to achieve photon-limited performance.

One way to look at the dependence of detector operation on temperature is through the photovoltaic (PV) detector model. For the PV effect to work, the energy bandgap E_g must be low enough so that the photons of energy $h\nu = hc/\lambda$ can excite the photoelectrons across the gap for the range of wavelengths being detected. The greater the wavelength, the lower the bandgap must be. This means we must have

$$\lambda_{\max} = \frac{C}{E_g} \quad (7.26)$$

for some constant C . On the other hand, the temperature T of the detector must be low enough that electrons are not thermally excited across the bandgap, increasing semiconductor conductivity without photon absorption. To meet this restriction, one must have

$$\frac{kT}{q_e} \ll E_g \quad (7.27)$$

where k is Boltzmann's constant ($1.38 \times 10^{-23} \text{ W}\cdot\text{s}\cdot\text{K}^{-1}$) and q_e is the charge of the electron. Solving (7.26) for E_g and substituting the result into (7.27) yields

$$\frac{kT}{q_e} \ll \frac{C}{\lambda_{\max}} \quad (7.28)$$

which shows explicitly the need for cooling of detectors for optimum operation at long-wavelengths.

7.7 FLUX CONDITIONING PRIOR TO THE DETECTOR

“Bare” detectors are seldom used. Usually they are enclosed in a housing with a window or preceded by an optical system or subsystem. One reason is to protect the sensitive surfaces of the detectors from damage and another is to limit the incident flux to a narrow range of angles of incidence or to otherwise modify the angular distribution of flux incident upon them. Variations in detector responsivity with incidence angle generally depart significantly from Lambert’s cosine law. Accordingly, either the angular response needs to be corrected or the range of incident angles approaching the detector needs to be restricted.

Values published by manufacturers for the responsivity and other characteristics of their detectors are approximate and can lead to significant errors if used for calibration purposes. For this and other reasons, detector performance figures of merit are used mainly for the selection of a detector with the proper characteristics for a given application—not for calibration purposes.

It is important to note that the smaller the detector, the lower the noise level produced. There is therefore usually a noise penalty to pay for using a detector with a sensitive surface significantly larger than the beam incident upon it. The unused area contributes to both the dark current and to the noise but not to the signal. The *SNR* of a detector can therefore be improved by using a detector only as large as is needed to match the beam of flux placed on the detector by the conditioning optics. Conversely, the optics can be designed to optimize the *SNR* for a given detector.

Once an appropriate detector has been selected, it will generally be placed in an optical-mechanical system that will have the effect of conditioning the flux prior to its receipt by the detector. This conditioning can consist of chopping, as described in Section 7.5, focusing incident flux into a narrow conical solid angular range of angles, and spectral filtering, as in the case of PV photodiodes used to measure luminous flux. In the latter case, a specially designed spectral filter is placed over the detector to correct its spectral sensitivity so that the overall spectral sensitivity approximately matches the shape of the CIE photopic spectral luminous efficiency function, $V-\lambda$, described in Chapter 2. A discussion of detector/filter combinations to match desired spectral response shapes is provided in Sections 7.7.2 and 7.7.3.

7.7.1 Cosine Response Correction

One special method of angular response conditioning deserves mentioning here. An ideal detector that can measure flux in a narrow beam of any angle of incidence as well as flux distributed over a hemispherical solid angle, must have excellent cosine response. Its response to constant and uniform collimated incident irradiance fully covering the detector aperture should decrease with the cosine of the angle of incidence. Such behavior is called "good cosine correction." It is called correction because, as we saw in Chapter 6, the reflectance of most materials increases with angle of incidence. Correspondingly, the absorptance of opaque materials generally varies with angle of incidence. Thus the sensitive surfaces of detectors, whether they be semiconductors or photon-emitting metals, exhibit angular sensitivity in their absorptance. It is only the absorbed radiation that can produce a signal from the detector. Such angular selectivity of absorptance causes the detector output to depart from Lambert's cosine law, and the detector must be corrected back to the proper angular response if it is to work properly as an irradiance or illuminance detector. Furthermore, the housings of many detectors shade their sensitive surfaces at some angles of incidence, again calling for some means of angular response correction.

A common method for correcting for cosine response errors is to cover the detector with a sheet of milk-white, highly diffusing semitransparent material having good (and ideally constant) conical-hemispherical spectral transmittance over the spectral range of good detector sensitivity. The idea is that no matter how the incident radiation falls on this sheet, a fixed and constant fraction of it will be delivered to the detector over a fixed and constant range of angles. In practice, no diffusing sheet has been found that satisfies this ideal.

What is done in practice is to experiment with a variety of diffusing materials, surface roughnesses, and geometrical configurations until a combination is found that provides reasonably good cosine correction. As one might expect, the greatest difficulty is in achieving good cosine response for angles of incidence greater than about 60 deg. The surface of the diffusing material, if it is polished, will have a reflectance that has approximately the shape shown in Figure 6.3. Even if the surface is roughened, the reflectance will still increase significantly at higher angles of incidence, causing the transmitted radiation to drop and the detector to under-measure the radiation at these angles.

An ingenious solution to this problem is to limit the size of the diffusing sheet and to allow it to extend above the detector housing so that at large angles of incidence some of the incident flux will be received by the *edge* of the sheet, this edge being perpendicular to the front surface. Thus, as more and more flux is reflected from the front of the sheet, more and more will be transmitted through its edge. At an 80-deg angle of incidence, for example, little flux will enter through

the front face of the diffuser, but quite a bit will enter through the edge because the angles of incidence here will be close to 0 deg.

One problem remains, however. True cosine response drops to zero at 90 deg, whereas the exposed edge of the diffuser receives considerable quantities of flux at this angle. The cosine corrector must be designed so that no flux can reach the detector for angles of incidence at and greater than 90 deg. The usual solution to this requirement is illustrated generically in Figure 7.18.

Notice that as the angle of incidence increases, more and more flux reaches the edge of the detector until the angle of incidence approaches 90 deg, at which point the shading ring begins to shade the edge. Finally, at 90 deg, the diffuser is shaded completely and no flux can reach it. The design of the specific dimensions of this cosine correction scheme depends strongly on the biconical optical properties of the diffusing sheet, the geometrical placement of the detector below it, and the angular response characteristics of the detector itself. Finding the right geometry is often a hit-or-miss proposition. Even if a good design is found, the quality of the corrected cosine response can suffer if the properties of the diffusing material change in time, from batch to batch of manufacture, or with the wavelength of incident radiation. Making a good cosine corrector is one of the most difficult problems in the manufacture of good quality, accurate irradiance and illuminance

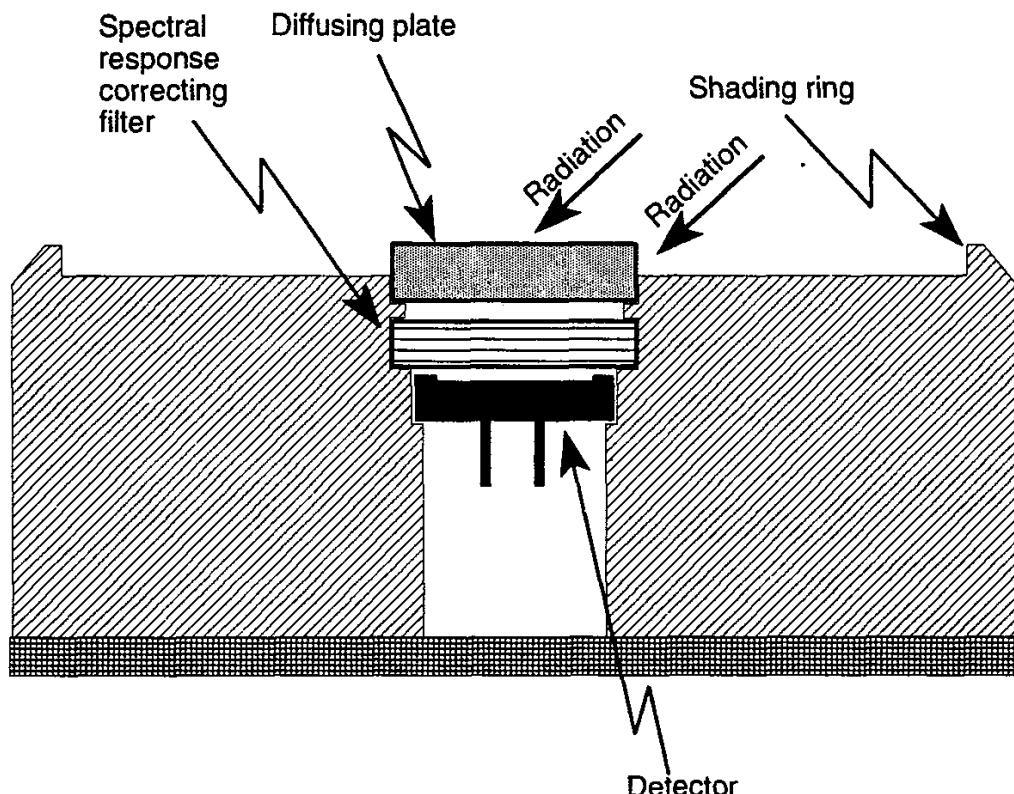


Figure 7.18 Diagram of the sensing element of an illuminance meter with both cosine-response correction and a V-lambda spectral matching filter.

meters intended for measurements of radiation received from within a hemispherical solid angle.

A way of providing better cosine correction than the one diagrammed in Figure 7.18 is through the use of an integrating sphere, also called an *Ulbricht sphere* in German-speaking countries [21,22]. An arrangement for utilizing the desirable properties of the integrating sphere is illustrated in Figure 7.19. The approach is based on the following idealized principle. Flux entering a small port in a hollow sphere whose interior surface is coated with a material of extremely high diffuse reflectance will be multiply reflected (scattered) an infinite number of times, in all directions, with little loss on each reflection. If a small hole is placed in the side of the sphere and shielded from flux that has not been reflected at least once by the sphere, the flux emerging from this hole will be a fixed and constant fraction of the flux entering the other hole, regardless how the flux entering the input port is distributed in angle.

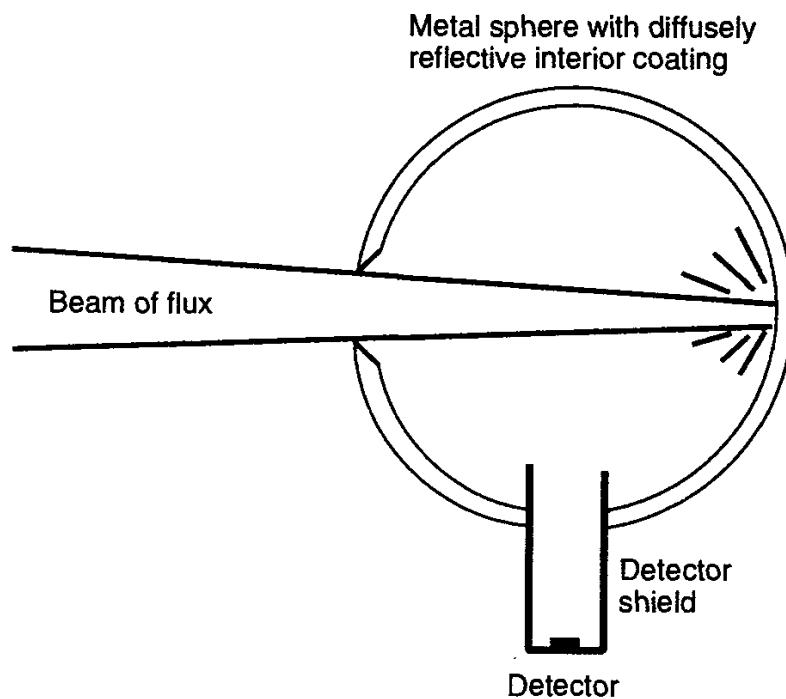


Figure 7.19 Cosine response correction with an integrating sphere.

Real integrating spheres, with reflectances less than 1.0 and with entrance and exit ports of finite areas cannot achieve this idealized performance. They can be made to approach it closely, but they are not efficient at delivering flux to the detector for measurement. An excellent treatment of the theory of the integrating sphere was offered by Goebel [23]. The relevant optical properties of a high-reflectance coating (PTFE powder) for such spheres have been reported by Weidner

and Hsia [24]. The hemispherical reflectance of this material exceeds 0.990 and is nearly constant over the wavelength range from 350 to 1,800 nm. Integrating spheres are discussed further in Section 8.11.

7.7.2 Photopic Correction

Also shown in Figure 7.18 is a V-lambda matching spectral response correcting filter. The goal in the design of this filter is to alter the spectral responsivity of the detector so that its shape is the same as that of the V-lambda curve shown in Figure 2.1 and Table 2.1. For optimum spectral matching, one or more glass filter materials are chosen so that the product of their spectral transmittances $T_1(\lambda)$ and $T_2(\lambda)$ and the spectral responsivity $R(\lambda)$ of the detector, less $V(\lambda)$, never exceeds some desired maximum value for all wavelengths from 380 to 770 nm. Once the proper filter thicknesses are determined by trial and error, the filters are cut to size, polished, and sandwiched together with a nonspectrally selective adhesive. (They can also be air-spaced—mounted together without adhesive.) Since changing the thicknesses of the filters exerts only a rather broad spectral influence, bringing the corrected spectral response within the desired tolerance at one wavelength may put it out of tolerance at another. A better approach might be to minimize the rms deviation Q over the complete wavelength range. The quantity Q to be minimized, for the case of two filters, is

$$Q = \sqrt{\int_{380\text{nm}}^{780\text{nm}} [T_1(\lambda) T_2(\lambda) R(\lambda) - V(\lambda)]^2 d\lambda} \quad (7.29)$$

Budde describes the use of this process to obtain the V-lambda correction of an EG&G UV44A silicon photodiode through the use of four filters [25]: Schott filters BG38 (3.38 mm), GG9 (3.23 mm), and KG3 (4.00 mm) and Corning filter 3307 (1.13 mm).

7.7.3 Spectral Filtering

The spectral filtering approach described in the previous section can be generalized to other desired detector/filter combination spectral response shapes. One of the most common of these applications is the use of filtering to “flatten” the spectral response of a solid-state detector over some portion of its wavelength range. Suppose a source of radiation to be measured has significant output flux only over the wavelength range from 450 to 770 nm. Suppose also that the spectral shape of the flux signal (or signals) to be measured is known to vary from time to time. One

would like to use a photovoltaic diode detector to measure this radiation, and to find a filter that can flatten the detector response over the range from 450 to 770 nm. Since it is known that there is no significant radiation outside this wavelength range, it does not matter what the detector/filter combination's response is outside this range. All that is needed is to make this response constant, or nearly so, from 450 to 770 nm.

The methods of Section 6.4 can be used to determine the spectral transmittance at normal incidence of various candidate spectral filters as a function of filter thickness. Sandwiching two or more filters together, each one having a different spectral transmittance, can result in a product transmittance that renders the sensitivity of the detector constant over the required spectral range. It is desired to determine the thicknesses t_i of the i th filter that best achieves the desired spectral response.

A modification of the approach used in Section 6.5 can be used to determine the spectral transmittance $T_N(t_1, t_2, \dots, t_N, \lambda)$ of a set of N filters, where the optical properties of the adjacent filters in contact with the i th filter are known. The first part of the modification is to use the refractive indices n_{i-1} and n_{i+1} for the filters in contact with either side of the i th one in (6.10) and (6.14) to obtain the reflectance ρ_i and transmittance τ_i at the i th interface. The second part of the modification is to use (6.47) to determine the transmittance T_1 of the first filter (into the second one) and the reflectance R_1 from the second filter back into the first, including all orders of multiple reflection in the first filter.

Next, one treats the first filter as if it were an interface having transmittance T_1 and reflectance R_1 . Equations (6.47) are then applied to the second filter, with the proper substitutions for the reflectances and transmittances at the interfaces. This process yields the transmittance $T_2(t_1, t_2, \lambda)$ of the first two filters in contact. These two filters are then treated as a single filter in contact with the third one, and the transmittance of the three is determined: $T_3(t_1, t_2, t_3, \lambda)$. The process is repeated until the transmittance of the complete N -filter stack is determined, $T_N(t_1, t_2, \dots, t_N, \lambda)$ as a function of the thicknesses of all the component layers.

If $S_w(\lambda)$ is the desired or wanted spectral response and $S_d(\lambda)$ is the real spectral response of the detector, then the goal is to minimize the function $Q(t_1, t_2, t_3, \dots, t_N)$ with respect to variations in the thicknesses t_i of the N filters making up the sandwiched combination of filters, where

$$Q(t_1, t_2, t_3, \dots, t_N) = \sqrt{\int_{\lambda_1}^{\lambda_2} [T_N(t_1, t_2, t_3, \dots, t_N, \lambda) S_d(\lambda) - S_w(\lambda)]^2 d\lambda} \quad (7.30)$$

with $[\lambda_1, \lambda_2]$ being the wavelength interval over which the spectral matching is to be achieved.

Various schemes are available for minimizing this function. Since real applications of this technique are likely to involve measured data on discrete intervals, the integral is often replaced by a sum. Since analytical relationships are unlikely to be available for the spectral variations in either $T_N(t_1, t_2, \dots, t_N, \lambda)$, $S_d(\lambda)$, nor $S_w(\lambda)$, the minimization of $Q(t_1, t_2, t_3, \dots, t_N)$ will most likely have to be a numerical process calculated on a high-speed computer.

At any wavelength, $T_N(t_1, t_2, \dots, t_N, \lambda)$ will be a monotonically changing function of each of the individual filter thicknesses t_i . A straightforward (but inelegant) strategy for minimizing $Q(t_1, t_2, t_3, \dots, t_N)$ would be to calculate $Q(t_1, t_2, t_3, \dots, t_N)$ for all combinations of a selected set of t_i values over a range of interest and to then select from the voluminous results that particular combination of thicknesses that results in the smallest value of $Q(t_1, t_2, t_3, \dots, t_N)$. More sophisticated techniques are available, but in this era of relatively inexpensive high-speed computers this approach should yield satisfactory results if N is no greater than about two or three.

The smallest value of $Q(t_1, t_2, t_3, \dots, t_N)$ selected at the end of this tedious process will be a single-number measure of the success of the matching of spectral responses. Q will be approximately proportional to the average of the mismatch between the two spectra. Dividing $Q(t_1, t_2, t_3, \dots, t_N)$ by the integral (or sum) of S_w over the same wavelength range will yield a dimensionless measure of the residual error in the spectral matching as a fraction of the desired sensitivity. Budde discusses a number of additional figures of merit that can be used for detector spectral matching [1].

There is another method that is used to alter a detector's inherent spectral response to match some desired spectral shape. Instead of placing two or more filters in series in front of the detector, two or more can be placed in parallel instead, side by side in front of either the detector or some other aperture in the optical system preceding the detector. This scheme is called *partial filtering*. The series method is sometimes called *subtractive filtering* and in consequence, the partial filtering method is sometimes called *additive filtering* [1]. With the latter method, the relative areas of the filters covering the aperture are adjusted to give the best match to the desired spectral shape for the response of the detector/filter combination. The advantage of this method is that there is no need to adjust the thickness of a filter layer, only the relative area of the aperture covered by it.

As might be expected, combinations of the two methods are also used, with both stacking and side-by-side layering of filters in a *filter mosaic* pattern. Discussions of the matching of detector/filter spectral responses to desired spectral shapes can be found in a book on color science by Wyszecki and Stiles [26] and in a report from the National Research Council of Canada [27].

This section would not be complete without a description of the novel filtering technique called *spectral masking*. One would like to be able to produce a filter of any spectral shape. A direct approach to the solution of this problem is to send

a beam of radiation into a spectrometer or spectrograph. The dispersive element of this device (described in Section 8.12) spreads the incident beam angularly, with the angle of spread depending upon the wavelength of the radiation. A full spectrum is therefore spread out on the exit plane of the spectrometer. By placing a mask over this exit plane, designed to block different wavelengths differently, and by then recombining the dispersed spectrum into a single beam again, almost any shape of spectral filtering can be achieved in principle.

An alternative to this is described in Chapter 9. A linear photodiode array is placed in the spectrum plane of a spectrograph and senses the whole spectrum essentially simultaneously. The relative weighting of different portions of the spectrum can be achieved electronically or by the computer software that analyzes the signals from the detector array.

7.8 SIGNAL CONDITIONING AFTER THE DETECTOR

The photocurrent or photovoltage produced by the detector itself is generally too small for most uses and is followed by an electronic amplification and/or measuring system. An exception might be relatively large-area semiconductor detectors with high responsivities whose output currents are sufficiently great to drive the coil movement in a sensitive current meter directly. The General Electric Lamp Marketing Department, for example, offers a "Light Meter Type 214" that measures incident illuminance (with some color and cosine correction) over three ranges, from 10 to 50, 50 to 250, and 250 to 1,000 FC, that requires no electrical input power to operate.

The electronic circuits to which a detector is connected further condition the signal output from the detector. Knowledge of the modifications to this signal imposed by the measuring circuit is important for accurate use of the detector signal.

There is great variety in the signal conditioning circuitries that are used. Their description is beyond the scope and intent of this chapter, but further information can be obtained from textbooks on electronics and from manufacturer literature. The Oriel Corporation's Volume II catalog, in particular, contains an extensive discussion of UV-VIS-IR detectors, detection systems, and output processors [28].

7.9 DETECTOR CALIBRATION

Chapter 9 contains a discussion of overall measuring instrument calibration. The calibration of detectors alone is not often of interest. Conceptually, however, the process is straightforward. One illuminates the detector with a known spectral distribution, ideally close to that of some standard spectral distribution, and then measures the output. The detector involved can be a bare detector such as a multiplier phototube or one with conditioning elements connected to it, such as a

PV photodiode with a spectrally selective filter and cosine correcting diffuser. More complicated detector/optics combinations are discussed in Chapter 9.

The varying spectral sensitivity of most detectors can make their spectral calibration difficult. Even “flat-response” thermal detectors exhibit some variation in their spectral sensitivity over their ranges of use and there are no known sources whose spectral intensity is constant with wavelength over a wide range. Three common approaches to detector calibration are in use. The first is to illuminate the detector with a standard source spectral distribution whose irradiance is known accurately at each wavelength, and to measure the integrated response of the detector to the flux incident upon it, calculated by multiplying the incident spectral irradiance by the sensitive area of the detector. The second is to perform a wavelength-by-wavelength measurement of the detector’s responsivity, using any arbitrary source and a scanning monochromator. In this case, the output of the monochromator at each wavelength must be known precisely. The spectral responsivity of the detector will be the ratio of the detector output to the known spectral radiant flux incident on it at each wavelength of the measurement. Third is to repeat this second experiment but instead of knowing precisely the monochromator output, a second, previously calibrated detector is substituted for the one being tested and the ratio of their outputs at each wavelength is determined. Since the absolute spectral responsivity of the reference detector is known, it is a simple matter to determine the spectral responsivity of the test detector from the ratio of the outputs of the two detectors.

The difficulty in the first case is in providing a real spectrum whose shape matches the desired standard spectral distribution precisely. Alternatively, one could use a spectrum that departs slightly but by known amounts from the standard spectrum, and correct the measured results to the standard spectrum. A high-quality blackbody radiation source at a known temperature qualifies as a standard source spectrum with which to calibrate a detector, but obtaining and operating such a source is not an easy task.

In order to use the second method, one needs a way of calibrating the monochromator’s output. If a previously calibrated source, such as an incandescent filament lamp, is used with a highly regulated power supply so that the current through the filament can be maintained at precisely the calibration level, then the spectral flux entering the monochromator will be known. However, the spectral transmittance of the monochromator at each wavelength of operation must also be known for this method to work. In practice this is difficult to come by. If the monochromator is a set of filters whose spectral transmittances are accurately known, then this method can be made to work, but the wavelengths of calibration are limited to those of the filter set.

The third method is the most common one. According to Budde, the thermopile is the most commonly used reference detector. It is made to be as non-spectrally selective as possible and assumed to be so over the spectral range of

calibration. However, some frequently used thermopiles have a definite selectivity, and extra care should be taken in using this approach. The detector comparison method is described at some length by Budde [29].

As described in Chapter 9, another calibration method has been developed recently. This method relies on the absolute calibration of certain PV diode detectors, independent of any source or other detector. This self-calibration, as it is called, can be used to provide reference detectors to which other unknown detectors can be compared for calibration purposes.

7.10 EXAMPLE DETECTORS AND THEIR CHARACTERISTICS

There are a number of manufacturers of detectors of optical radiation. A representative, but not exhaustive, list is provided in Appendix 7. A more complete listing can be found in Part 2 of the August 1993 issue of *Physics Today* [30].

Probably the primary parameter of interest in choosing detectors for an intended application is the spectral range of usable sensitivity. Next would be the relative sensitivity of different devices having spectral ranges overlapping the appli-

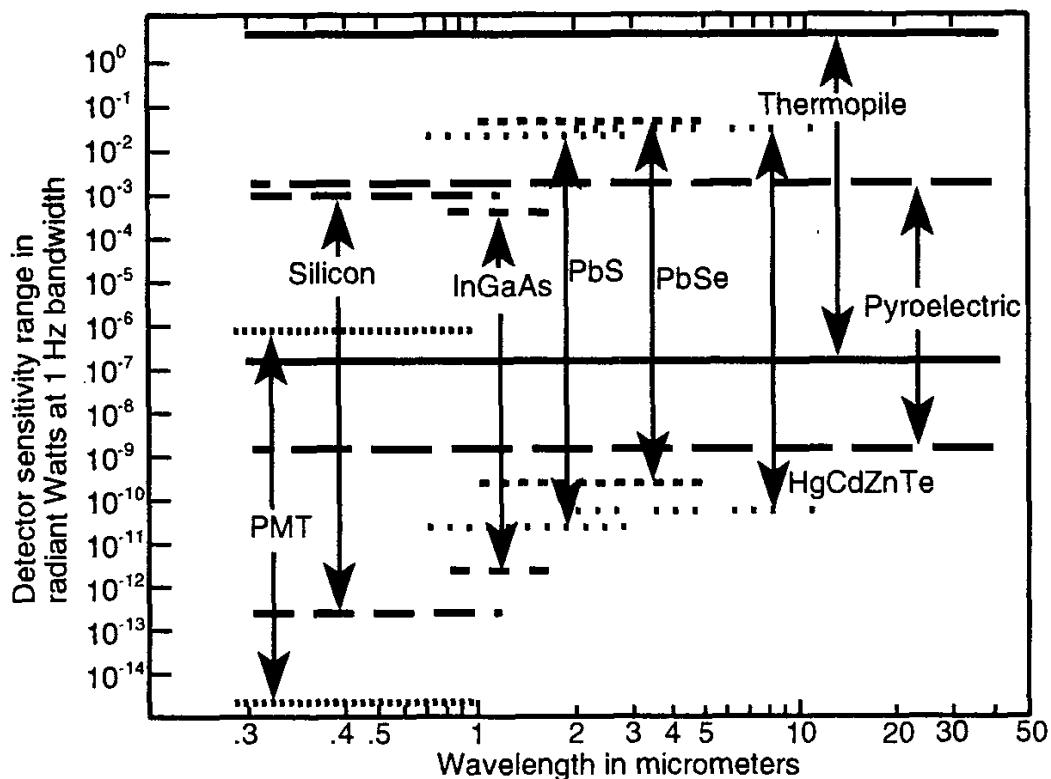


Figure 7.20 Ranges of usable spectral sensitivity for a variety of detector types. (Courtesy Oriel Corporation.)

cation's spectral range of interest. Figure 7.20 provides a graphic illustration of these two parameters for a variety of detector types.

The next parameter of interest would be the specific spectral responses of candidate detectors. Figure 7.21 shows the approximate generic spectral responses of a wide variety of detector types, using normalized detectivity D^* as the figure of merit.

The photocathode of vacuum photodiodes and photomultiplier tubes determines the spectral responses of these devices. A variety of photocathode spectral responses is plotted in Figure 7.22. Additional specialized photocathode materials are available. In the past a series of "S-numbers" were used to identify different photocathode materials having characteristic spectral responses. This scheme grew out of the existence of a limited number of efficient photocathode materials whose relative spectral responsivities were well known. The Electronics Industries Association registered several of these spectral responsivity functions according to their S-numbers. Budde presents a table giving the photosensitive materials, envelope materials, detector type, and notes for S-numbers 1 through 24. The relative spectral responsivities of six of these materials are plotted in Budde's Figure 5.1 and the relative spectral responsivities of five more photoemissive cathode materials not having S-number designations are plotted in his Figure 5.2. The S-number designation can still be found in the catalogs of some manufacturers, but many more photocathode materials and different formulations are now available and not all

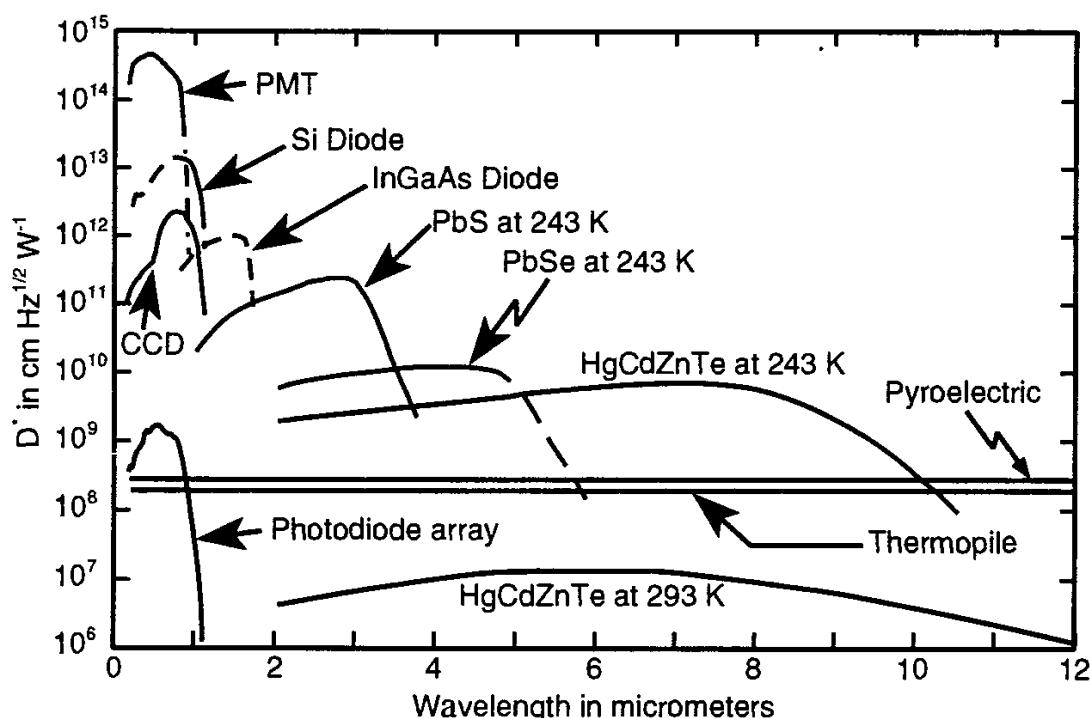


Figure 7.21 Normalized detectivity, D^* for a variety of detector types. (Courtesy Oriel Corporation.)

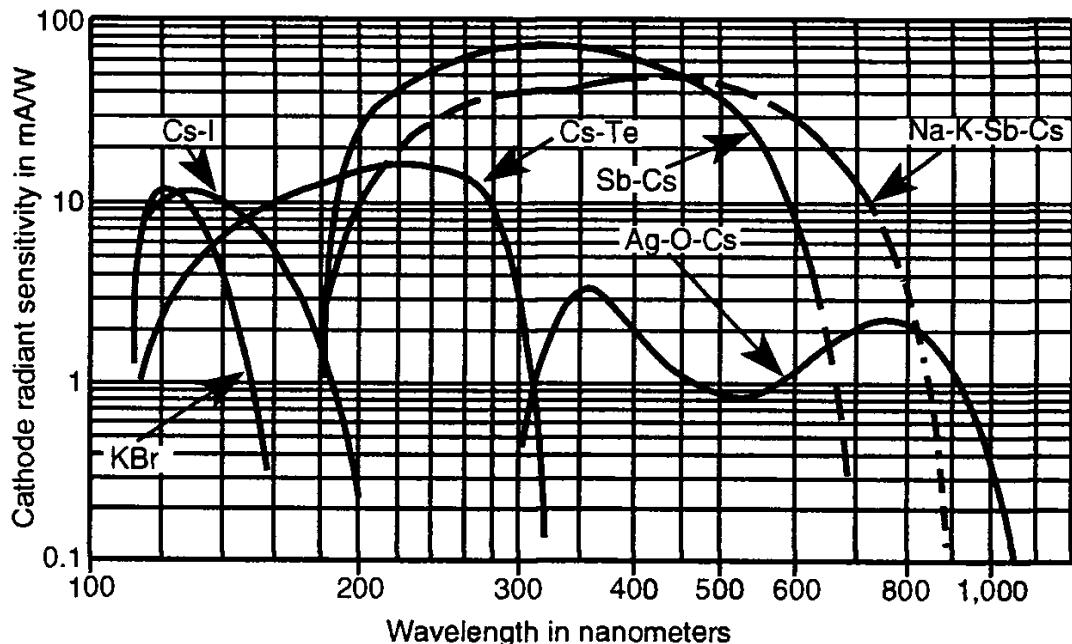


Figure 7.22 Photocathode spectral radiant sensitivities for a number of materials. (Courtesy Hamamatsu Corporation.)

manufacturers use the S-number designation. Specific PMT manufacturers can be consulted for additional information.

Several solid-state photodiode spectral response characteristics are plotted in Figure 7.23, in relation to the visible portion of the spectrum. The relative sensitivities a UV-enhanced silicon photovoltaic detector and a cooled InGaAs detector are plotted in Figure 7.24 showing continuous coverage by these detectors of the spectral range from 200 to 1,600 nm.

Once a variety of candidate detectors covering a spectral range of interest has been identified, their noise characteristics should be examined to properly match the detector to the radiant signal strengths anticipated for the application. Some plots of *NEP* for two different infrared detectors, PbS and PbSe, are shown in Figure 7.25. Not only does the minimum detectable signal vary with wavelength, but the temperature of the detector is also important.

Since pyroelectric detectors require a chopped input flux, their response to the chopping frequency is an additional parameter of importance in the selection of these detector types. As mentioned in Section 7.3.1, special pyroelectric detectors have been developed with the ability to respond to slowly varying infrared radiation. Detectors of this type can be very effective as proximity sensors, sensing the long-wavelength infrared radiation emitted by the warm bodies of mammals, including humans. Applications include intrusion alarm systems, automatic light switches, and a variety of other applications involving the detection of slow movement of

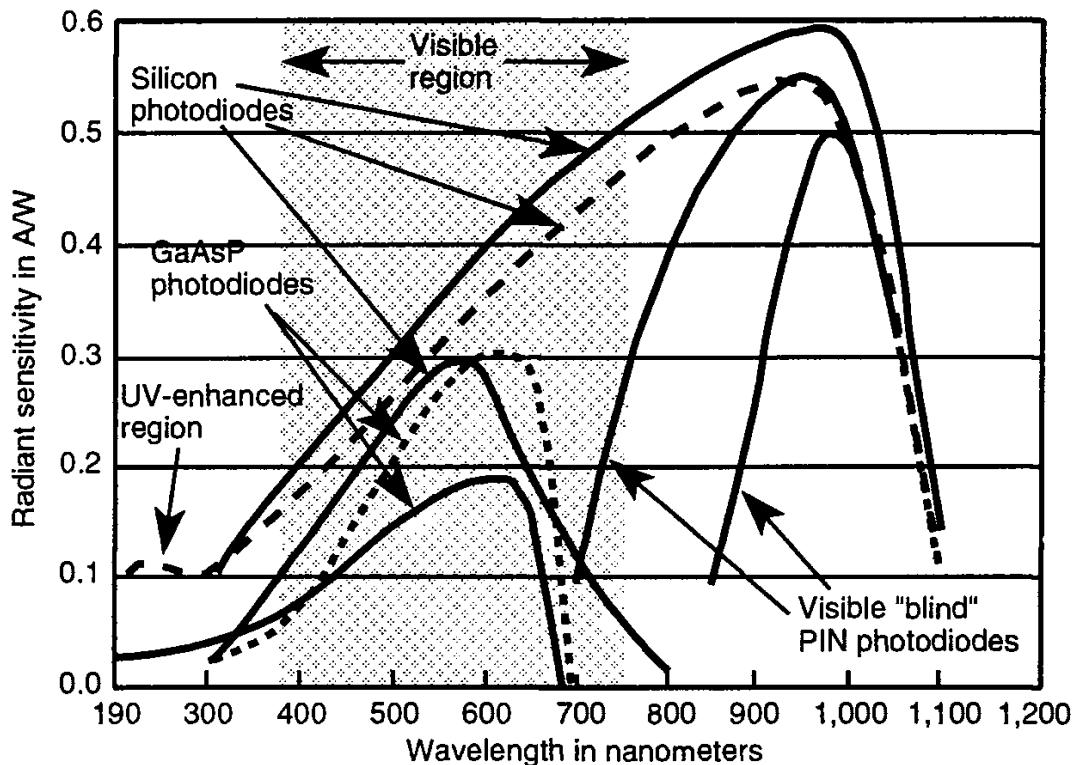


Figure 7.23 Spectral sensitivities of a variety of solid-state semiconductor detector types. (Courtesy Hamamatsu Corporation.)

an infrared source. A representative plot of the temporal response of such detectors is shown in Figure 7.26.

Finally, the wide range of linearity in the response of silicon photovoltaic diode detectors was mentioned in Section 7.3.3.3. A plot of output current versus input radiant flux for such a device is shown in Figure 7.27. It can be seen that excellent linearity is preserved over most of the range plotted. Departure from the linear range occurs only at the upper end, and this departure is seen to depend upon the effective load resistance “seen” by the photodiode. The lower the impedance (the closer to true “short circuit” operation), the greater the linearity at the upper levels. The linearity exhibited in Figure 7.27 extends nearly two orders of magnitude below the range plotted in the figure, but in this lower range the operational limit of the detector is determined by the *NEP* of the device or the noise limitations of the external circuits to which it is attached.

In spite of the general value of silicon photovoltaic detectors over their spectral range of useful response, additional detector types continue to be developed with almost constantly improving performance, including extended spectral range of coverage. Linear and two-dimensional diode arrays are advancing as well, and are making possible new application areas. For example, mercury-cadmium-telluride photovoltaic as well as photoconductive arrays have been developed [31] that are

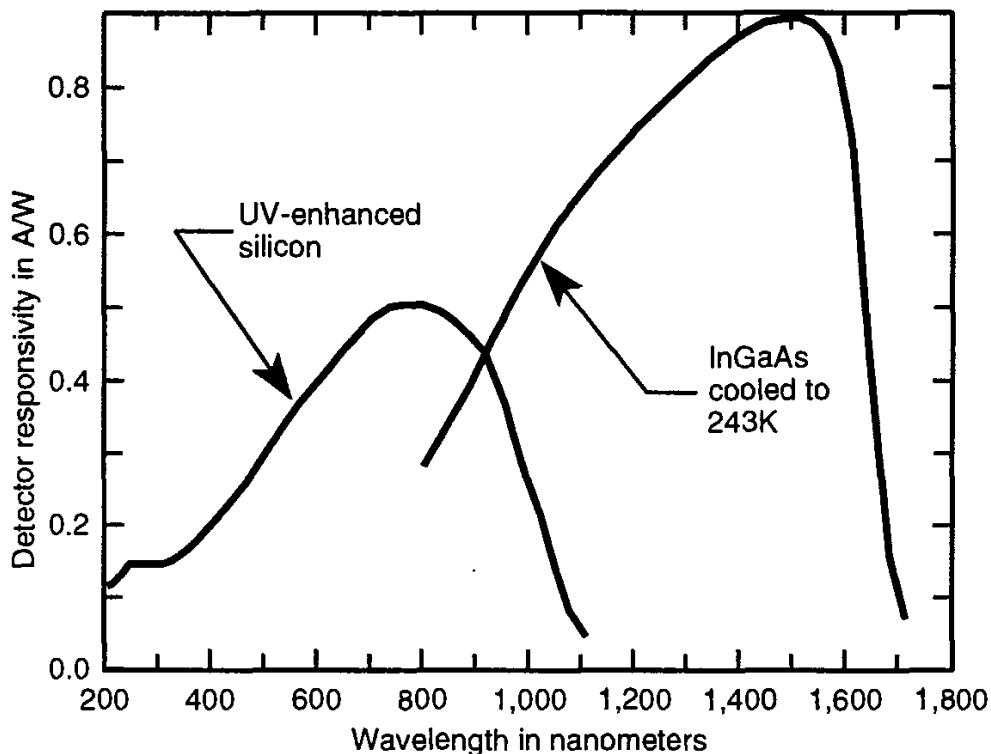


Figure 7.24 Typical spectral responses of UV-enhanced silicon photovoltaic and cooled InGaAs detectors. The latter is cooled to 243K. (Courtesy Oriel Corporation.)

suitable for use in the infrared out to 10 or even 13 μm , covering an important “window” in the earth’s atmosphere from 8 to 13 μm .

Another interesting detector can be mentioned. It is the misnamed “two-color” detector that takes advantage of the transparency of silicon photovoltaic diodes at wavelengths above about 1.1 μm to place another detector below the silicon one in the same housing. The lower detector has historically been PbS, but others using Ge or InGaAs are also available that provide coverage of the spectral range from 0.4 to 1.6 or 1.8 μm , with a fairly narrow gap between the spectral ranges of the two detectors at about 1 μm .

One can also purchase detectors with pre-amplifiers built into the detector housing, for reduced noise operation, detectors with thermo-electric temperature stabilizing elements built into the housing, and detectors with small narrow band-pass interference filters at preselected wavelengths also integrated into the detector package [32].

Finally, a new type of detector has recently been developed [33]. Called the “silicon carbide UV photodiode,” it is based on SiC and has a wide 3-eV bandgap, making it unresponsive to infrared wavelengths and giving it very low responsivity for wavelengths above about 420 nm. The range of useful spectral response is from below 200 to the lower edge of the visible portion of the spectrum near 380 nm,

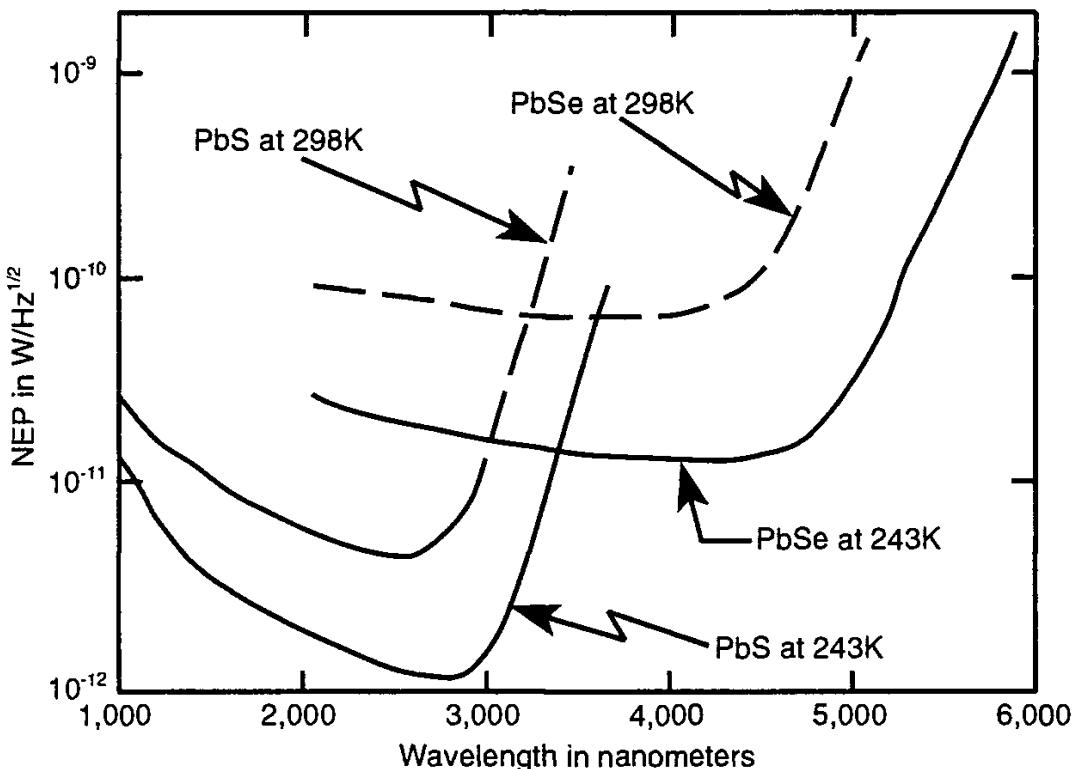


Figure 7.25 Plot of the wavelength dependences of noise equivalent power (NEP) for lead sulfide and lead selenide detectors at two different temperatures. (Courtesy Oriel Corporation.)

covering both UV-A and UV-B portions of the spectrum. The new detector boasts a dark current of 2 femtoamperes and a short-circuit responsivity of just under 200 mA/W [34].

EXAMPLE PROBLEM 7.1

Problem: Your company is hired by the U.S. Secret Service to design a system to detect intruders around the edges of the White House lawn as well as the fenced in space around some other federal buildings in Washington, DC. You are asked to design a solid-state laser source and detector combination that will provide a signal whenever the beam from the source to the detector is interrupted by an object. The system must be able to operate outdoors, in full daylight and in the presence of fog, over a distance of 100 meters.

Solution: Before doing the final system design, you desire to make a “breadboard” mockup of a system to see if you can get such a system operating quickly. You find in an Edmund Scientific Company catalog a solid-state laser having a 3.0 mW output at 670 nm and a beam diameter of 5 mm with a divergence of less than 0.5 mrad. Your idea is to select a detector with reasonably good sensitivity

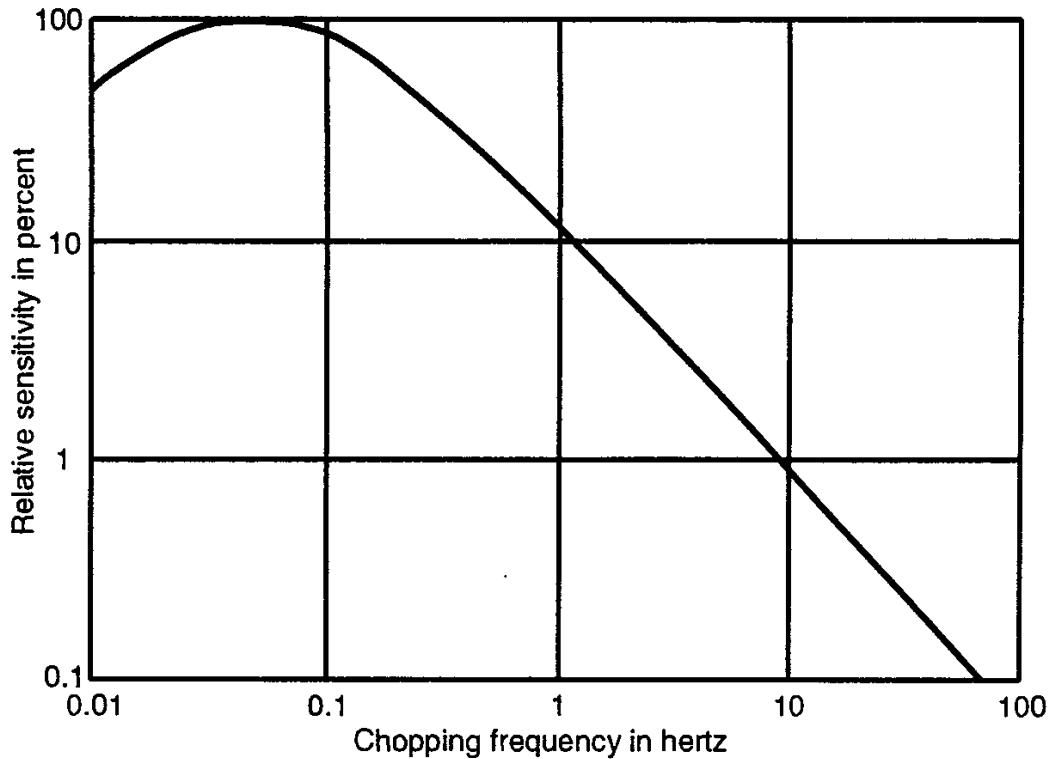


Figure 7.26 Plot of the relative sensitivity of a pyroelectric motion sensing detector versus signal frequency. (Courtesy Hamamatsu Corporation.)

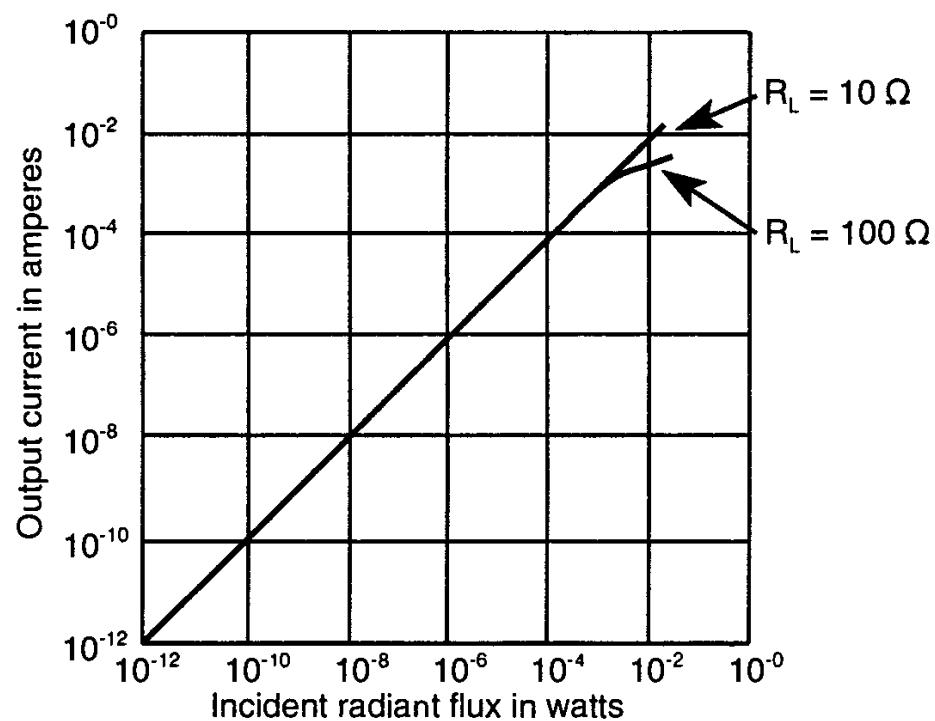


Figure 7.27 Range of linearity for silicon photovoltaic diode detectors. (Courtesy Oriel Corporation.)

at this wavelength and cover it with a narrowband interference filter to exclude flux from wavelengths much different from 670 nm. From the Oriel filter catalog you find an interference filter centered at 670 nm with a peak transmittance at this wavelength of 50% and a bandwidth of 10 nm at half maximum. From Figure 7.23 you see that a Si photodiode detector can be found with a sensitivity of from 0.4 to 0.45 A/W. You find one in a catalog having a sensitive area 10 mm in diameter and a responsivity at 670 nm of 0.42 A/W. In 100 meters you can expect the 5-mm diameter laser beam to expand to the following diameter:

$$\text{Diameter} = 5 + 0.5 \times 10^{-3} \times 100 \times 10^3 = 55 \text{ mm}$$

This means that the diameter of the beam at the detector 100m away will be 5.5 to 6 cm in diameter, 5.5 to 6 times the size of the detector. If the signal strength is adequate, this overfilling of the detector will allow for some misalignment of the laser beam, a desirable feature. Assuming that the 3 mW of laser power is insignificantly attenuated by the atmosphere and that the beam irradiance will be uniformly distributed over a 6-cm diameter (3-cm radius) area, you calculate that the beam irradiance will be approximately

$$\frac{0.003 \text{ W}}{\pi 0.03^2 \text{ m}^2} = 1.06 \text{ W/m}^2$$

The area of the detector is $\pi 0.005^2 \text{ m}^2 = 7.9 \times 10^{-5} \text{ m}^2$ so that without filtering, the flux received by the detector will be approximately $1.06 \times 7.9 \times 10^{-5} \text{ W} = 8.4 \times 10^{-2} \text{ mW} = 0.084 \text{ mW}$. After passing through the filter, this will drop to approximately 0.042 mW. Multiplying this by the 0.45 A/W responsivity of the detector yields a current output from the photodiode of 35 mA.

From Figure 7.27, it is clear that signals of this magnitude are in the upper range of detector output, well above the noise levels encountered in such detectors, so no special noise suppression circuits will be needed. This strong signal also allows for significant attenuation by the occasional fog that might be present. From Figure 7.27, you decide to set the threshold for the alarm system to $3.5 \times 10^{-9} \text{ A}$, a full seven orders of magnitude below the normal clear atmosphere operation level of the detector and three orders of magnitude above the noise regime of the detector. This setting will allow for significant attenuation by the fog but still a signal well above the inherent noise level of the detector.

To provide additional attenuation of daylight besides the narrowband filter, you decide to place the detector inside a short baffled tube to shield it from ambient daylight and also to surround the laser source with a black surface shaded from daylight. There is still one problem, however. A person with highly reflecting clothing might attempt to walk through the beam. Sunlight reflected from the

clothing might provide a high enough signal to prevent the alarm from going off. To estimate the magnitude of this potential problem, you assume that the clothing has a diffuse reflectance of 100% at 670 nm and that in a worst case the entire field of view of the detector will be filled by the clothing fully illuminated by the sunlight.

Looking at Figure 3.8, you see that at 670 nm, solar direct beam irradiance is about $1.2 \text{ W} \cdot \text{m}^{-2} \cdot \text{nm}^{-1}$. After reflection from the clothing, assuming a perfectly Lambertian surface, from (1.19) the spectral radiance L_λ of the clothing at 670 nm will be $1.2/\pi = 0.38 \text{ W} \cdot \text{m}^{-2} \cdot \text{nm}^{-1} \cdot \text{sr}^{-1}$. Assuming that the geometry of the baffle tube enclosing the detector permits a field of view of only 5 deg, (1.52) gives a solid angle of 6 msr. The spectral irradiance reaching the filter/detector combination will therefore be given by (4.11) with the spectral radiance L_λ assumed to be constant over the solid angle range of integration. This computation gives an incident spectral irradiance of $0.38 \times 6 \times 10^{-3} = 2.28 \text{ mW} \cdot \text{m}^{-2} \cdot \text{nm}^{-1}$.

The spectral bandwidth of the interference filter is about 10 nm. Assuming that the overall transmittance of this filter is approximately the same as if its transmittance was half (0.5) its peak transmittance (0.5) over this 10-nm bandwidth and zero outside, the filter would transmit approximately $0.5 \times 0.5 \times 10 \text{ nm} \times 2.28 \text{ mW/m}^2 \cdot \text{nm}^{-1} = 5.7 \text{ mW/m}^2$ of direct beam sunlight to the detector. Multiplying this by the area $7.9 \times 10^{-5} \text{ m}^2$ of the detector yields an incident flux level of $0.45 \mu\text{W}$. Multiplying by the detector's 0.45 A/W responsivity yields a signal level of $0.2 \mu\text{A}$. This is significantly higher than the alarm threshold value of $3.5 \times 10^{-9}\text{A}$.

From this result you decide to modulate the laser source electronically at a frequency of 100 Hz and to use a detector post-processing circuit that responds only to signals modulated at this frequency, ± 10 Hz. Any interruption in this "modulation coded" signal, even if by a constant bright source with significant flux at 670 nm, will trigger the alarm. You complete your breadboard mockup of this system and even attempt to foil the alarm by using a mirror to direct sunlight into the beam, saturating the detector. However, this beam is unmodulated and the modulated laser beam is interrupted so the alarm goes off. Your boss is impressed by your demonstration of the system, congratulates you on a job well done, and directs you to turn the project over to the production engineering team.

REFERENCES

- [1] Budde, W. *Optical Radiation Measurements*, Vol. 4, *Physical Detectors of Optical Radiation*, Academic Press: New York, NY, 1983.
- [2] Dereniak, E. L., and D. G. Crowe. *Optical Radiation Detectors*, John Wiley & Sons: New York, NY, 1984.
- [3] Kruss, P. W., L. D. McGlauchlin, and R. B. McQuistan. *Elements of Infrared Technology*, John Wiley & Sons: New York, NY, 1962.

-
- [4] Fenton, Q. W., "How Thermocouples Work," *Nucl. Energy* (Br. Nucl. Energy Soc.), Vol. 19, 1980, p. 61.
 - [5] Budde, W. *Optical Radiation Measurements*, Vol. 4, *Physical Detectors of Optical Radiation*, Academic Press: New York, NY, 1983, pp. 101–104.
 - [6] Oriel Corporation. "Optics & Filters," Vol. III, Oriel Corporation, 250 Long Beach Blvd., P. O. Box 872, Stratford, CT 06497, 1990.
 - [7] Cohen, M. L., and J. R. Chelikowsky. *Electronic Structure and Optical Properties of Semiconductors*, Springer Verlag: New York, NY, 1988.
 - [8] Henderson, B., and G. F. Imbush. *Optical Spectroscopy of Inorganic Solids*, Oxford University Press: New York, NY, 1989.
 - [9] Hummel, R. E. *Electronic Properties of Materials: An Introduction for Engineers*, 2nd ed., Springer Verlag: New York, NY, 1993.
 - [10] Hunter, L. P. *Introduction to Semiconductor Phenomena and Devices*, Addison-Wesley: Reading, MA, 1966.
 - [11] *Encyclopedia of Chemical Technology*, 3rd ed., Vol. 20, *Refractories to Silk*, John Wiley & Sons: New York, NY, 1982, p. 602.
 - [12] Hunter, L. P. *Introduction to Semiconductor Phenomena and Devices*, Addison-Wesley: Reading, MA, 1966.
 - [13] Dereniak, E. L., and D. G. Crowe. *Optical Radiation Detectors*, John Wiley & Sons: New York, NY, 1984, p. 26.
 - [14] Viemann, W. "Thin Film Scintillators for Extended Ultraviolet Response Silicon Detectors," *SPIE*, Vol. 186, 1979, p. 90.
 - [15] Harris, R. E. "Si...Mr. Safeguard of the Detector World," *Electro-Optical Systems Design*, Milton S. Kiver Publications, Dec. 1977.
 - [16] Harris, R. E. "PbS...Mr. Versatility of the Detector World," *Electro-Optical Systems Design*, Milton S. Kiver Publications, Inc, Dec. 1976, pp. 47–50.
 - [17] Harris, R. E. "PbSe.....Mr. Super Sleuth of the Detector World," *Electro-Optical Systems Design*, Milton S. Kiver Publications, Inc., March 1977, pp. 42–44,
 - [18] Amelio, G. F., M. F. Tompsett, and G. E. Smith. "Experimental Verification of the Charge Coupled Diode Concept," *Bell System Tech. Jour.*, Vol. 49, 1975, p. 593.
 - [19] van der Ziel, *Noise in Measurements*, John Wiley & Sons: New York, NY, 1976.
 - [20] Budde, W. *Optical Radiation Measurements*, Vol. 4 *Physical Detectors of Optical Radiation*, Academic Press: New York, NY, 1983, pp. 34–38
 - [21] Ulbricht, ETZ 21, 1900, p. 595.
 - [22] Rosa, E., and A. Taylor, Sci. Papers NBS, 281, 1922, p. 447.
 - [23] Goebel, D. G. "Generalized Integrating-Sphere Theory," *Applied Optics*, Vol. 6, 1967, pp. 125–128.
 - [24] Weidner, V., and J. J. Hsia. "Reflection Properties of Pressed Polytetrafluoroethylene Powder," *J. Opt. Soc. Am.*, Vol. 71, 1981, pp. 856–861.
 - [25] Budde, W. *Optical Radiation Measurements*, Vol. 4 *Physical Detectors of Optical Radiation*, Academic Press: New York, NY, 1983, p. 285.
 - [26] Wyszecki, G., and W. S. Stiles. *Color Science: Concepts and Methods, Quantitative Data and Formulae*, 2nd ed., John Wiley & Sons: New York, NY, 1982.
 - [27] National Research Council of Canada, Report APRO-253.
 - [28] Oriel Corporation, "Light Sources, Monochromators & Spectrographs, Detectors & Detection Systems," *Fiber Optics*, Vol. II, Sec. 3, Oriel Corporation, 250 Long Beach Blvd., Stratford, CT 06497, 1994,
 - [29] Budde, W. *Optical Radiation Measurements*, Vol. 4, *Physical Detectors of Optical Radiation*, Academic Press: New York, NY, 1983, Ch. 4, and pp. 39–59.

-
- [30] Lubkin, G., ed. "Tenth Annual Physics Today Buyers' Guide," *Physics Today*, American Institute of Physics, 335 East 45 St., New York, August, 1993.
 - [31] Goldminz, L., et. al. "Mercury-Cadmium-Telluride Photovoltaic and Photoconductive Focal-Plane Arrays," *Optical Engineering*, Vol. 32, 1993, pp. 952-957
 - [32] EG&G Optoelectronics Canada, 22001 Dumberry Road, Vaudreuil, Quebec J7V 8P7, Canada.
 - [33] Brown, D. M., et al. "Silicon Carbide UV Photodiodes," *IEEE Trans. on Electron Devices*, Vol. 40, 1994, pp. 325-333.
 - [34] Cree Research Inc., Durham, NC 22713.

APPENDIX 7A

DETECTOR MANUFACTURERS

Hamamatsu Corporation
360 Foothill Rd., P. O. Box 6910 Bridgewater,
NJ 08807 USA
Phone: 908/231-0960
Fax: 908/231-1539

Hamamatsu Photonics, Solid State Div.
1126-1, Ichino-cho, Hamamatsu City
435 Japan
Phone: 053-434-3311
Fax: 053-435-1037

Oriel Corporation
250 Long Beach Road, P. O. Box 872
Stratford, CT 06497 USA
Phone: 203/377-8282
Fax: 203/378-2457

Infrared Industries, Inc.
12151 Research Parkway
Orlando, FL 32826
Phone: 407/282-7700
Fax: 407/273-9046

CI Systems, Inc.
5137 Claretton Dr., Suite 220
Agoura Hills, CA 91301
Phone: 818/865-0402
Fax: 818/865-0403

Advanced Photonix, Inc.
1240 Avenida Ascaso
Camarillo, CA 93012
Phone: 805/484-2884
Fax: 805/484-9935

Katsura Co., Ltd.
2612-155 Kanai-chu
Machida-City
Tokyo 194 Japan
Phone: (0427) 36-4533
Fax: (0427) 36-4534

Graesby Infrared
12151 Research Parkway
Orlando, FL 32826
Phone: 407/282-7700
Fax: 407/273-9046

Graesby Infrared, Ltd.
Exning Road
Newmarket, Suffolk CB8 0AU, U.K.
Phone: 0638 663963
Fax: 0638 668238

EDO Corp., Barnes Engineering Div.
88 Long Hill Cross Rd., P. O. Box 867
Shelton, CT 06484
Phone: 203/926-1777

Bicron
12345 Kinsman Rd.
Newbury, OH 44065
Phone: 216/564-8000
Fax: 216/564-8047

E G & G Ortec/Berthold
100 Midland Rd.
Oak Ridge, TN 37831
Phone: 615/482-4411
Fax: 615/482-0396

Galileo Electro Optics
Sturbridge Division
Galileo Park, Box 500
Sturbridge, MA 01566
Phone: 508/347-9191
Fax: 508/347-3849

BEA Electro-Optics
640 Pearson St., Suite 200
Des Plaines, IL 60016
Phone: 708/298-1420
Fax: 708-298-1423

Centronic, Inc., Electro-Optics Division
2088 Anchor St.
Newbury Park, CA 91320
Phone: 805/499-5902
Fax: 805/499-7770

Edinburgh Instruments, Ltd.
Riccarton, Carrie
Edinburgh, EH144AP U. K.
Phone: (031) 449-5844
Fax: (031) 449-5848

Infrared Engineering, Ltd.
Galliford Rd.
The Causeway
Maldon, Essex
CM9 7XD, U. K.
Phone: (621) 852244
Fax: (621) 840534

Hilger Analytical
Westwood Industrial Est.
Ramsgate Road
Margate, Kent
CT9 4JL, U. K.
Phone: 0843 225131
Fax: 0843 224402

Silicon Sensors, Inc.
Old Highway 18E
Dodgeville, WI 53533
Phone: 608/935-2707
Fax: 608/935-2775

Eltec Instruments, Inc.
P. O. Box 9610
Daytona Beach, FL 32120
Phone: 904/253-5328
Fax: 904/258-3791

Ealing Electro-Optics
New Englander Industrial Park
Holliston, MA 01746
Phone: 508/429-8370
Fax: 508/429-7893

EMR Photoelectric (Schlumberger)
P. O. Box 44
Princeton, NJ 08542
Phone: 609/799-1000

EG&G Optoelectronics Canada
22001 Dumberry Road
Vadreuil, Quebec J7V 8P7 Canada
Phone: 514/424-3300
Fax: 514/424-3411

OptoElectronics Div. of Textron
1309 Dynamic St., P. O. Box 750039
Petaluma, CA 94975
Phone: 707/763-5900
Fax: 707/762-7383

McPherson Div. of S. I. Corp.
530 Main Street
Acton, MA 01720
Phone: 508/263-7733
Fax: 508/263-1458

Thorn EMI
100 Forge Way, Unit F
Rockaway, NJ 07866
Phone: 201/586-9594
Fax: 201/586-9771

Chapter 8

Optical Systems

8.1 INTRODUCTION

Many traditional optical systems are designed to produce an image of an object to be captured on photographic film, on an imaging detector array, or to be viewed by the eye through an eyepiece. The simplest of these optical systems is a lens, transferring some object or source of radiation to a magnified or demagnified image on an imaging surface, usually planar and in this case called the *image plane*. The image can be placed on a detector, a detector array, a ground glass, or an imaginary plane in a subsequent optical system such as an eyepiece. In designing and using optical systems, one frequently needs to speak of the radiance, irradiance, luminance, or illuminance of the image or of the flux distribution at other places in the optical system.

In addition to traditional imaging optical systems a number of nonimaging systems, such as those found in optical-fiber communication systems, optical data processing systems, laser surgery systems, UV exposure apparatus (for accelerated weathering of materials such as plastic that tend to degrade when exposed to solar UV radiation), and nonimaging concentrators of solar radiation, have useful applications. The methods of radiometry and photometry apply to these systems as well.

This chapter begins with a definition of the *optical axis* and a brief review of simple, idealized lens theory. *Aberrations* are physical processes and their consequences that cause the imaging properties of real optical systems to depart from those of ideal systems. Although it is assumed for most of this chapter that all imaging optical systems are perfect imagers, without aberrations, aberrations can affect the outcomes of radiometric or photometric measurements and one must be aware of these potential problems in the design and/or operation of optical systems intended for optical radiation measurements. Consequently, short descriptions of some lens and mirror aberrations are included. Following this, some characteristics of nonimaging optical systems are discussed, in the context of applying radiometry and photometry to these application areas. The chapter concludes with discussions of integrating spheres, monochromators, windows, radiation sources, goniometers,

and optical systems for measuring the transmittance, reflectance, and scattering properties of objects.

8.2 OPTICAL AXIS

Most optical systems are composed of components with homogeneous optical properties bounded by surfaces of revolution, producing a well-defined center line about which all aspects of the component are radially (or axially) symmetric. The center line of such components is called the *optical axis* of the component. A system of geometrically centered components is a *centered system* having a single well-defined axis of symmetry, which serves as the optical axis for the optical system [1].

For idealized perfect simple lenses composed of surfaces of revolution, having circular edges in planes parallel to each other, and radially symmetric refractive index distributions, it is easy to find the optical axis. It is the line through the centers of the circular shapes of these lenses and perpendicular to the planes of the circles. The optical axis is coincident with the center ray through the simple lens, shown schematically in Figure 8.1. It is the line of perfect geometrical symmetry through the lens. A similar definition applies to mirrors whose surface shapes are radially (axially) symmetric and whose edges are circular. The optical axis is the line through the geometrical center of the mirror's circular edge and perpendicular to the mirror surface at the point of intersection with the reflecting surface.

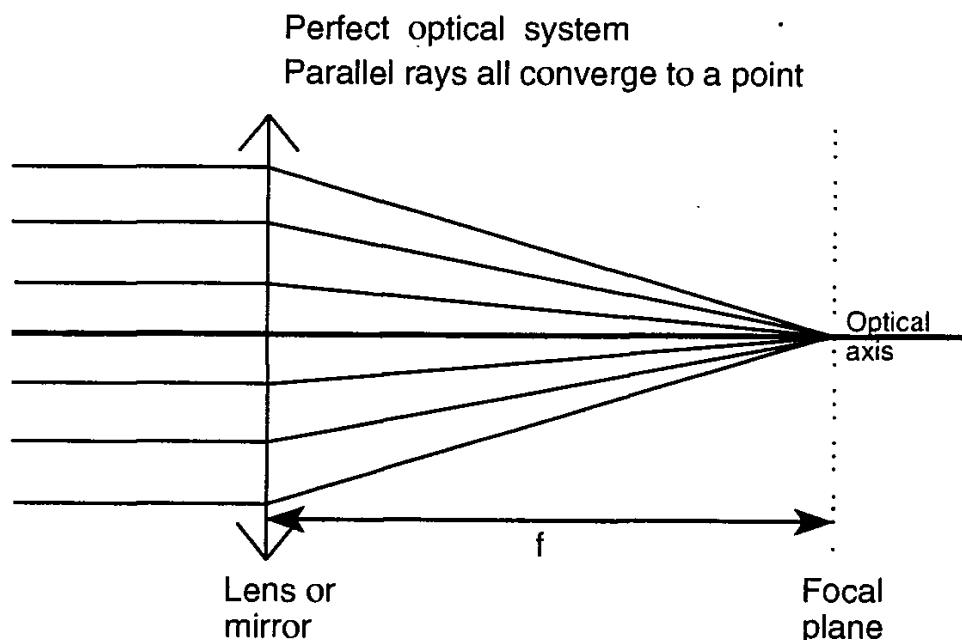


Figure 8.1 A collimated beam of parallel rays incident on the aperture of a perfect lens will converge to a point in the focal plane of the lens after passage through it.

For more complex optical systems, and those with asymmetrical aberrations, the job of defining the optical axis precisely is not as straightforward. A definition that might be somewhat more accurate would be the line through the “center of mass” of the optical system that also passes through the center of symmetry of the aberrations. This may or may not be useful for different applications. For asymmetrical aberrations, however, even this definition will not suffice. The reader is referred to textbooks on lens design and aberrations for a more detailed discussion of this topic [2,3].

8.3 IDEALIZED (THIN) LENS THEORY

Parallel rays entering an idealized perfect imaging system parallel to the optical axis will converge to a geometric point in what is called the *focal plane*, as illustrated for a simple lens in Figure 8.1. Parallel rays entering the optical system at some angle to the optical axis will converge to a different point in the focal plane, as illustrated in Figure 8.2. The distance f along the optical axis from the center of this perfect system to the focal plane is called the *focal length* of the system. This idealized perfect imaging system is approximated by monochromatic radiation passing through a thin lens, one whose thickness is small compared with its focal length. In simplified, thin lens theory, it is assumed that all lenses are very thin, effectively having zero thickness. As shown in Figure 8.1, they are therefore drawn as straight lines with arrow points on each end.

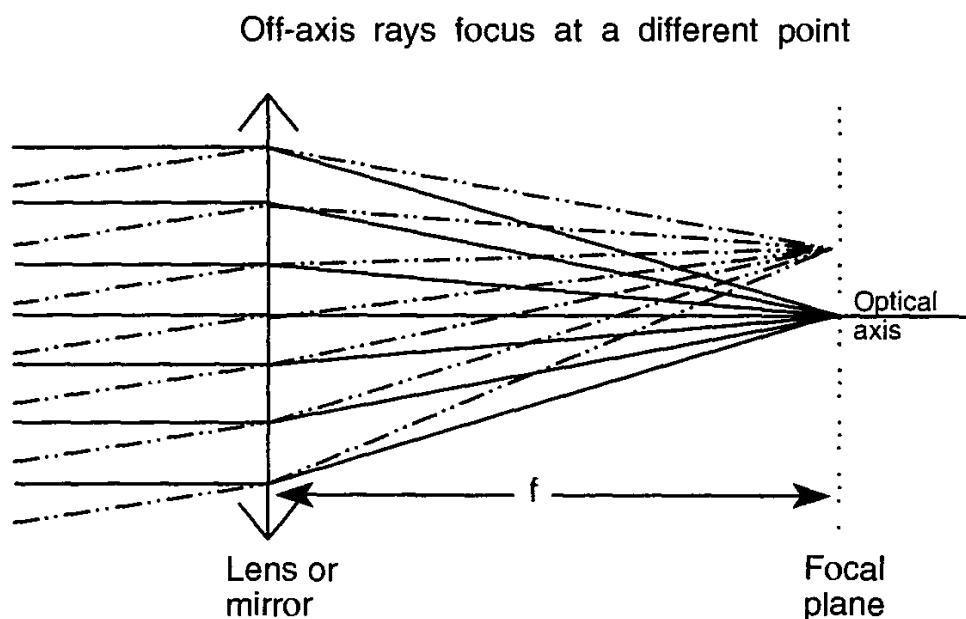


Figure 8.2 Parallel rays incident on a perfect lens at some angle to the lens optical axis will converge to an off-axis point in the focal plane of the lens.

In thick lenses and multi-element lenses, the geometrical “center” of the lens is not necessarily the axial location from which the focal length is measured. In general, a lens can have a different focal length for parallel rays entering it from one side than from the other and the reference marks, called *principal planes*, from which the (focal) distance is measured may not coincide or even lie inside the lens. These complexities are not addressed in this text, since the primary focus is on the use of optical systems in radiometry and photometry.

The focal length f and the lens diameter D are the two most important characteristics of circular aperture lenses and imaging mirrors. The ratio f/D of the focal length to the diameter is called the *F-number* or speed of the lens. This quantity is sometimes designated as *f/no.* or *f/#*, terminology that can lead to great confusion. The terminology of Kingslake [4] is used in this book, with “F-number” sometimes being abbreviated as “F-no.” The lower the F-number, the greater the image irradiance, all other things being equal. (This is demonstrated mathematically in Section 8.4.) In a camera system, greater image irradiance leads to shorter exposure times and a “fast” photograph, a picture taken in a short time interval. This is why imaging systems with large apertures and short focal lengths (small F-numbers) are referred to as fast systems.

Although the imaging properties of real optical systems do not match those of the idealized system defined above, idealized simple system approximations are very useful in preliminary optical system design, for determining approximate design parameters quickly and in assessing the approximate consequences of various modifications to proposed designs. Once preliminary optical system designs have been completed, a more detailed and careful procedure must be followed to ensure that the final design will meet overall system requirements. Lens design and more general optical system design are discussed in numerous textbooks on the subject [5–8]. The simplified lens theory offered here is intended only to illustrate basic relationships in this field and how they affect radiometric system performance generally.

Mirrors are frequently used as the primary or the only optical element in optical systems because their imaging properties are the same for all wavelengths and, as shown in Figure 6.1, most mirrors chosen for these applications have high reflectances that are essentially constant over a reasonably large spectral range. As illustrated in Figure 8.3, the main disadvantage of mirrors is that any finite-sized detector or other device placed at the focus of simple symmetric mirrors blocks some radiation approaching the center of the mirror. This is not a serious problem in most radiometric and photometric applications or in the operation of astronomical telescopes. When it is a problem, however, off-axis mirrors can be used that send images of objects off to one side, as illustrated in Figure 8.4. Designing and fabricating off-axis imaging mirrors without aberrations can be an expensive but viable option for many scientific and other applications.

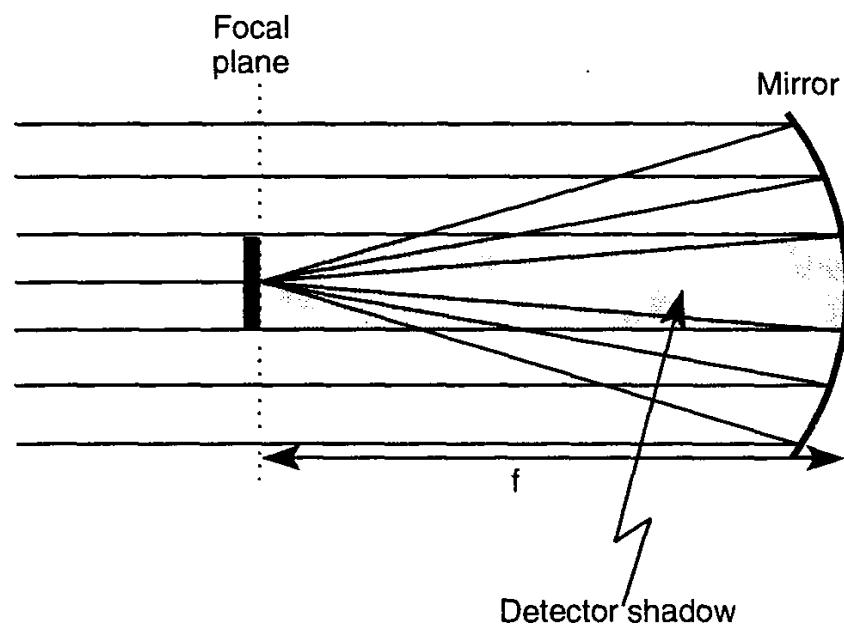


Figure 8.3 Parallel rays incident on a perfect mirror will converge to a point in the focal plane of the mirror. Any detecting device placed on axis in the image plane will block a portion of the flux contributing to the image.

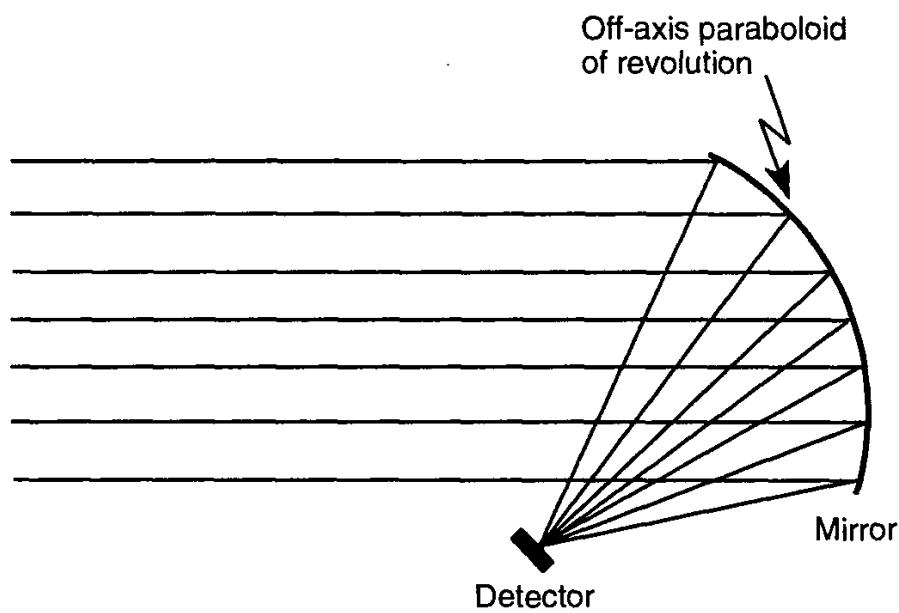


Figure 8.4 Mirrors operated off-axis can avoid blockage of incident flux but are more difficult and expensive to make.

For pedagogical reasons the remainder of this chapter deals only with lens-based optical systems. It is to be understood that all-reflective optics can be substituted for refractive systems in most cases without significantly altering the radiometric and photometric concepts presented.

Parallel rays approaching the lenses and mirrors in Figures 8.1 through 8.4 can be thought of as emanating from a point source of radiation an infinite distance away. If this point is brought to a finite distance S_1 from the lens (or mirror), the image plane will move away from the lens, to the greater distance S_2 from the lens. A pencil of rays emerging from the point source P in the “object,” in what is called the *object plane*, will pass through the lens and converge to point P' in the *image plane*. See Figures 8.5 and 8.6.

The relationship between the object distance S_1 and the image distance S_2 for a lens of focal length f is given by [9]

$$\frac{1}{S_1} + \frac{1}{S_2} = \frac{1}{f} \quad (8.1)$$

This relationship is called “the lens formula in Gaussian form,” and the magnification m of the image is given by

$$m \equiv -\frac{S_2}{S_1} = \frac{f}{(f - S_1)} \quad (8.2)$$

As in previous definitions, the symbol \equiv is used here to mean “equivalence.” The convention of signs used by Jenkins and White [9] is followed in this section:

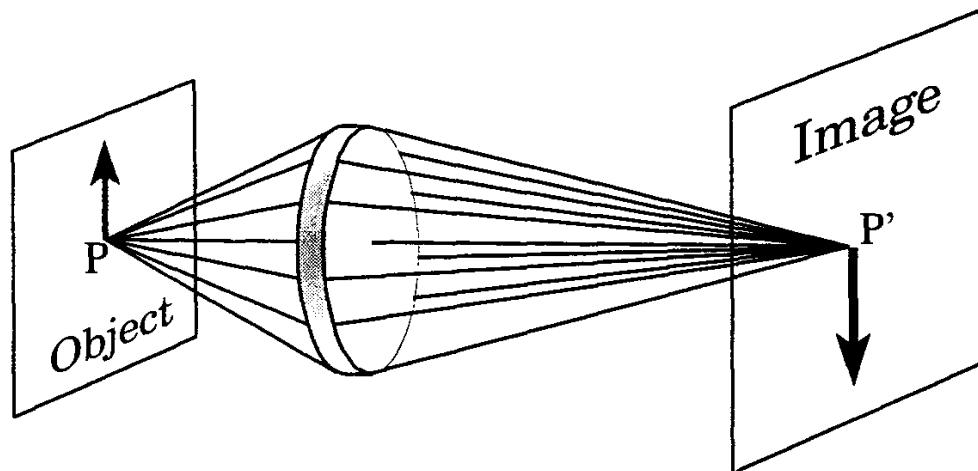


Figure 8.5 Illustration of image formation by a perfect lens.

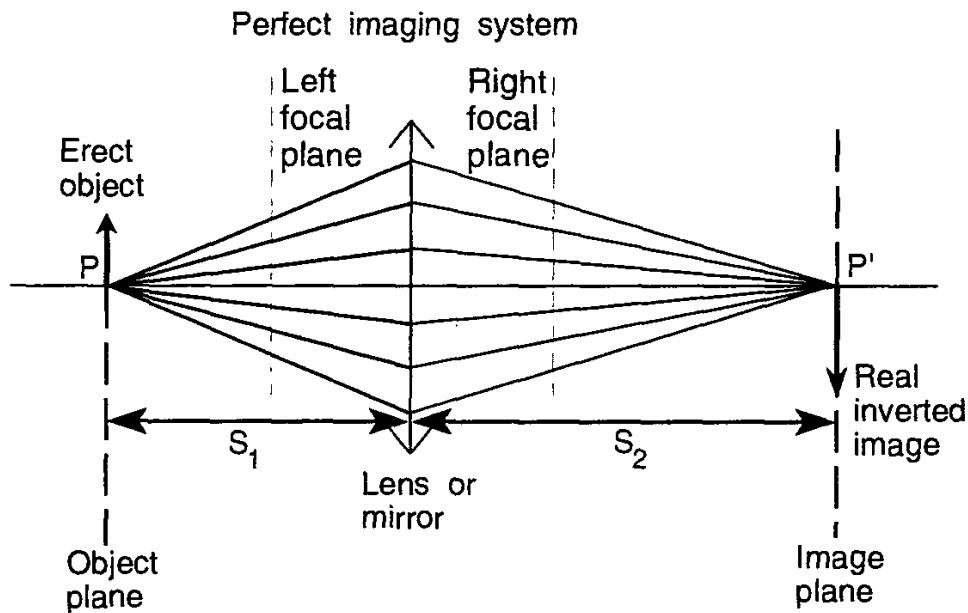


Figure 8.6 Rays from point P in the object plane of a perfect lens will converge to corresponding point P' in the image plane. S_1 is the object distance and S_2 is the image distance.

1. All figures are drawn with the light traveling from left to right.
2. All *object* distances (S_1) are considered as positive when they are measured to the left of the vertex (axial center of the lens), and negative when measured to the right.
3. All *image* distances (S_2) are positive when they are measured to the right of the vertex, and negative when to the left.
4. Both left and right focal lengths are positive for a converging system and negative for a diverging system, leading to the terminology: *positive lenses* and *negative lenses*.
5. Object and image dimensions are positive when measured upward from the axis and negative when measured downward.
6. All convex surfaces encountered are taken as having a positive radius and all concave surfaces are taken as having a negative radius.

A *real image* is one to which real rays converge or from which they diverge, as illustrated in Figure 8.6. Rays actually reach and pass through a real image point. Real images can be made visible (at visible wavelengths) by placing a flat, diffusely reflecting or transmitting surface at their position. The image will be visible due to the light diffusely reflected from or transmitted by the surface. A *virtual image* is one through which actual rays do not pass. A diffusely reflecting surface cannot be positioned to visualize rays to or from a virtual image. A set of rays will appear to diverge from a virtual image, as illustrated schematically in Figure 8.7, but they do not actually pass through the image.

The magnification is unity whenever $S_1 = S_2$. Substituting this equality into

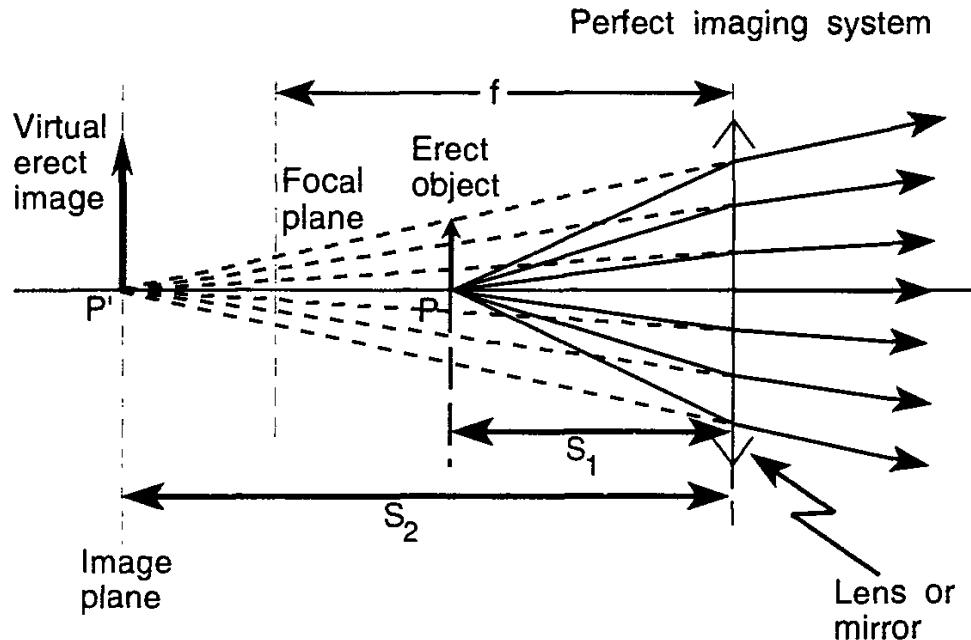


Figure 8.7 Illustration of the formation of a virtual image.

(8.1) yields the condition $S_1 = S_2 = 2f$ for unity magnification. Thus $|m|$ will be greater than 1 when $S_1 < 2f$.

According to the above sign convention, if $|m| > 1$, then the image is magnified (larger than the object). If $|m| < 1$, then the image is demagnified. This is illustrated in Figure 8.8. If m is positive, then the image is virtual and erect, while if it is negative, the image is real and inverted.

Thus, it can be seen that m in Figure 8.6 will be positive for positive lenses whenever $S_1 < f$ and negative whenever $S_1 > f$. S_1 is greater than f in Figures 8.6 and 8.8 and $S_1 < f$ in Figure 8.7.

One goal of this chapter is to derive equations for the radiance and irradiance (luminance and illuminance) of an image, given the radiance (luminance) of the object. The emphasis is on the radiometry and photometry of these processes rather than on accurate representation of the optical image forming properties of the lenses and mirrors involved.

There are optical systems intended primarily for the measurement of radiation using a single detector at the focus of an imaging optical system or at the receiving surface of a variety of nonimaging concentrators. The first of these is addressed in the next section. The second is discussed in Section 8.9.

8.4 RADIANCE AND IRRADIANCE OF IMAGES

One key to determining the radiance and irradiance of images is the radiance invariance, discussed in Chapter 5. It was shown there that, for a lens of transmittance 1.0, the radiance or luminance of the image is equal to the radiance or

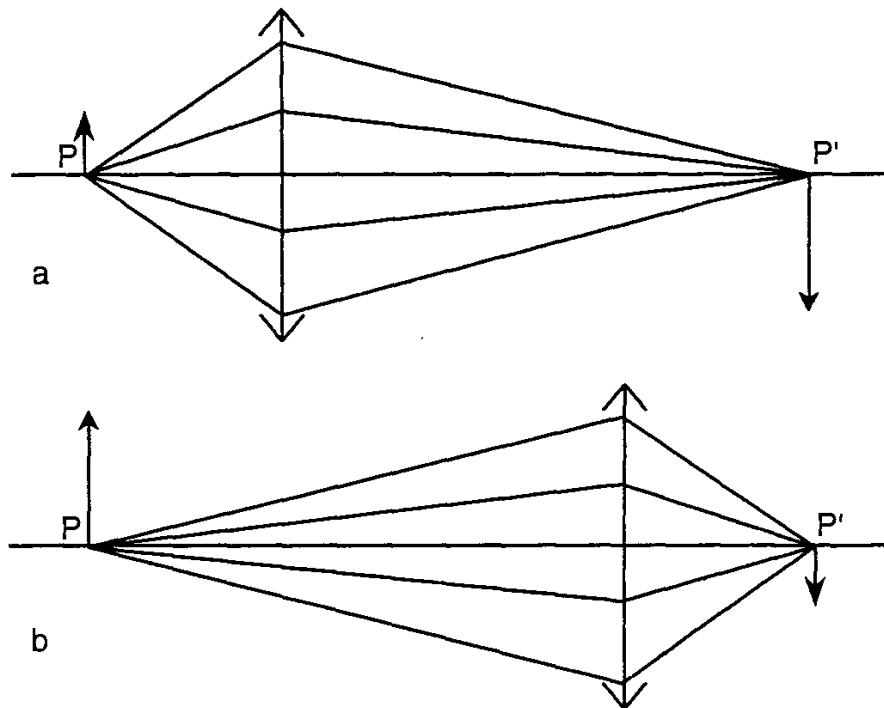


Figure 8.8 Illustration of (a) magnification and (b) demagnification with a perfect lens.

luminance of the object for corresponding points in the object and image. This result may appear to violate intuition. Consider the case shown in Figure 8.8(a, b), where a lens of transmittance T is used to image a source object of radiance L_o , onto the image plane in two different ways. It is natural to ask how the radiance can be the same in the two situations when the image in Figure 8.8(a) is spread out over a larger area. The answer is that in Figure 8.8(a) the lens subtends a larger solid angle at the source so that it collects greater flux, but this flux is spread out over a larger image area and the radiance of the image remains the same as that of the source, ignoring losses in the lens. In Figure 8.8(b), the reverse is true. The lens subtends a smaller solid angle at the source, collecting less flux, but the flux is concentrated into a smaller image area so that the radiance of the image remains the same as that of the source.

The problem is a conceptual one that results in some cases from the common experience of using a magnifying glass to burn a hole in a piece of paper by focusing the sun's rays on the paper. The larger the diameter of the lens (called a *burning glass*), the more the lens appears to concentrate the sun's rays and the hotter the burning spot.

If the lens focal length remains the same, a larger diameter collects a larger solid angle of flux from the object and delivers a larger solid angle of flux to the image, producing a hotter burn spot. The radiance of the source and image have not changed, but the image *irradiance* has increased. This issue is discussed in Section 2.3, where it is stated that the action of an image-receiving surface conceptually is to “convert” the spatial distribution of radiance or luminance incident

on it into a spatial distribution of flux over the image plane, characterized by the irradiance or illuminance of the image.

If Ω_i is the solid angle containing the pencil of rays converging to point P' on axis in the image plane of a lens, then the irradiance E_i of the image, in terms of the incident radiance distribution $L_i(\theta', \phi')$ is given by

$$E_i = \int_{\Omega_i} dE_i = \int_{\Omega_i} L_i(\theta', \phi') \cos \theta' d\omega \quad (8.3)$$

showing how radiance L_i in solid angle Ω_i is converted into irradiance E_i of the image. The *irradiance* of the image is different in (a) and (b) in Figure 8.8. It is the irradiance we are thinking about when we use a burning glass to concentrate rays from the sun and burn a hole in a piece of paper. One can think of it as if it is the irradiance that burns the hole, not the radiance.

Letting $L_o = TL_i$, for an object of radiance L_o that is independent of direction, the integral of (8.3) is

$$E_i = \pi T L_o \sin^2 \theta \quad (8.4)$$

where θ is the angle subtended by the radius of the circular lens at a point on the axis of the lens and T is the lens transmittance, here assumed to be the same for all rays. (See (4.22) and (5.19)–(5.21).) If D is the diameter and f the focal length of the lens, related by

$$\tan \theta = \frac{D}{2f} \quad (8.5)$$

then (8.4) can be written as

$$E_i = \pi T L_o \left(\frac{D^2}{D^2 + (2f)^2} \right) \quad (8.6)$$

$$E_i = \pi T L_o \left[\frac{1}{1 + (2F\text{-no.})^2} \right]$$

showing that image irradiance increases inversely with one plus the square of the F – no.

So far, only the on-axis image irradiance has been discussed. Next we turn to the distribution of irradiance in the image plane at nonaxial locations. Suppose that a perfect lens is used to image a uniform large planar source of constant and Lambertian luminance. The irradiance of the image on the optical axis in the image

plane is given by (8.4). Let θ' be the angle of a point in the image away from the optical axis. The distance from the center of the lens to the field point P' will be given by $f/\cos \theta'$. The square of this distance is proportional to $\cos^{-2} \theta'$. Thus, the irradiance off-axis will be reduced by $\cos^2 \theta'$ by this inverse square law mechanism. It will also be reduced by two more factors of $\cos \theta'$ due to Lambert's cosine law, first for the projected area of the lens as viewed from P' and second for the projected area of the receiving surface in the image plane as viewed from the center of the lens. Because of these effects, the irradiance $E_i(\theta')$ of the image at the field angle θ' off-axis can be expressed in terms of the axial image irradiance E_{io} with the equation

$$E_i(\theta') = E_{io} \cos^4 \theta' \quad (8.7)$$

There are many exceptions to this rule in real optical systems. For example, the placement of aperture stops can cause an angle-dependent blockage of rays and the lens transmittance itself can vary with angle of incidence. However, this $\cos^4 \theta'$ dependence is significant and contributes to the phenomenon known as *field darkening*, the loss of image irradiance or illuminance at the edge of the image plane.

8.5 VIGNETTING

Vignetting (pronounced “vinyetting”) is a term applied to the blockage of off-axis rays by aperture stops or other beam-delimiting components. Vignetting is illustrated in Figure 8.9 for the case of two lenses being used together. Vignetting,

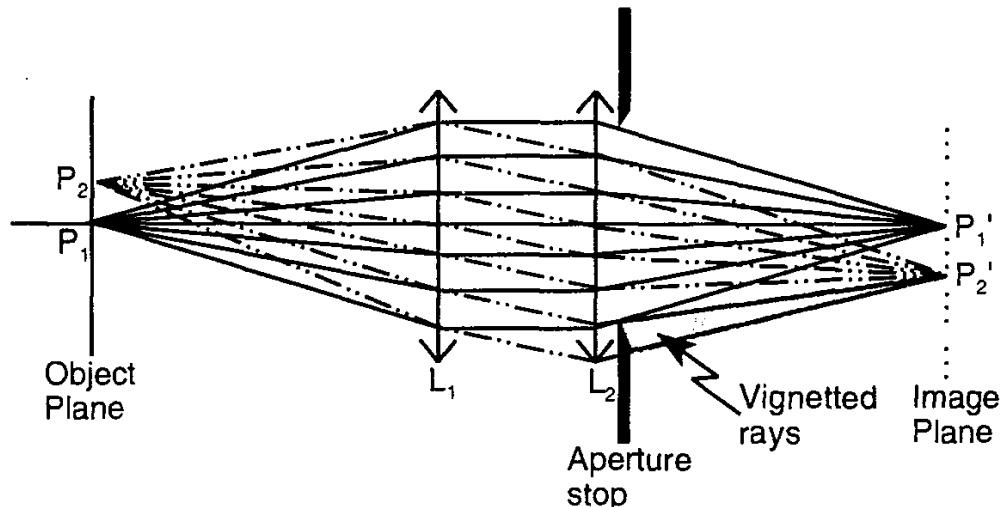


Figure 8.9 Illustration of the loss of flux through an optical system due to vignetting of off-axis rays by an intermediate aperture stop.

when it is present, can produce an additional decline in the off-axis image irradiance beyond that produced by the $\cos^4 \theta'$ effects. Vignetting produces a blockage by the edge of the lens on the right (or some other aperture stop) of some rays approaching that portion of the lens.

A drop in image irradiance or illuminance with the fourth power of the cosine of the off-axis angle is described in Section 8.4. The $\cos^4 \theta'$ factor in (8.7), combined with any vignetting that might be present, combine to produce field darkening.

8.6 ABERRATIONS

Even the very best lenses or mirrors do not make perfect imagers. The spatial distribution of irradiance in the image plane of a very small (nearly infinitesimally small) point source does not perfectly match that of the source. The general term given to the many effects leading to the failure of an image to be a true replica of the object is *aberration*. The reader is referred to many excellent texts on lens and optical system design for a comprehensive treatment of lens and other optical system aberrations [3,10–13]. Here are listed the major types of aberrations, along with brief descriptions.

8.6.1 Spherical Aberration

Lenses and mirrors in most common use have surfaces in the shapes of portions of spheres, mainly because these shapes are relatively easy to fabricate. However, the resulting optical elements exhibit an aberration that is directly attributable to their spherical shapes. *Spherical aberration* arises from the fact that parallel rays reflected from or refracted by a spherical surface do not all converge to the same point. This is illustrated in Figure 8.10, where it is shown that rays approaching the center of the mirror cross the optical axis farther from the mirror than do those approaching the mirror's periphery. This causes a "smearing out" of the focal point, both laterally (perpendicular to the axis) and longitudinally (parallel to the axis).

The effect is also present with homogeneous refractive index lenses having spherical surfaces, as illustrated in Figure 8.11. There is an axial position in the image space, where the image of a point source is as small as possible. One name given to this image (for axially symmetric systems) is "circle of least confusion" and its position can be referred to as the "best focus" position.

In both of these figures it is seen that there is a range along the optical axis over which one can define the image to be "in focus;" that is, a range over which the diameter of the image of the point source is smaller than some acceptable maximum. (The "best" focus is somewhere in this range.) If the diameter of the

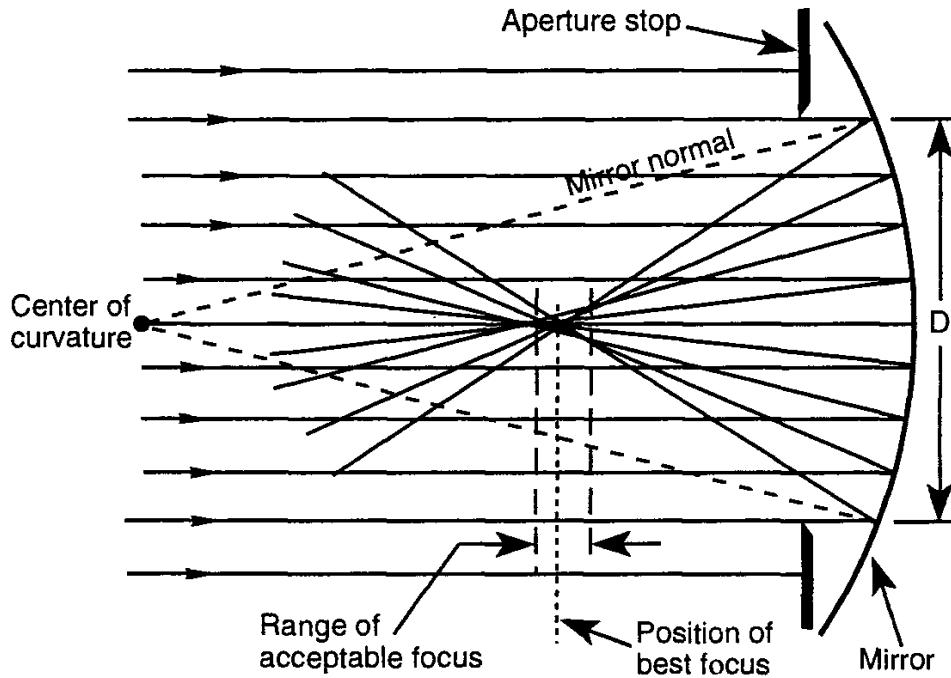


Figure 8.10 Illustration of spherical aberration in a mirror.

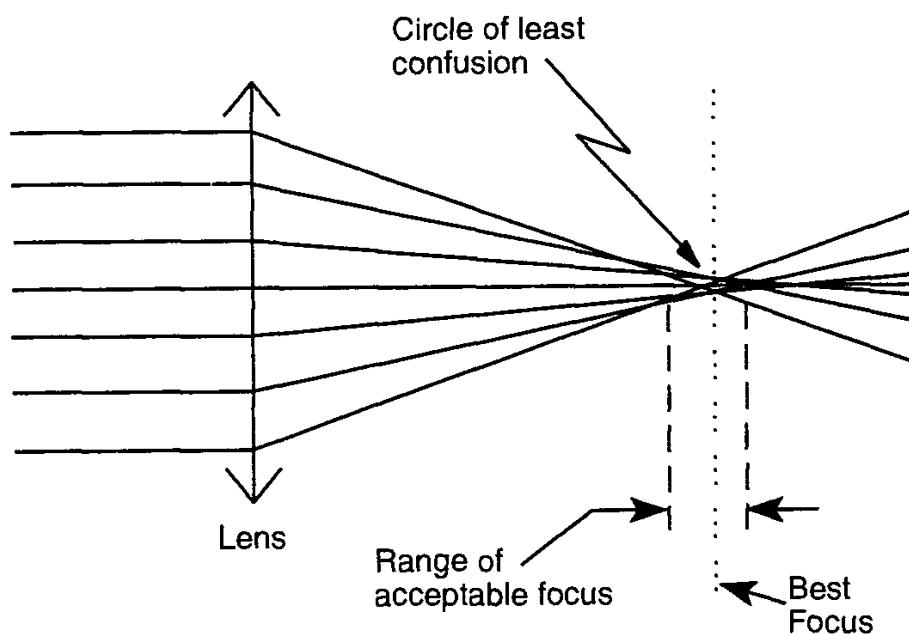


Figure 8.11 Illustration of spherical aberration in a lens.

lens or mirror is reduced without any other changes, the in-focus range is extended because some of the rays leading to this defect are blocked, the circle of least confusion is made smaller (and moved to a different axial location), and the axial range of acceptable in-focus image diameter is increased. This is the reason some people squint their eyes when they lose their eyeglasses. The squinting reduces the effective aperture size in their eyes.

Readers wearing eyeglasses can demonstrate this effect by removing their glasses, forming a small “pinhole” by bringing their thumbs and forefingers together, and looking through the pinhole at what should appear to be an out-of-focus object without corrected vision. The “pinhole” needs to be very close to the cornea, without touching it. By then squeezing the fingers together, the size of the pinhole can be made smaller than the iris of the eye, blocking peripheral rays through the lens of the eye, and permitting only central rays to pass. The observed result is a loss of image brightness and a sharpening of the edges of normally fuzzy, out-of-focus retinal images. This effect is less pronounced in high-illumination areas, where the iris diameter is already smaller and uncorrected vision may already be improved naturally, leading to less noticeable improvement in vision when the pinhole is made smaller.

Now we return to the concept of an axial range over which an image can be said to be in focus. Although Figure 8.11 shows parallel rays entering the lens from the left, we can as easily think of them as having come from a point on axis in the object space to the left of the lens, some distance S_1 from the lens.

From the relationship between S_1 and S_2 shown in (8.1), it is clear that any particular range ΔS_2 of values for S_2 has a corresponding range ΔS_1 of values for S_1 . This means that if an image-capturing device is placed at some point to the right of the lens in Figure 8.11, objects over a range of values of S_1 will be in focus on the image capturing device. This range ΔS_1 is called the *depth of field* of the imaging system. The depth of field of an imaging system can be extended by reducing the diameter of the aperture. This is most commonly accomplished by placing what is called an *aperture stop*, having an adjustable diameter, in front of, behind, or in the middle of a multi-element lens system. This adjustable aperture stop, usually an iris diaphragm, is used to reduce spherical aberration by blocking peripheral rays, keeping them from passing through the lens. The tradeoff with this is the loss of flux that results. Image irradiance decreases as the depth of field is increased by this means. In a photographic camera, this means that exposure times have to be increased when the depth of field is increased.

Photographers desiring to stop fast action, such as at an athletic event, need short exposure times and will therefore open the lens as wide as possible. They must be very careful in focusing, however, to make sure that the object or person being photographed falls within the resulting shallow depth of field. Increasing depth of field without sacrificing exposure times can be accomplished by reducing lens aberrations through better lens design and fabrication (thereby permitting

large aperture areas without image degradation) and by using faster photographic film, film capable of producing a good image with less light.

Spherical aberration is frequently the most prominent aberration in imaging systems. One means of correcting for spherical aberration is to change the *figure* or shape of the optical surface from spherical to aspherical. The resulting lens or mirror is called an *asphere*. Aspherical surfaces with high quality have been very difficult and expensive to fabricate in quantity, but they are becoming more common in the optical industry. Another way to correct for spherical aberration is to introduce a radial refractive index gradient into the lens, forcing peripheral rays to focus at the same place as paraxial rays, even with spherical lens surfaces. Gradient index lenses are also becoming more common. Spherical aberration can be completely eliminated through the use of parabolic mirrors, mirrors whose surface shapes are parabolas of revolution about the optical axis.

8.6.2 Chromatic Aberration

Chromatic aberration is a consequence of the natural and unavoidable variation of refractive index with wavelength (color), called *dispersion*. As refractive index increases, rays are bent through greater angles. This means that wavelengths with higher refractive indices are brought to a focus closer to a lens than those corresponding to a lower refractive index. The effect is illustrated in Figure 8.12 for the case of lens glass, whose dispersion causes the refractive index to decrease with increasing wavelength. Blue light (shorter wavelengths) is brought to a focus closer to the lens than red light (longer wavelengths). The effect is to produce a colored

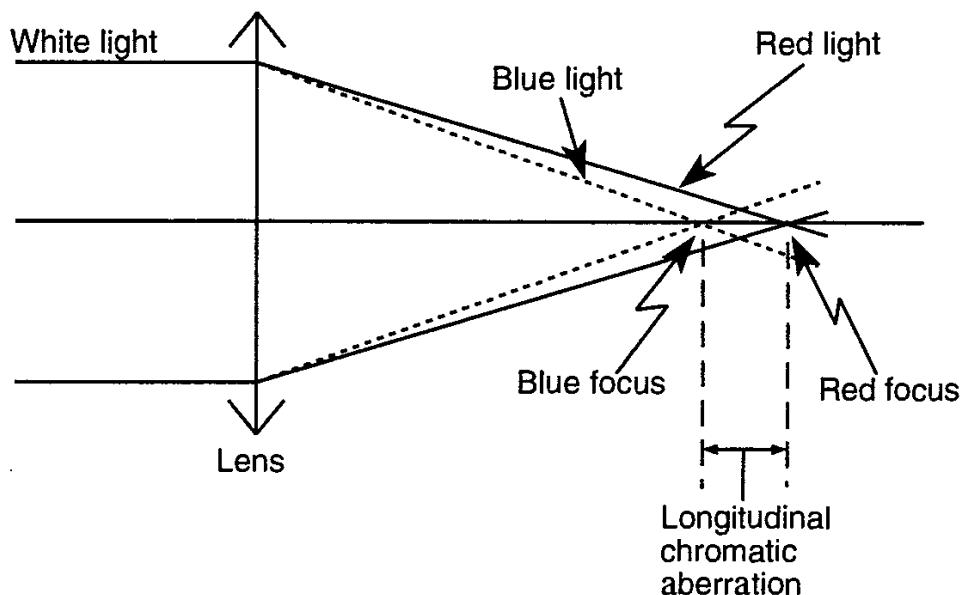


Figure 8.12 Illustration of chromatic aberration.

appearance to images of white light produced by lenses suffering from chromatic aberration.

Some glasses have greater dispersion than others. It is possible to put two or more glasses with different dispersions together in a lens system in such a manner as to partially correct for chromatic aberration. Such lenses, usually doublets or triplets, are called *achromats* because they are relatively free of chromatic aberration, at least over the visible portion of the spectrum. Many achromats also have partial correction for spherical aberration and are therefore overall better imagers. All-reflective optical systems have no chromatic aberration and are recommended for applications involving large-wavelength ranges where optical glasses exhibit substantial dispersion.

It is noted that the chromatic aberration illustrated in Figure 8.12 is termed *longitudinal chromatic aberration*. A related but separate aberration, called *transverse chromatic aberration*, results from lateral (in the image plane) dispersion of rays converging to off-axis image points. This is also called *lateral color* and *chromatic difference of magnification*. The latter term arises from the dependence of magnification (8.2) upon refractive index and hence wavelength [14].

8.6.3 Distortion

Distortion is the failure of an image to be a perfectly (linearly) scaled replica of the object. Distorted images have parts that are magnified or demagnified by different amounts than other parts. The most common kinds of distortion are given the names *barrel* and *pin cushion* distortion because of their characteristic appearances. They are illustrated in Figure 8.13. Other forms of distortion are possible. Distortion is particularly bad for certain imaging systems, such as aerial cameras used in map making, where all parts of the image should remain in true proportions.

8.6.4 Coma

With the exception of lateral color and distortion, the above aberrations are exhibited on-axis and do not greatly change for off-axis points in the image. *Coma*, however, is zero on-axis and increases with distance away from the axis of the imaging system. The problem off-axis is that rays through the center of the imaging system come to a focus at one point in the image plane while those passing through the system away from the axis come to a focus at different points. If the lens or mirror aperture is divided into a series of concentric annular rings and the rays through each ring are considered separately, it is found that the off-axis rings produce not a point focus in the image plane but a circular one. As one goes from

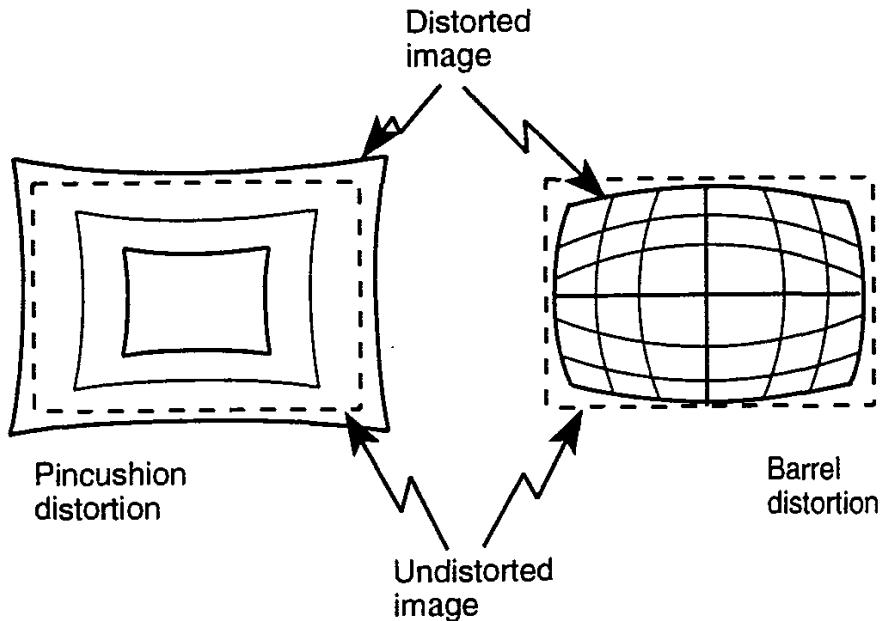


Figure 8.13 Illustration of image distortion produced by an imperfect optical system.

the innermost to the outermost circular ring, the corresponding pattern in the image plane is a circle that changes in both diameter and off-axis location.

Coma is illustrated in Figure 8.14. Since the image irradiance is spread over larger circular areas for off-axis rings of rays, image brightness decreases for the larger image circles. The total image is a superposition of the enlarged and displaced circles from all parts of the aperture. The net effect is illustrated in Figure 8.14. Both positive and negative coma are possible, depending on whether off-axis rays are imaged closer to or farther from the axis than are central rays.

8.6.5 Astigmatism

Astigmatism results when the image of a point source changes from a line in front of the “best focus” position to a line behind the best focus position and perpendicular to the former line. In between these line focuses, the image is elliptical or circular in shape, varying from a line to an ellipse with high eccentricity (very oblong), to one with less eccentricity, to one with least or no eccentricity, to one with increasing eccentricity but with the principle axis now perpendicular to what it was before. This effect is illustrated in Figure 8.15.

Astigmatism generally increases with the distance of the object off-axis. The best-focus image surface is seldom a perfect plane perpendicular to the optical axis of an imaging system. It is more generally parabolic in shape, but other shapes are also possible.

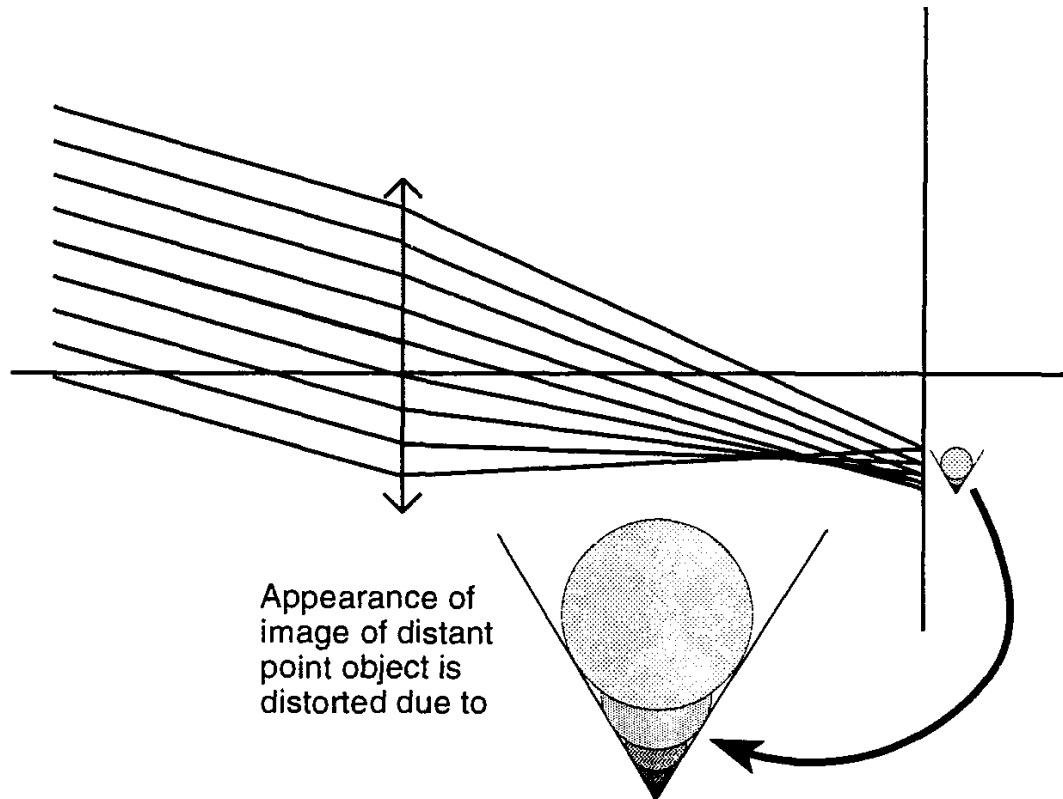


Figure 8.14 Illustration of the lens aberration called *coma*. Although rays close to the central ray produce a small point image, rays at increasing distances from the central ray produce a circular image with increasing size and decreasing distance from the optical axis for the lens shown in the drawing.

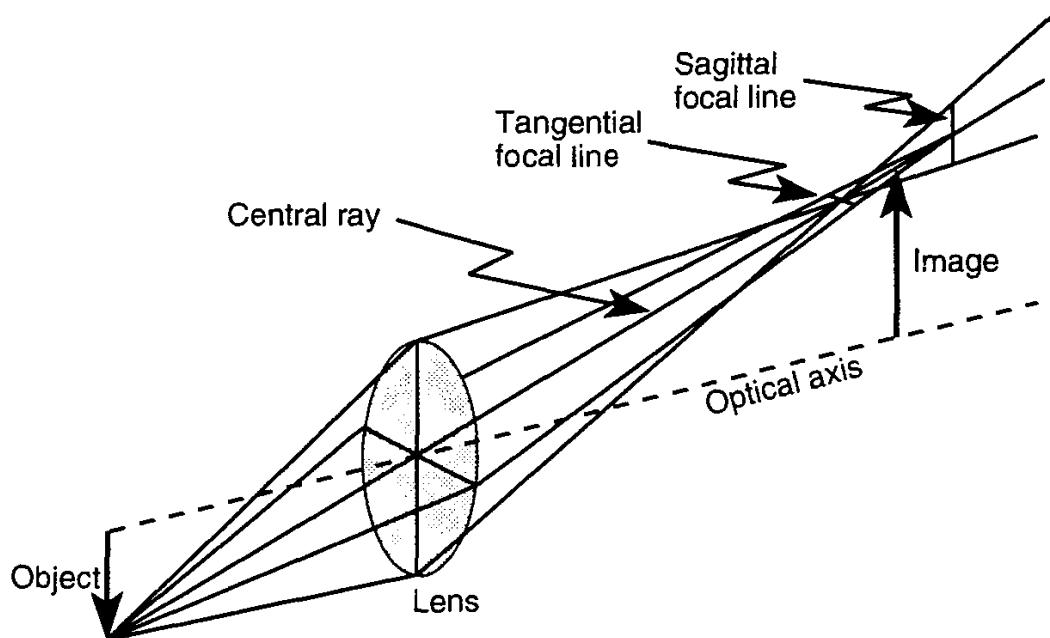


Figure 8.15 Illustration of lens astigmatism.

8.6.6 Field Curvature

The term *field curvature* refers to curvature of the surface of best focus for an image. With this aberration, the image “plane” is not planar but curved. Attempts to capture the image on a planar surface result in parts of the image being out of focus. The curvature is *inward* if it is concave towards the lens and *outward* if it is convex toward the lens.

8.6.7 Correcting Aberrations

Correcting for all the above aberrations in one optical system can present a nearly insurmountable problem for the lens designer. In practice, the process is something like attempting to herd cats. As one aberration is improved, another may get worse. Better materials, improved fabrication techniques, and sophisticated computer programs can help, but high-quality lens design still remains as much an art as it is a science. The task is greatly eased if the specific application permits reasonable tolerances in one or more of the aberrations. In the case of a radiance or luminance meter, for example, the aberration tolerances of the optical system are generally not nearly as stringent as for high-resolution imaging systems or those that must provide good images over a wide angular field of view or over a large flat image plane.

For a typical luminance meter, the spectral range of use is limited, so chromatic aberration is not as serious, and a small angular field of view is usually mapped onto a relatively modest-sized single-element detector in the image plane. As long as the width of the image of a point source is less than, say, 10% of the detector diameter, the residual aberrations of the lens system should not produce serious measurement errors. The result of the various aberrations in this case is a slightly larger field of view than that delimited by the edge of the detector.

On the other hand, substantial aberrations in the imaging system of a heat-seeking missile, for example, could cause it to miss the target altogether. It is important in designing instruments, or optical systems intended to collect and measure electromagnetic radiation, to know the effects of aberrations on the measurement and to match the quality of the system to the requirements of the application.

8.6.8 The Diffraction Limit

Suppose all the above aberrations could be reduced to zero. The lens or mirror system would be considered a perfect imager. It should therefore be able to distinguish between (or *resolve*) two points in the object arbitrarily close together. There is one final remaining effect, however, that limits the resolution of even

perfect optical systems—the *diffraction limit*. In previous sections, the properties of optical systems have been analyzed in terms of their effects on infinitesimally small rays of photons propagating strictly according to the rules of reflection and refraction. Electromagnetic radiation also exhibits wave properties and one of these is its ability to diffract or bend around opaque objects. The reader unfamiliar with this phenomenon can find adequate descriptions in any textbook containing a section or chapter on physical optics. The result for circularly symmetric lens systems is an unavoidably finite-sized image of a point source that has a characteristic shape and appearance.

The problem of calculating this shape for a circular aperture was first solved by Sir George Airy in 1835. The result is given in terms of the Bessel function of unity order [15], $J_1(x)$. The result is

$$E(x) = E(0) \left[\frac{2J_1(x)}{x} \right]^2 \quad (8.8)$$

where $E(x)$ is the irradiance as a function of the angular distance from the geometrical center of the pattern, $E(0)$ is the irradiance at the center of the pattern, $J_1(x)$ is the Bessel function of order 1, and $x = ka \sin \delta$, with k being $2\pi/\lambda$, a the radius of the aperture, and δ the angle from the center of the aperture to the field point where the value of $E(x(\delta))$ is desired. The function in (8.8) is plotted in Figure 8.16.

A central bright maximum dominates this irradiance pattern and is called Airy's disk. Since the secondary maxima are so much reduced in irradiance, it is this central disk that contributes most to the diffraction limit of the resolution of imaging systems. The first dark minimum of the function plotted in Figure 8.16 occurs at $x = 1.220 \pi = 3.832$, and for the small angles involved, $\sin \delta$ is nearly equal to δ , giving an angular radius δ in radians for Airy's disk, in terms of the diameter $D = 2a$ of the aperture, of

$$\delta = \frac{1.22 \lambda}{D} \quad (8.9)$$

The practical effect of (8.9) is that the best possible image of a point source is an Airy disk of finite size.

Imaging systems whose aberrations are small enough to produce resolutions better than this are termed *diffraction limited*, since diffraction remains as the primary contributor to image degradation. As is made evident by (8.9), the larger the aperture of such systems, the greater the resolving power, which is the reason that large aperture, diffraction-limited astronomical telescopes give better resolution than small ones.

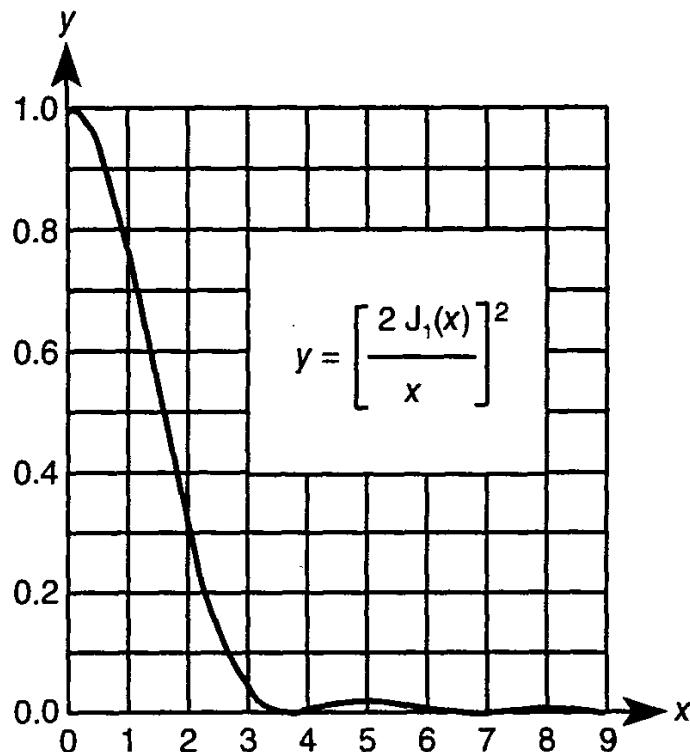


Figure 8.16 Plot of the relative irradiance of the diffraction pattern produced by a circular aperture of radius a . $x = (2\pi a \sin \delta)/\lambda$ and δ is the angle from the optical axis.

It can be noted in passing that imaging systems have been devised that exceed the diffraction limit. One way is to sample the irradiance distribution in the image plane with very high resolution detector arrays and use the function in (8.8) to process the image in such a way as to partially remove the effects of diffraction on the image. Other image enhancement techniques have been developed in recent years that can be used to partially overcome a variety of image degrading processes, including the effects of atmospheric perturbation on images produced by telescopic systems of many kinds [16].

8.7 IMAGE QUALITY

A number of measures of image quality and resolving power (the ability to distinguish two closely spaced objects in an image of them) are available and the field of image analysis is extensive. One indicator of image quality, an easy one to understand and one which is useful in radiometry and photometry, is the *point spread function*, the irradiance as a function of radial distance or angle outward from the optical axis of the image of a point source. The point spread function is illustrated in Figure 8.17. The more rapidly this function drops to zero for increasing distance from its center, the better the image quality, at least in terms of sharpness

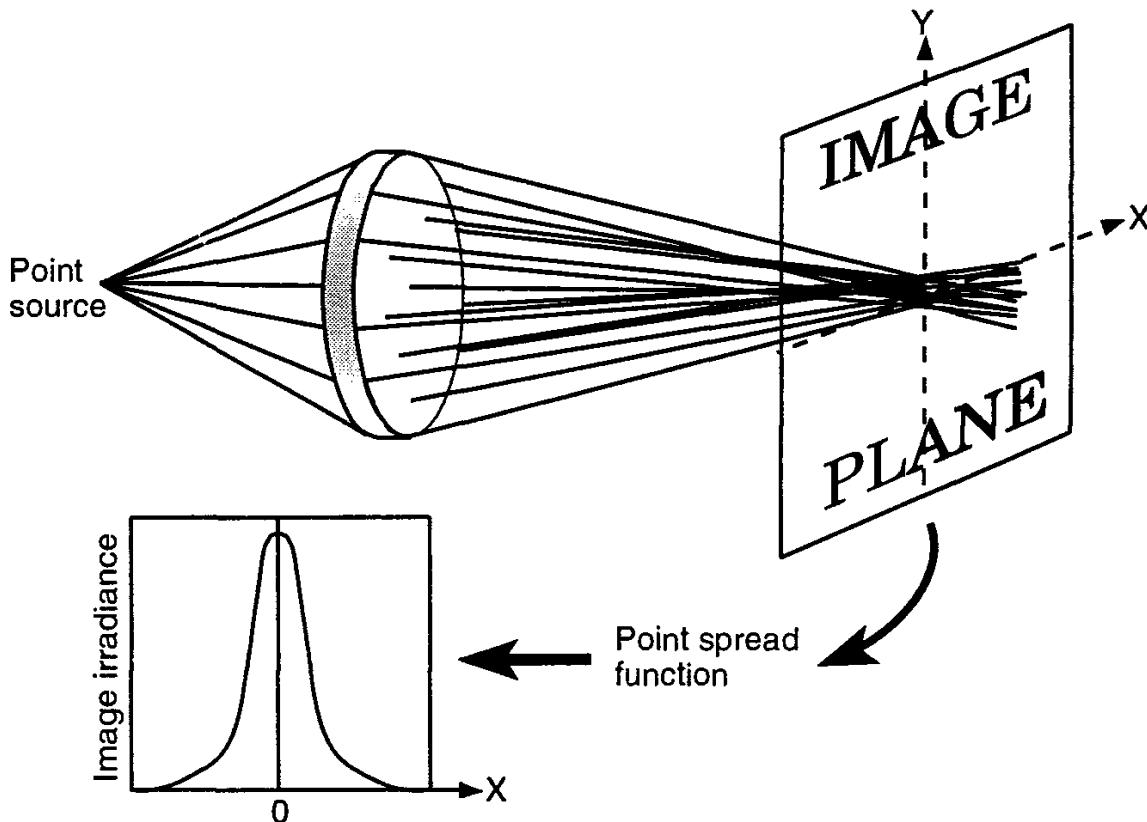


Figure 8.17 Illustration of the concept of point spread function, a plot of the irradiance of the image of a point source.

of focus and resolving power. One can use this measure of image quality to assess the impact of aberrations on the operation of radiometers and photometers. If the width of the point spread function is much smaller than a single detector element in a radiant flux measuring system, then the effects of aberrations on the measurement will be small. If this width is comparable to or greater than the size of the detector element, then attention must be paid to the nature and magnitude of the aberrations in assessing their impact on measurement accuracy and one's confidence in measurement results.

Other measures of image quality can be found in textbooks on lens design and optical engineering [3,11,17,18]. They include the optical and modulation transfer functions (OTF and MTF). These functions provide information about how well an imaging system preserves in its image spatial variations in object radiance or luminance, as functions of the spatial modulation frequency (having units of cycles per unit length) in the object. The relationship between image quality and the distribution of flux through an optical system is mentioned briefly in the next section.

8.8 FLUX DISTRIBUTION

The main characteristic of an optical system of interest to students of radiometry and photometry is its effect on the incident flux distribution. Previous chapters have provided the basic tools needed to determine the propagation of flux through an optical system. Chapter 4 showed how flux propagates through homogeneous media, from one surface to another. Chapters 5 and 6 showed how the flux propagation is modified along rays through the system each time an interface with another medium is encountered, at least in the absence of aberrations. For real, imperfect optical systems, systems that deliver inaccurate, altered versions of the object flux distribution to the image plane, determining the final flux distribution at the end of the optical system can be difficult and time consuming.

The methods in most general use are based upon the conceptual process of launching a large number of rays into the system and arranging these input rays so that they cover the solid angle and spatial range of the intended field of view of the optical system and so that their initial spatial density variation with angle and/or across the entrance aperture matches the variation in the incident flux density. The subsequent paths of these rays are followed mathematically until they reach the final image surface. If the intersection of each ray with the image plane is plotted as a spot on a graph of points in the image plane, the resulting plot of spot density around the optical axis in this plane provides a map of image plane irradiance. Such a plot is referred to as a *spot diagram*. A typical spot diagram for a limited number of rays is illustrated in Figure 8.18.

The spatial density of the spots across the image plane is proportional to the image irradiance distribution. The effects of losses of flux along the rays as they propagate through the system can be handled by a weighting process, with a different weighting factor for each ray that is affected differently by the various loss mechanisms. Rays that fail to reach the image surface at all can be assigned a weighting factor of zero. The ratio of the weighted sum of the image rays to the sum of the incident rays is a measure of the total flux transmittance factor for the optical system. The greater the number of rays used in this calculation, the better will be the final result.

A detailed and comprehensive discussion of this process is beyond the scope of this text. The reader is referred to several textbooks on optics, lens, design, and optical instrument design for more comprehensive and detailed information [3,4,7].

Fortunately, the art and science of optical system design have advanced sufficiently to permit the development of several commercially available computer programs that do a reasonable job of calculating the flux distribution at the image surface of optical systems composed of mirrors, lenses, and beamsplitters, even in the presence of aberrations, assuming that one knows the properties of the component elements of the system. The vendors of these programs include Stellar

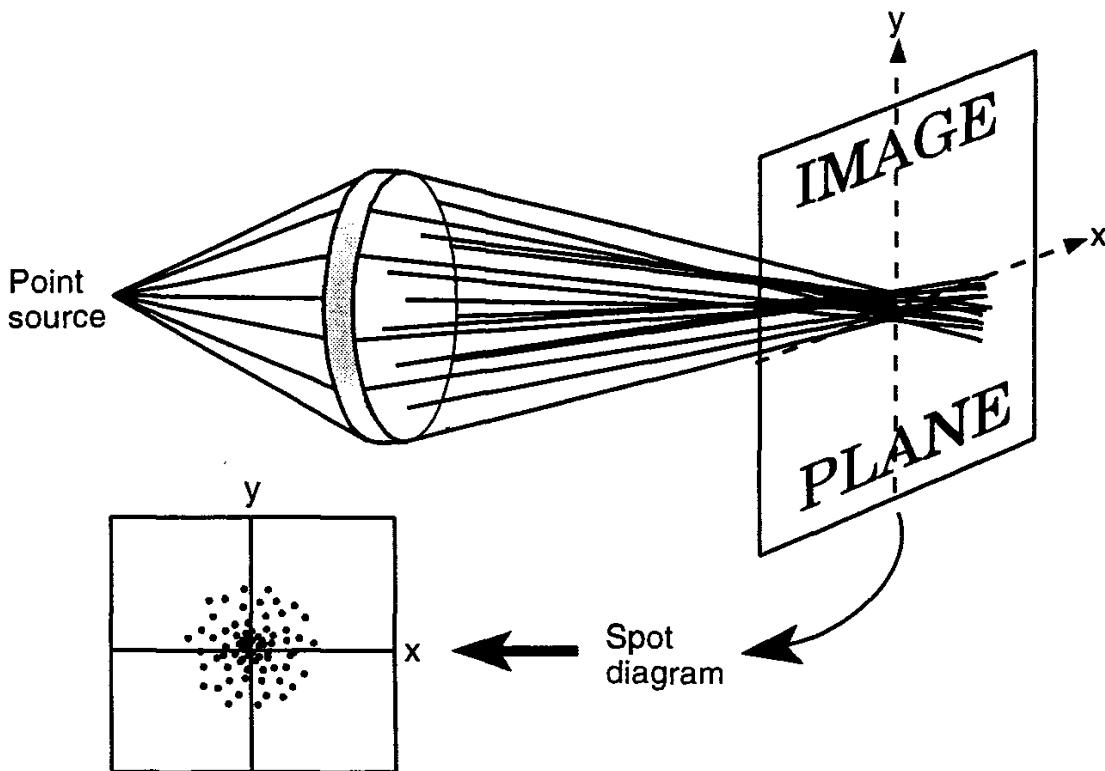


Figure 8.18 Illustration of the concept of spot diagram as a map of point source image irradiance distribution.

Software of Berkeley, California, and Optical Research Associates of Pasadena, California. In preparing to use these programs, or others like them, it is important to have at least a rudimentary knowledge of lens design concepts and terminology.

Most radiometers and photometers employ some kind of optical system to deliver incident radiation to a detection device where the detected radiation can be converted into an electrical signal that is related to the radiation incident upon the entrance aperture of the instrument. The presence of aberrations in the optical system preceding the detection system can produce errors of measurement or errors in understanding the meaning of the measurement results. Some of these errors can be eliminated or reduced through proper calibration procedures. The design of radiometers and photometers is discussed in Chapter 9.

8.9 NONIMAGING OPTICAL SYSTEMS

According to Welford and Winston [19]: "In the mid-1960s it was realized in at least three different laboratories that light could be collected and concentrated for many purposes, including solar energy, more efficiently by nonimaging optical systems than by conventional image-forming systems. The methodology of design-

ing optimized nonimaging systems differs radically from conventional optical design. The new collectors approach very closely the maximum allowable theoretical concentration; and for two-dimensional geometry, which is important for solar energy collection, this limit is actually reached."

The main objective of nonimaging concentrators is to collect the maximum amount of flux possible, consistent with a number of design constraints. The concept of concentration ratio is central to any discussion of this class of optical system. The concentration ratio, designated C can be defined many ways, but the simplest conceptually is the ratio in the absence of any system losses of entrance (input) A_i to exit A_e aperture areas of the generalized optical system illustrated schematically in Figure 8.19:

$$C \equiv \frac{A_i}{A_e} \quad (8.10)$$

Except for the trivial case of $C = 1.0$ with no optical system present, there will in reality be losses due to absorption, reflection, and/or scattering in going between these two apertures in a real optical system. Thus, the actual concentration ratio will be somewhat less than the theoretical value expressed in (8.10).

In Section 5.10 (Equation (5.32)) it was shown that the theoretical maximum concentration ratio for going from a medium with refractive index = 1.0 to one whose refractive index is n is given by

$$C_{\max} = \frac{n^2}{\sin^2 \delta} \quad (8.11)$$

where δ is the half-angle of the conical solid-angle field of view at the entrance aperture. This means that the smaller the size of the source whose flux is being concentrated, the greater the possible concentration ratio. Conversely, if the source solid angle increases to a hemispherical one, the denominator in (8.11) becomes 1.0 and n^2 is the maximum concentration ratio possible in this situation. If $n = 1$, no concentration is possible in this case.

Both imaging and nonimaging optical systems can be used in the design of radiometers and photometers. The conventional lensless Gershun tube radiance meter described in Problem 1.2 is a nonimaging radiometer. The lensed version of this meter, introduced in Section 5.9, is an imaging radiometer since it places an image of the field of view on the detector area.

In conventional (imaging) optical system design, the surfaces of the optical elements involved are usually either planar or spherical in shape. They could also be aspherical surfaces such as paraboloids that are surfaces of revolution about the optical axis. Special lens design methods are available that take advantage of the

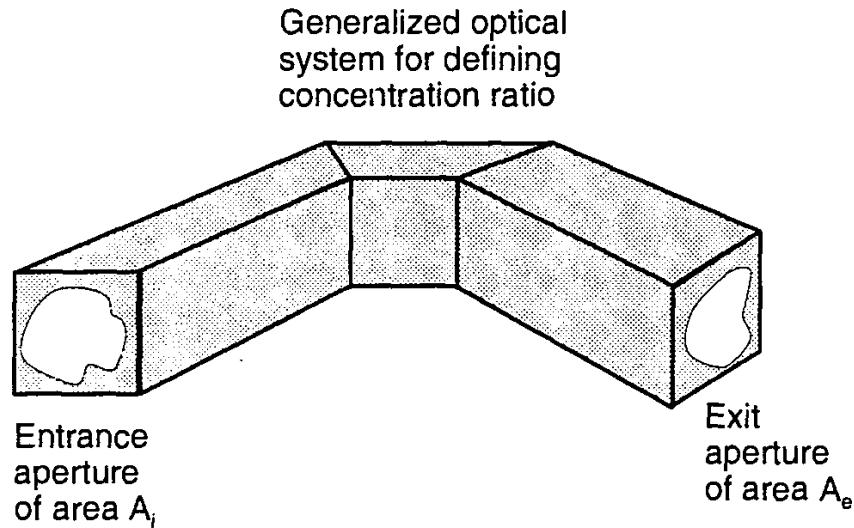


Figure 8.19 Illustration of a generalized optical system with arbitrary entrance and exit apertures for the purpose of defining concentration ratio.

symmetries involved in these systems, but other methods are available that are more general.

According to Welford and Winston [19], the surfaces of nonimaging concentrators do not generally have spherical surfaces and sometimes no explicit analytical form is available for the surface. Designing nonimaging optical systems therefore is best accomplished using ray-tracing methods that make no specific assumptions about the surface shapes of optical elements within the system. Ray-tracing can be used for nonimaging systems if the procedure properly accounts for alterations in ray direction as each ray is propagated through the system, properly accounting for the positions and/or shapes of the intervening optical elements. Thus, the density of image rays can be used to provide a map of image irradiance distribution. Welford and Winston show how to cast the laws of reflection and refraction into a vector form to assist in the process of computerized ray tracing through optical systems [19].

The most thoroughly studied form of nonimaging concentrator is something called a *compound parabolic concentrator*, or CPC, one version of which is illustrated in Figure 8.20. It is a characteristic of a parabola that all rays parallel to its axis and reflected from it pass through its focal point. A parabola of revolution about the parabola's axis has this same property in three dimensions, at least for rays parallel to the axis.

Figure 8.20 shows a section through a compound parabolic concentrator. All rays entering the aperture parallel to the one labeled "b" in this drawing will be reflected to the focus of the upper parabola at the lower edge of the exit aperture. Ray "a" in the drawing and all those parallel to it will be reflected to the focus of

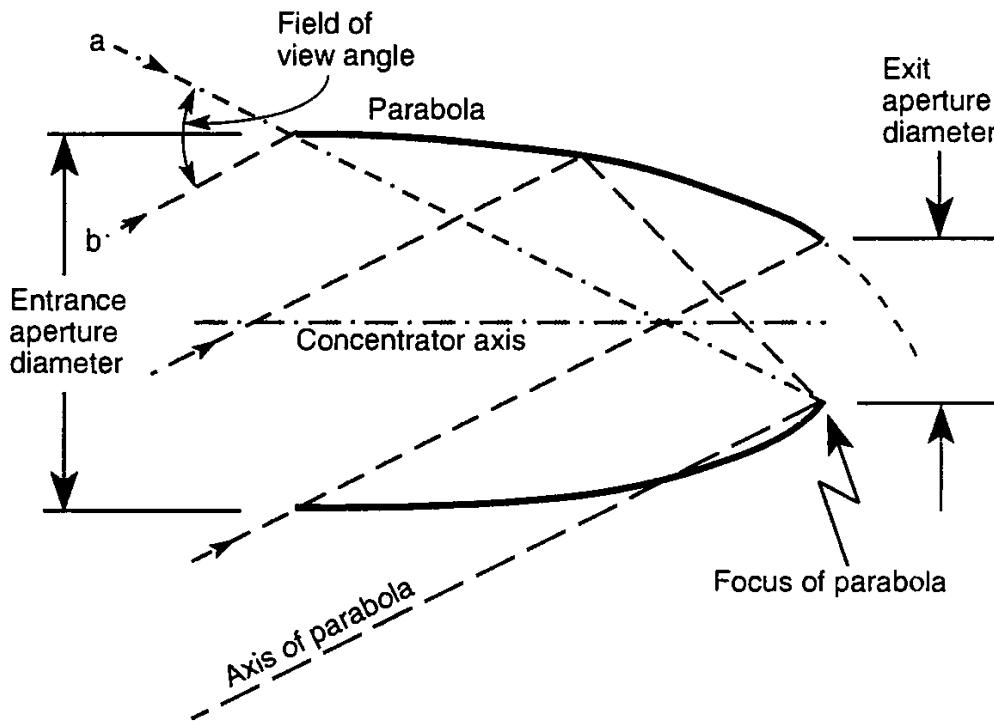


Figure 8.20 Design of a representative compound parabolic concentrator.

the lower parabola at the upper edge of the exit aperture. All rays in the plane of the drawing with angles between rays a and b will be reflected out of the exit aperture. Thus, all rays entering the entrance aperture between the angles of a and b will leave the (smaller) exit aperture. Concentration will have therefore been accomplished without imaging, with the concentration ratio being the ratio of the entrance to exit aperture areas. A complete, three-dimensional CPC can be formed either by extending the parabolas shown in Figure 8.20 vertically out of the plane of the paper, forming what is called a two-dimensional concentrator, a concentrator with long rectangular entrance and exit apertures, or by rotating the parabolas around the line labeled “concentrator axis” thereby forming a three-dimensional concentrator with circular entrance and exit apertures.

8.10 THROUGHPUT

There is a concept of interest in the design and evaluation of optical systems that applies to both imaging and nonimaging systems. The concept is not crucial in the design of radiometers and photometers, so only a brief abstract of its definition is provided here. The radius r of the aperture of the optical system shown in Figure 8.21, multiplied by the half-angle α of the angular extent of the field of view of

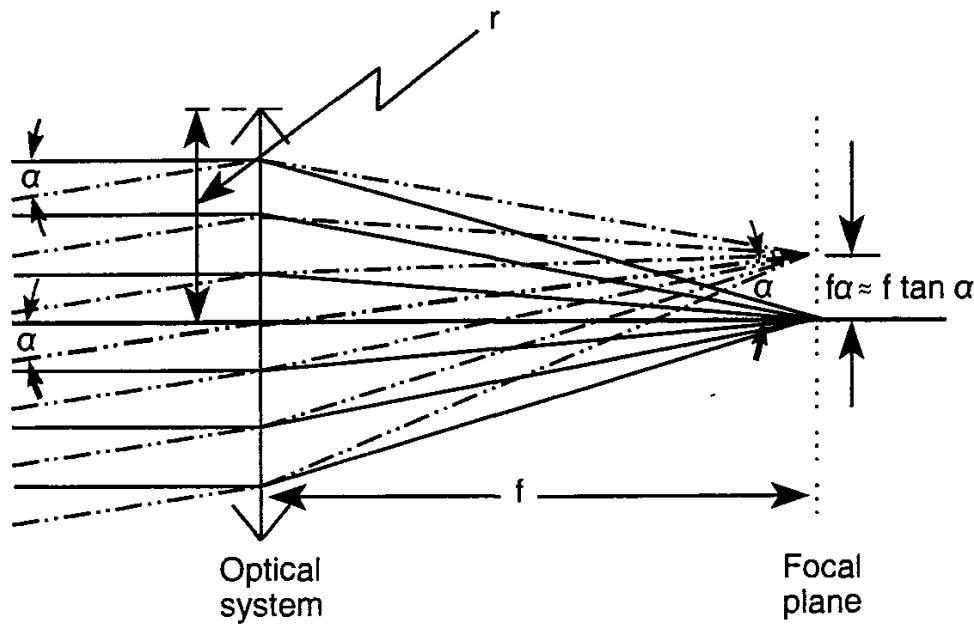


Figure 8.21 Geometric construction for the definition of optical system throughput or *etendu*.

the system yields a quantity $r\alpha$ called the *throughput* of the system. It is also known as the *etendu*, *extent*, or *acceptance* of the optical system [20,21]. Neglecting losses due to absorption, reflection, and scattering, and assuming no blocking elements in the system, the throughput is constant throughout the system, from element to element along the optical axis. The throughput is $r\alpha$ at the lens and in the focal plane it is the image height $f\alpha$ times the angle r/f subtended by the radius of the aperture at the focal point.

A more generalized form of the throughput, usually designated by Θ , is given by

$$\Theta = \int_A \int_{\Omega} \cos \theta d\omega ds \quad (8.12)$$

where A is the area through which a beam of radiation filling solid angle Ω is passing. Following the derivation of Nicodemus [20], it is noted that the differential element of throughput is given by

$$d^2\Theta = \cos \theta d\omega ds \quad (8.13)$$

and from (1.21) the element $d^2\Phi$ of flux can be written in terms of the element of throughput as follows:

$$d^2\Phi = L d^2\Theta \quad (8.14)$$

Since flux is conserved in a lossless, isotropic, uniform medium, $d^2\Phi$ remains unchanged along an elementary beam associated with a ray propagating through an optical system. In Section 5.5, it is proved that the radiance L is also conserved along such a ray. If both L and $d^2\Phi$ are conserved along a ray, then according to (8.14), so must be the element $d^2\Theta$ of throughput. Now, if it is assumed that the solid angle Ω in (8.12) (that is, filled with rays of the beam) is the same for all points over the area A and that there are no losses due to vignetting as the beam propagates through an optical system, then the two integrals in (8.12) are independent and separable, leading to the expression

$$\Theta = \int_A ds \int_{\Omega} \cos \theta d\omega = A \Psi \quad (8.15)$$

where Ψ is the value of the second integral on the right side of (8.15). Ψ is called a *projected solid angle* or a *weighted solid angle* and we see that the throughput is equal to the product of the area A of a beam's intersection with a reference surface in an optical system and the projected solid angle Ψ .

The above discussion shows that to keep the throughput (and the flux) constant through an optical system, increasing either A or Ψ requires a corresponding decrease in the other value. Thus, in Figure 8.21, an increase in either r or α will increase the flux transfer through the system unless the other of these two quantities is decreased a like amount. Throughput can be used in the design of radiometers and photometers to keep track of flux transfer through optical systems and in evaluating the tradeoffs between angular field or view and flux concentration. For more information, the reader is directed to [20] and [21].

8.11 INTEGRATING SPHERES

Integrating spheres are necessary components of some optical systems, especially those intended for the measurement of hemispherical optical properties of materials. The integrating sphere is valuable because of its ability to deliver to a detector a fixed fraction of the flux entering it regardless of the angular distribution of the entering beam of flux. This was mentioned in Section 7.7. An expression for the efficiency of an integrating sphere, the fraction of flux entering a small entrance port that emerges from an exit port, is now derived following a procedure described in detail by Goebel [22].

Entrance and exit ports represent areas whose reflectances are not the same as that for the sphere wall. Sometimes an "exit" port may be filled with a surface whose reflectance is to be measured by illuminating it directly through an entrance port on the opposite side of the sphere and placing a detector in another "exit" port.

In order to handle portions of a sphere wall having different reflectances, the derivation begins with the assumptions that the inner-coated surface of the sphere has $n + 1$ areas (the i th one being of magnitude a_i), each having Lambertian reflectance ρ_i , and that the remaining area of the sphere (the coated wall with highest reflectance possible) is a_w having reflectance ρ_w . Any open ports in the sphere are included in the areas a_i . Open ports are assumed to be areas with zero reflectance. The sphere's inner radius is R so that its total area is $A_s = 4\pi R^2$ or πD^2 with D being the sphere inner surface diameter. The reason for including areas in the sphere that are neither ports with zero reflectance nor the sphere wall with reflectance ρ_w is to permit the inclusion in the sphere of reflectance samples and/or port plugs with reflectances different from that of the wall.

The ratio of each area a_i to the total area A_s of the sphere is called f_i . The average reflectance of the sphere's interior surface is

$$\langle \rho_w \rangle = \sum_{i=0}^n f_i \rho_i + \rho_w \left(1 - \sum_{i=0}^n f_i \right) \quad (8.16)$$

Let Φ_o be the flux (from the sphere's entrance port) incident upon area a_o having reflectance ρ_o . Let Φ_i be the total flux reflected from a_o , and subsequently multiply reflected from all other portions of the sphere with nonzero reflectance, that is incident on a_i . Using an infinite number of multiple reflections, Goebel shows [14] that

$$\Phi_i = \frac{f_i \rho_o \Phi_o}{1 - \langle \rho_w \rangle} \quad (8.17)$$

Thinking of Φ_i as the flux leaving exit port of area a_i (or falling on a detector having this area), for some value of i , the ratio $F_i = \Phi_i / \Phi_o$ is called the sphere efficiency for transferring flux to the i th area, currently the exit port.

It can be shown that

$$F_i = f_i \rho_o \left[1 - \rho_w \left(1 - \sum_{i=0}^n f_i \right) - \sum_{i=0}^n f_i \rho_i \right]^{-1} \quad (8.18)$$

According to Goebel, equivalent relations were derived by Jacques and Kuppenheim [23] using integral equation methods and by Hisdal [24] using matrix relations. Consider the case of an integrating sphere with two holes, an entrance port of area a_1 and an exit port (or a nonreflecting detector placed in the exit port) of area a_2 . Let the flux entering the entrance port illuminate area a_o of the sphere wall of

reflectance $\rho_w = \rho_o$. One has $\rho_1 = \rho_2 = 0$. The fraction of flux entering the entrance port (illuminating area a_o) that escapes out the exit port is therefore

$$F_2 = \frac{f_2 \rho_w}{1 - \rho_w(1 - f_1 - f_2)} \quad (8.19)$$

with ρ_w being the sphere wall reflectance, and f_1 and f_2 being the fractions of total sphere area attributed to the entrance and exit port areas, respectively. This function is plotted in Figure 8.22 for entrance and exit ports with equal area fractions f_1 and f_2 ranging from 0.002 to 0.032.

Coatings with reflectances as great as 0.99 are available [25]. As was mentioned in the previous chapter, the optical properties of a high-reflectance coating (called PTFE powder) for integrating spheres have been reported by Weidner and Hsia [26]. The hemispherical reflectance of this material exceeds 0.990 and is nearly constant over the wavelength range from 350 to 1,800 nm.

The validity of the assumptions made in performing the above derivation decreases as the fractional areas of the entrance and exit ports increase and the wall reflectance decreases. Thus, ideally one wants a sphere with very high wall

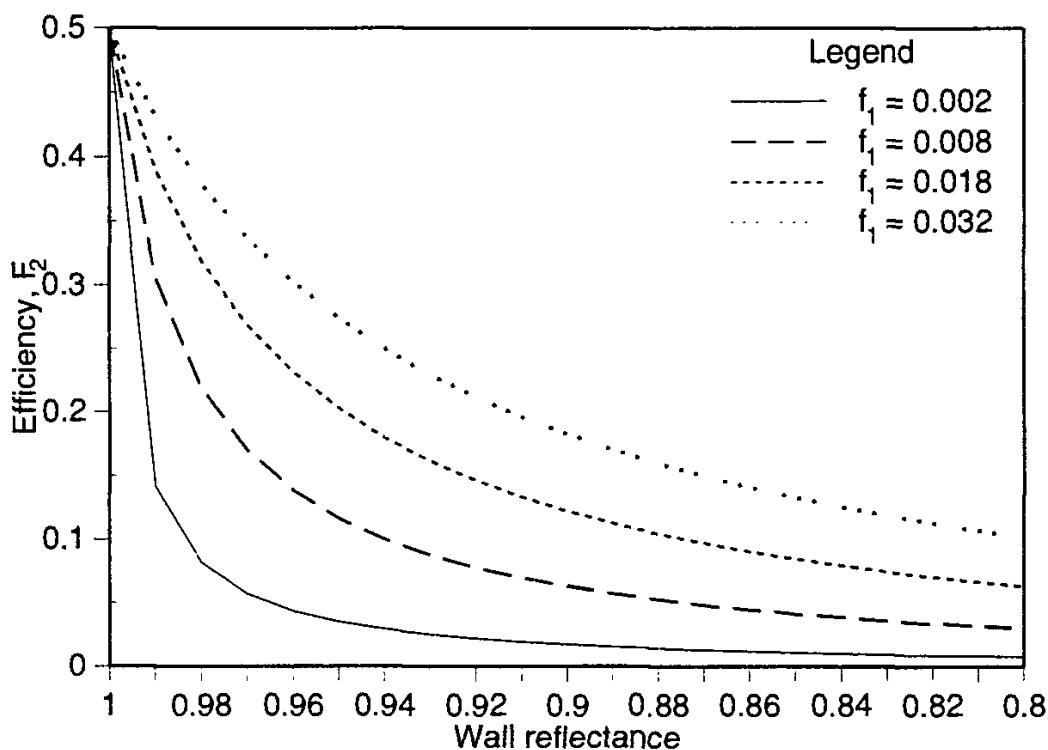


Figure 8.22 Plot of integrating sphere efficiency versus wall reflectance for several different entrance and exit port sizes. $f_1 = f_2$ is the fraction of total sphere area occupied by the equal area entrance and exit ports.

reflectance and very small entrance and exit ports. High wall reflectance is achievable practically. The sphere diameter should be chosen large enough so that the total area of the sphere is much greater than that of the ports. Integrating sphere design involves a trade-off between sphere efficiency and its ability to deliver a fixed fraction of the flux entering it to the detector, regardless how the entering flux is angularly distributed.

Expressions similar to (8.19) for the cases of integrating sphere reflectometers of different designs are provided by Goebel [27]. Integrating spheres are used for a variety of purposes in optical systems that are used in the fields of radiometry and photometry. Several are now described.

8.11.1 Cosine Correction

It is pointed out in Section 1.5 that the output of a perfect irradiance meter illuminated uniformly with collimated radiation fully filling the sensing area will decrease with the cosine of the angle of incidence. This point is repeated in Section 7.7.1 and a method of approximating the ideal “cosine response” is described and illustrated (Fig. 7.18).

It is very difficult to achieve true cosine response for irradiance and illuminance meters using cosine correcting schemes like the one illustrated in Figure 7.18. Although it does not generally deliver much flux to the detector, an integrating sphere can provide nearly perfect cosine response. A representative physical arrangement for this is shown schematically in Figure 8.23. Radiation coming from any portion of a hemispherical solid angle can be incident upon the entrance port of the sphere. The inner wall is coated with a diffusely reflecting material of high reflectance.

The detector is shielded from receiving radiation directly from the entrance port. The shield is a small baffle that is coated with the same material as the inner wall of the sphere. The radiation reaching the detector is the result of many inter-reflections over the whole inner surface of the sphere, and a *fixed* fraction of the flux entering the sphere reaches the detector. Experiments have shown that properly designed spheres of this type exhibit very close to the theoretically optimum cosine response to the angle of incident radiation except at very large angles of incidence.

8.11.2 Transmissometers and Reflectometers

There are many possible configurations for measuring specular and diffuse transmittance and reflectance. For measurements of samples with significant nonspecular components, integrating spheres are commonly used for planar samples of the

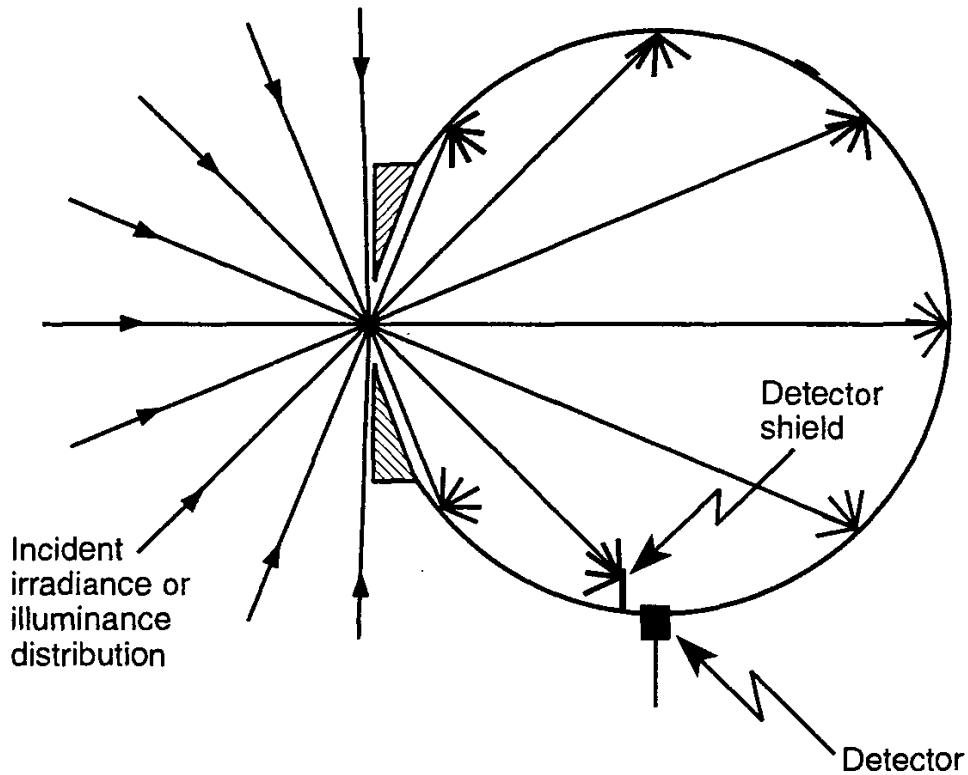


Figure 8.23 Integrating sphere and radiation detector combination used for good cosine response correction in an irradiance or illuminance meter.

materials being measured. For example, a sample could be placed over the entrance port of the sphere illustrated in Figure 8.23 and the illuminating beam could be collimated, made to contain a narrow range of incidence angles, as illustrated in Figure 8.24. The result is a measurement of the conical-hemispherical transmittance of the sample.

As indicated in the drawing, the sample is illuminated at normal (or near normal) incidence and because the sample is diffusely transmitting, radiation leaving the sample is spread over a relatively large portion of the sphere wall. Again, the detector is shielded so that it receives only flux that has been multiply reflected within the sphere. The sphere is used to collect transmitted radiation leaving the sample over a full hemispherical solid angle of 2π sr and deliver a fraction of this to the detector. If the sample exhibits little scattering, the directly transmitted beam will strike the sphere wall, causing a bright spot at this location. As is the case with the collimated beam entering the sphere without the sample present, better results are obtained in such situations if the detector is shielded from receiving radiation reflected (diffusely) from this bright spot directly.

The direction of incidence of the collimated beam on the sample illustrated in Figure 8.24 can be changed between measurements so that the conical-hemi-

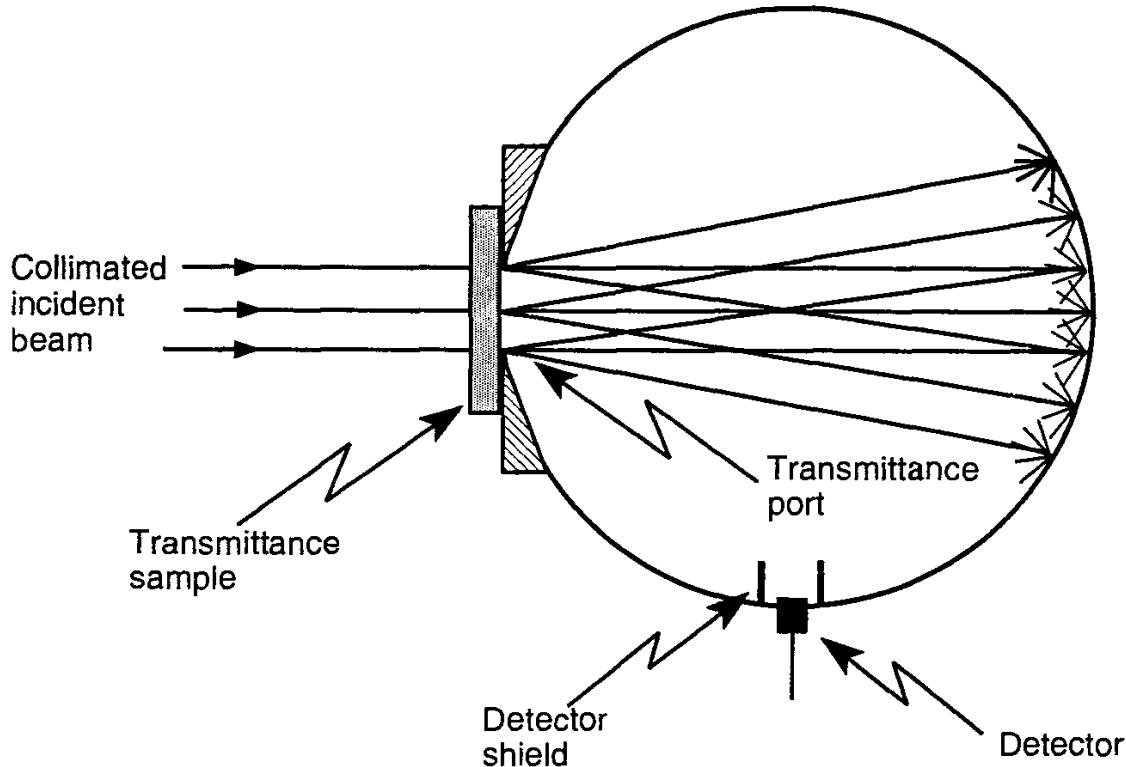


Figure 8.24 Integrating sphere and detector combination used for measuring the conical-hemispherical transmittance of a sample of diffusely transmitting material.

spherical transmittance can be obtained for a wide range of incidence directions. At large angles of incidence, however, edge effects and other problems can reduce the quality of the measurement.

Calibration is obtained with this configuration by taking a reading first without the sample in place, the "reference reading," and then another, the "sample reading," with it in place. The first reading gives the signal level for 100% transmittance. Dividing the second reading by the first gives a good approximation to the conical-hemispherical transmittance of the sample. Of course, the presence of the sample alters the reflectance of the entrance port for radiation incident upon it from the inside of the sphere. This changes the sphere efficiency and can introduce a (systematic) measurement error. (If the reflectance of the sample is known by previous measurement, however, a correction for change in sphere efficiency can be obtained through a suitable application of (8.19), thereby eliminating or at least partially compensating for this error.)

The reflectance of samples, including the one whose transmittance is to be determined, can be obtained with the arrangement illustrated in Figure 8.25. In this case, the entrance port is left open and the sample is placed in a second port, the reflectance port, immediately opposite the entrance port. Calibration in this case is obtained by taking readings first with a known standard reflectance sample and then with the sample to be measured. The ratio of the sample reading to the

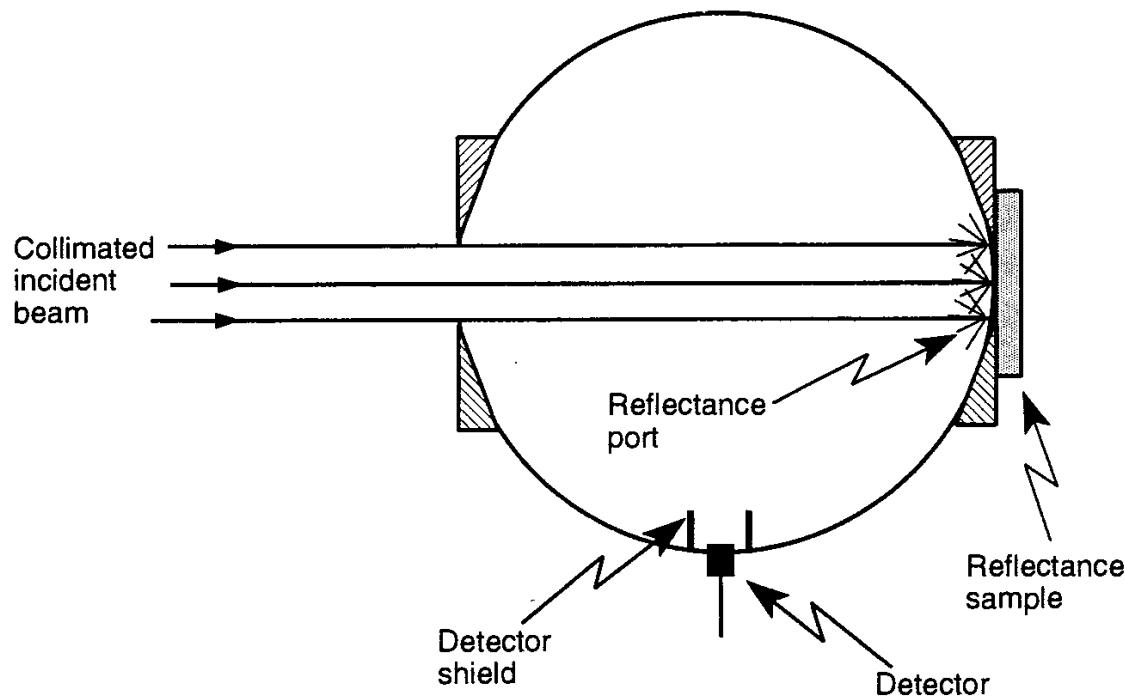


Figure 8.25 Integrating sphere arranged for the measurement of reflectance.

reference reading, multiplied by the reflectance of the standard, is close to the true conical-hemispherical reflectance of the sample. If the sample is partially transmitting, care must be taken to ensure that radiation transmitted by the sample is not reflected back through it into the sphere. It is customary therefore to place some kind of light trap behind the sample. The light trap can be as simple as a surface having very low reflectance over the wavelength range of the measurement, or as complex as a mirrored-baffled enclosure designed to prevent rays transmitted by the sample from returning to it.

The differing reflectances of the two reflectance samples will alter the efficiency of the integrating sphere for the two measurements, and a correction for this error can be made with the aid of (8.19). An extra baffle is needed to prevent the detector from receiving radiation directly reflected from the samples. Frequently the same sphere is used for both transmittance and reflectance measurements. For transmittance measurements, the reflectance port is filled with a plug whose surface reflectance and shape are made to match that of the sphere wall.

The measurements depicted in Figures 8.24 and 8.25 use what is called the *single-port substitution method*. It can be very time-consuming to make separate reference and sample measurements sequentially in time, especially if the source of radiation is a monochromatic beam from a scanning monochromator. It can take a long time to scan a full spectrum twice, once for the reference measurement and once for the sample measurement. This approach is subject to drift in the instru-

mentation. Either the source output strength or the detector and its electronic processing circuit sensitivity can drift between measurements.

One way of avoiding these problems is to use what is called the *dual-port comparison method*, diagrammed in Figure 8.26. The incident beam is split into two beams, entering two separate ports in the sphere. A reference beam passes unattenuated through the reference port and strikes the sphere wall (or a calibrated reflectance standard) opposite this port. A test beam is directed through the sample placed over another port. A reflective chopping wheel (a sectored mirror disk), is used to alternatively send first the reference and then the test beam into the sphere. The detector therefore measures first the reference signal and then the test signal a very short time later. The detection electronics keeps track of these separate signals and they are processed further to yield the transmittance. The measurement cycle happens so quickly that the source and detection systems have little time to drift between reference and test measurements. If this system is made part of a scanning spectroradiometer, the wavelength of the incident beams can be made to vary continuously during the measurement. The result is a direct plot of the spectral transmittance of the sample. Other measurement configurations can be used [28].

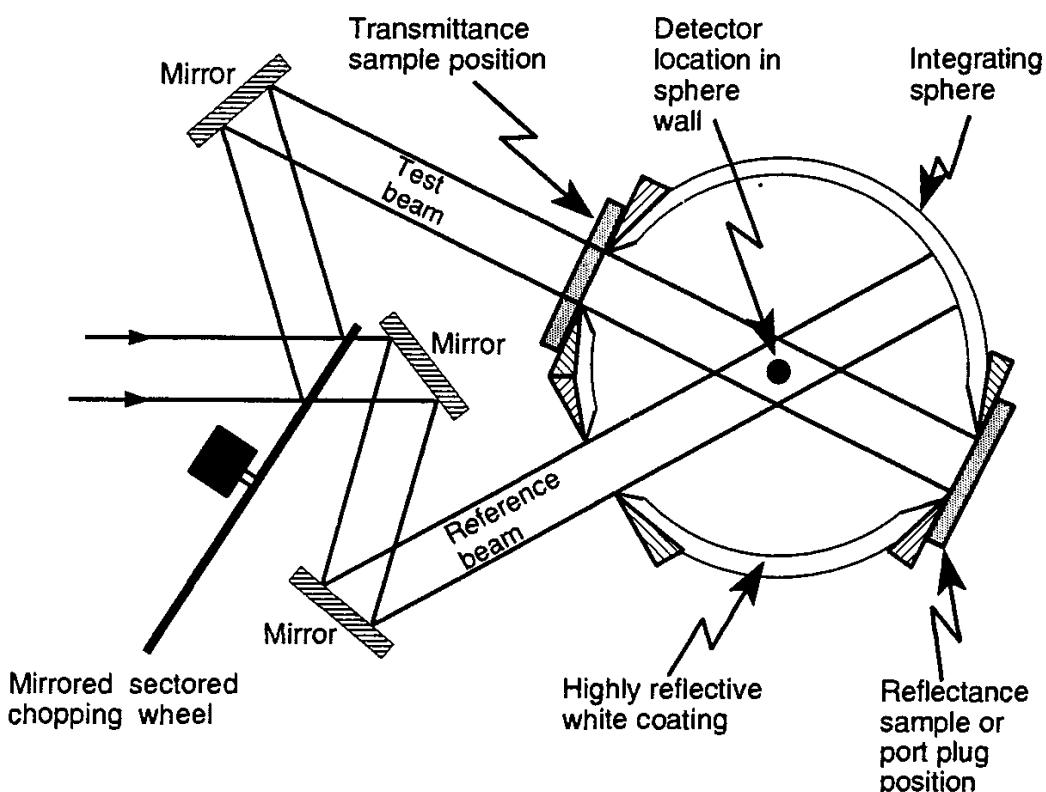


Figure 8.26 Dual-port comparison method for measuring transmittance and reflectance.

In some cases, the sectored chopping wheel has open, mirrored, and nonreflecting segments. The open segment sends the test beam into the sphere, the mirrored one sends the reference beam in, and the nonreflecting one produces nearly zero flux in the sphere, providing a dark reference to the detection electronics.

The arrangement shown in Figure 8.26 can easily be used to make reflectance measurements, by leaving the transmittance (entrance) port open and putting a sample to be measured in the reflectance port shown in Figure 8.26. This scheme does not result in an *absolute* reflectance measurement, only a relative one, since the reflectance of the portion of the sphere illuminated by the reference beam must be known.

Methods for determining the absolute reflectance of materials have been developed [29,30], but they are rather more involved than the one described above. For more comprehensive discussions of measurement methods using integrating spheres, the reader is referred to publications from Labsphere and the National Institute of Standards and Technology (formerly the National Bureau of Standards) [31–33]. Several other references to the use of integrating spheres can be cited [34–37]. The subject is discussed further in Section 9.7 on the calibration of radiometers and photometers.

For the calibration and evaluation of a variety of imaging systems, from satellite multispectral remote sensing systems to video and photographic cameras, it is frequently desirable to have a source of radiance/luminance, spatially (angularly) uniform and known over a wide angular range. There are several methods of filling the interior of an integrating sphere with spatially uniform radiation or illumination so that this radiation, upon emerging from an exit port meets the criterion. One is to place a single source of radiation of known spectral properties inside the sphere and next to the sphere wall opposite the exit port, with an opaque but highly reflective and uniform flat baffle in between so that the front of the baffle seen through the exit port receives only multiply reflected diffuse radiation of high uniformity. Another is to ring the view port with a series of sources, properly baffled so that only indirect, multiply reflected radiation escapes through the port.

Another application that takes advantage of the properties of integrating spheres is the design of instruments intended to measure the nonspecular component of radiation transmitted by sheet materials. These instruments are called *haze meters*. One way to use an integrating sphere in this method is to send collimated beam radiation through the sample and the entrance port, as illustrated in Figure 8.24, but to have an open port just opposite the entrance port to allow the specularly transmitted portion of the beam to escape the sphere and not be measured [38]. In this case, the integrating sphere collects only the diffuse, non-specular portion of the transmitted radiation, delivering a fixed fraction of it to the detector. Through proper design and operation of these devices, the relative magnitudes of the diffusely scattered and unscattered directly transmitted com-

ponents can be determined for samples of translucent and transparent sheet materials.

8.12 MONOCHROMATORS

It is frequently important in optical systems intended for radiometric or photometric measurement to perform measurements over restricted wavelength intervals or over very narrow intervals that are then scanned through a spectral range of interest.

There exists a number of optical elements and subsystems for selecting out a limited portion of the spectrum of an incident beam for further propagation through an optical system. Several of the most prominent of these, called *monochromators*, are now described.

8.12.1 Spectral Filters

Probably the simplest spectrally selective device available is the colored-glass filter, consisting of a parallel plate of clear glass containing imbedded molecules that impart to the natural transmittance spectrum of the clear glass a strong variation in spectral transmittance over a wavelength range of interest. These filters are good for isolation of broad spectral regions. Many have more than one spectral passband. Rejection of the unwanted passband can frequently be achieved by judicious choice of a source whose output over the unwanted wavelength region is low or nonexistent or the use of a detector whose spectral response rejects the unwanted wavelengths. Alternatively, one can use two filters in series to reject unwanted wavelengths while passing the desired ones.

According to an Oriel Corporation catalog [39], two processes are involved in spectrally selective transmission with colored-glass filters: ionic absorption and colloidal scattering. At room temperature, glass is a (very) viscous liquid and can therefore hold ionic materials in "solution." Some common compounds used for this purpose are nickel oxide, cobalt oxide, chromium oxide, and ferric oxide. The absorption of radiation as it propagates through such solutions follows Beer's law, (6.28) described in Section 6.4. The application of Beer's law to a parallel plate of material is described in Section 6.5. The absolute magnitude of the spectral transmittance of a colored-glass filter can be adjusted with minimal change in its shape by changing the thickness of the filter.

With colloidal scattering, inorganic elements or salts are thermally produced to form microscopic crystals imbedded in the glass that scatter and absorb certain wavelengths while allowing others to pass unattenuated. Common inorganic materials used include sulfur, cadmium sulfide, cadmium selenide, and gold.

A large number of optical companies manufacture colored-glass filters and several offer catalogs with listings for a variety of glass filters along with the trans-

mittance spectra of each. One of the best catalogs, because of the detailed tutorial information relating to its products, is the one published by Oriel Corporation [39]. A limited selection of companies that offer spectral filters of many kinds is provided in Table 8.1. Many more can be found in the Physics Today Annual Buyers' Guide [40].

A sampling of a few assorted filter-glass spectral transmittance curves from the Oriel catalog are plotted in Figures 8.27 through 8.29. Oriel catalog numbers are indicated for each curve. The spectral transmittances shown are due only to internal optical processes and do not include reflection losses at the air/glass interfaces.

Some of the substances used in making glass filters can produce reasonably sharp transitions between high and low transmittances, but narrow bandpass filters, with spectral transmittances as sharply defined as the one shown in Figure 8.30, are not generally available at a variety of wavelengths with this type of filter. The width of a spectrally selective filter's passband is frequently defined in terms of the spectral range between values of the transmittance that are 50% of the peak spectral transmittance. This *spectral width at half maximum* is illustrated in Figure 8.30. The passband width is said to be equal to $\Delta\lambda = \lambda_2 - \lambda_1$, with λ_1 and λ_2 being the wavelengths where the spectral transmittance is half its maximum value.

In order to achieve narrow bandpass spectral transmittance, it is usually necessary to use multilayer or composite filters. The latter are made of two or more separate filters sandwiched together. The total spectral transmittance of two filters combined in this manner has approximately the shape of the mathematical product of the spectral transmittances of the two filters, on a wavelength-by-wavelength basis. By putting a high-pass filter together with a low-pass filter one can produce either a bandpass or a band exclusion filter.

Table 8.1
Spectral Filter Manufacturers

Oriel Corporation, 250 Long Beach Blvd., P. O. Box 872, Stratford, CT 06497

Schott Glass Technologies, Inc., 400 York Ave., Duryea, PA 18642

Hoya Optics, Inc., 3400 Edison Way, Fremont, CA 94538-6190

Acton Research Corp., P. O. Box 2215, 525 Main St., Acton, MA 01720

Barr Associates, Inc., 2 Lyberty Way, P. O. Box 557, Westford, MA 01886

Andover Corp., 4 Commercial Dr., Salem, NH 03079

Melles Griot, Inc., 1770 Kettering St., Irvine, CA 92714

Graseby Optronics, 12151 Research Parkway, Orlando, FL 32826-3207

Guernsey Coating Laboratories, 4464 McGrath St., Unit 107, Ventura, CA 93003

Optics for Research Inc., P. O. Box 82, 315 Bloomfield Ave., Caldwell NJ 07006

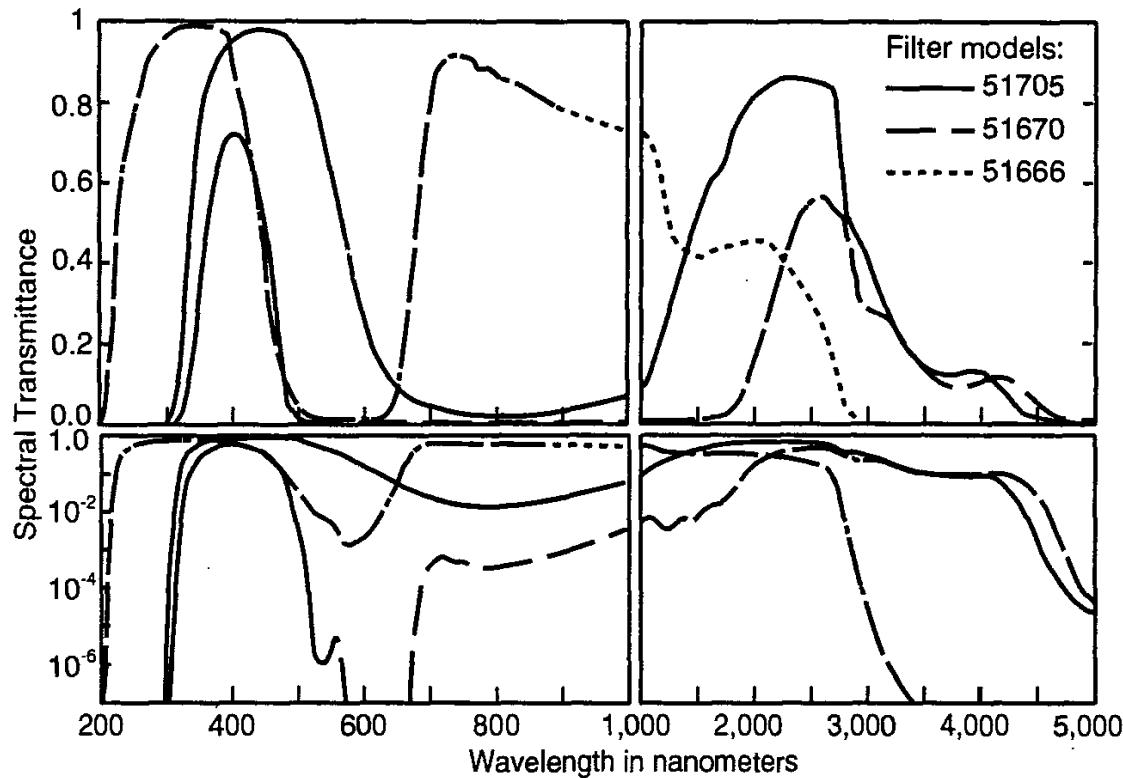


Figure 8.27 Spectral transmittances of selected UV-VIS bandpass glass filters. (Courtesy Oriel Corporation.)

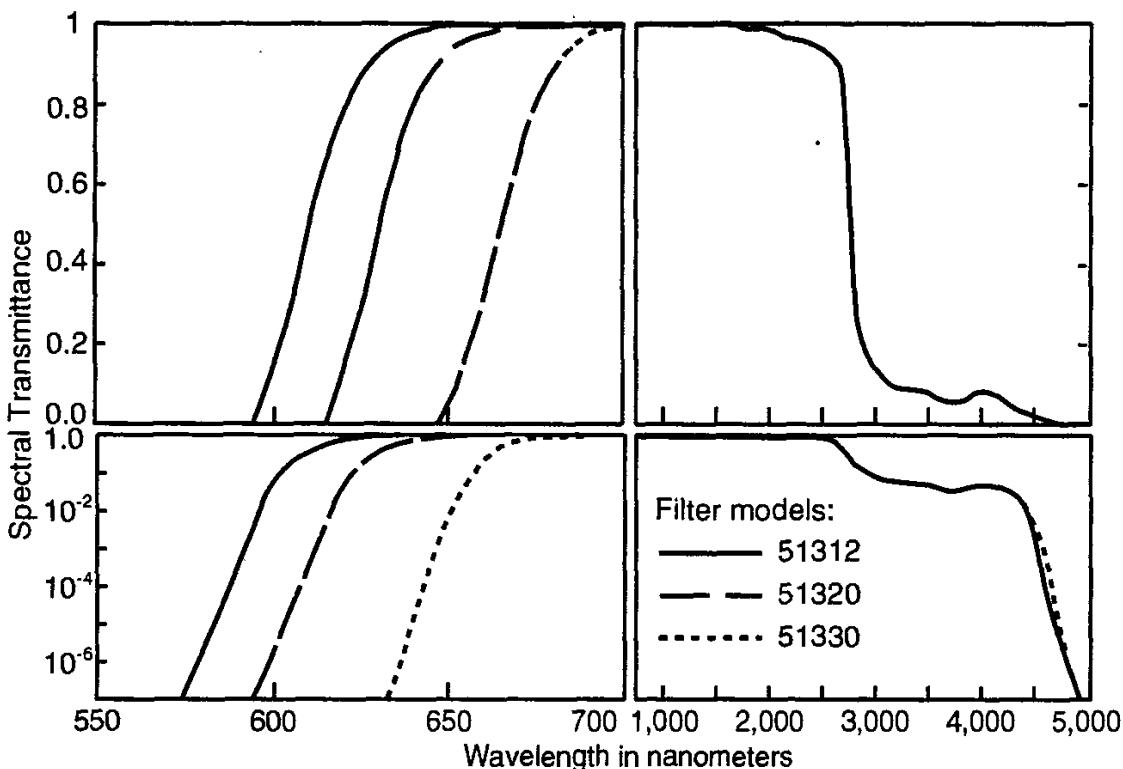


Figure 8.28 Spectral transmittances of selected visible long-pass glass filters. (Courtesy Oriel Corporation.)

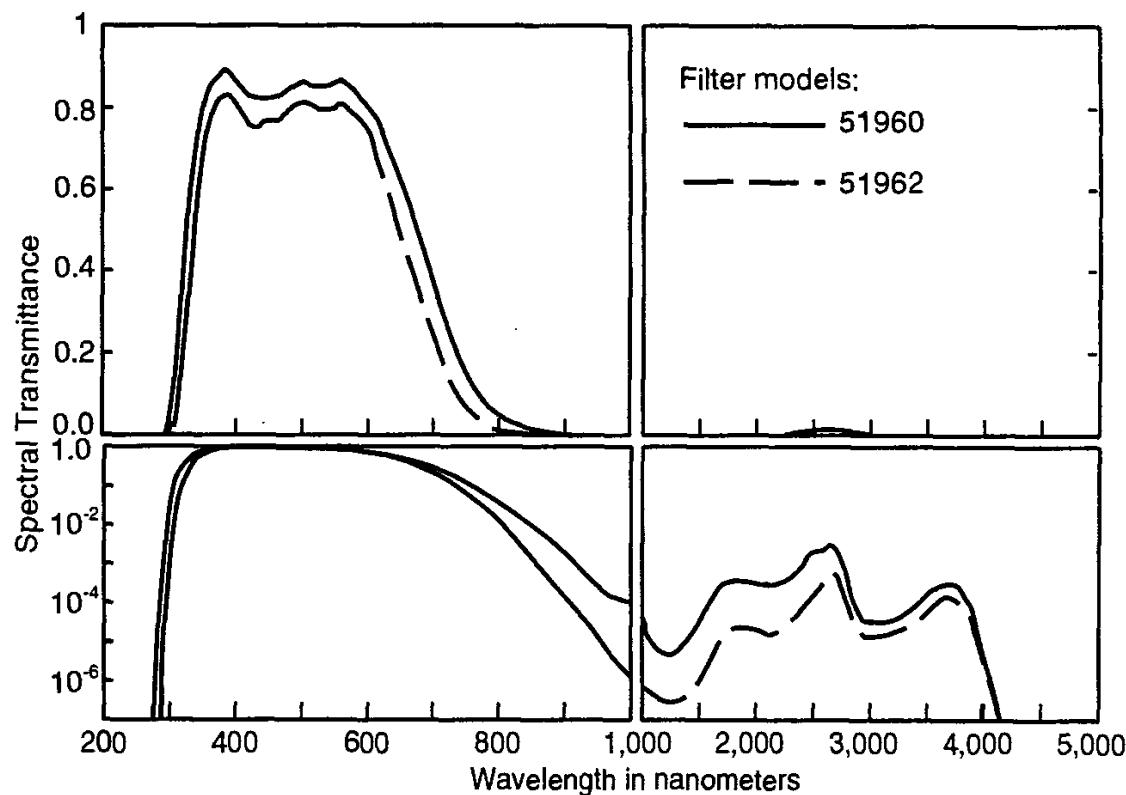


Figure 8.29 Spectral transmittances of selected IR blocking glass filters. (Courtesy Oriel Corporation.)

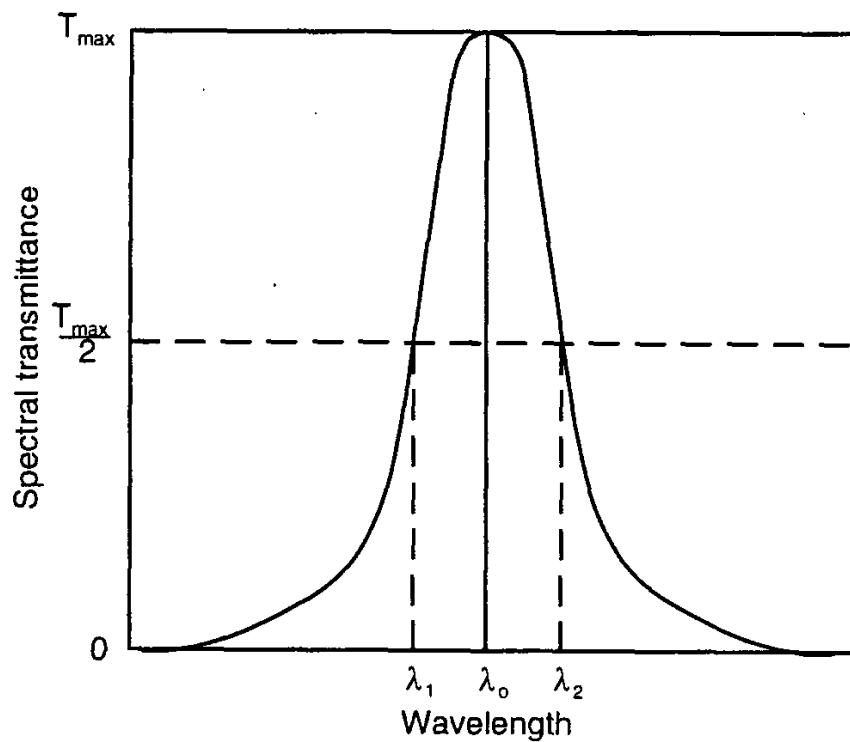


Figure 8.30 Plot of the transmittance of a filter showing the spectral width at half maximum. $\Delta\lambda = \lambda_2 - \lambda_1$

For extremely narrow bandpass filters, and those for which the width and center wavelength can be specified arbitrarily, a different technique is used. The *multilayer interference filter* is composed of a number of dielectric and/or metallic layers deposited in carefully controlled thin-film thicknesses on a piece of clear or colored glass. Through the phenomenon of destructive and constructive interference, the shape of the spectral transmittance curve of the filter can be controlled almost arbitrarily.

As the number of layers in a multilayer interference filter increases, the shape of the spectral passband generally improves. Many of these filters are made of metallic film "cavities" spaced apart by clear dielectric layers. A two-cavity filter, for example, might have a seven-layer dielectric reflector sandwich that is 1/4 wavelength thick, separated by a half-wavelength dielectric spacer, followed by another pair of dielectric-spaced reflectors, using a coupling layer to hold the two cavities together [39]. Figure 8.31 illustrates the effect on bandwidth and band shape of increasing the number of "cavities" in an interference filter.

Because their principle of operation depends on physical spacings accurate to a fraction of a wavelength, interference filters are sensitive to temperature changes. Too rapid a temperature change or too great a temperature (typically

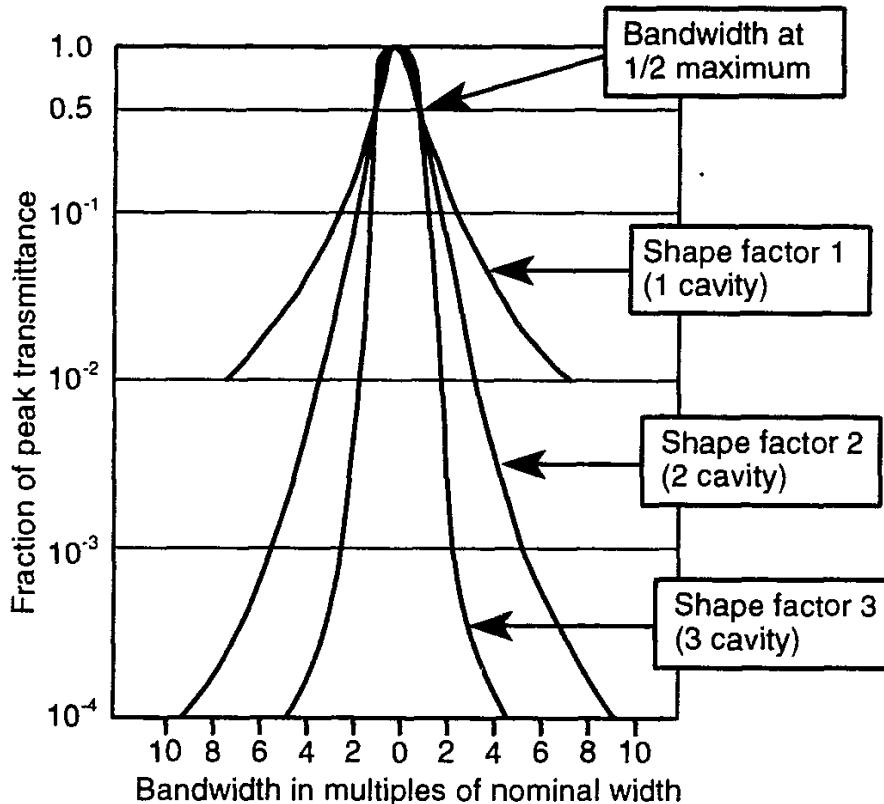


Figure 8.31 The spectral passband shape for an interference filter having one, two, and three cavities. (Courtesy Oriel Corporation.)

exceeding 100°C) can produce permanent changes or even destruction of the filter. Within temperature limits, permanent damage may not be produced, but a change in the transmittance and central wavelength generally occurs, typically at a rate of from 0.01 to 0.03 nanometers per °C, for the movement of the central wavelength. It is important to seal the edges of interference filters to protect their layers from moisture effects, otherwise permanent changes in filter optical properties can result.

Due to the multiple-reflection interference effects upon which interference filters depend, their spectral transmittances depend upon the angle of incidence of the radiation passing through them. A representative change in filter spectral transmittance for illumination at a 20-deg angle of incidence is shown in Figure 8.32.

Interference effects can be used not only to produce narrow spectral passbands but also to make high and low-pass filters. Characteristic spectral transmittance plots for high and low-pass filters are illustrated in Figures 8.33 and 8.34, respectively.

Oriel Corporation offers a line of what are called "holographic notch filters" with high transmittance at wavelengths above and below a narrow range over which the filters are essentially opaque. These filters are characterized by very sharp spectral transitions at the edges of their spectral rejection bands. They are ideal

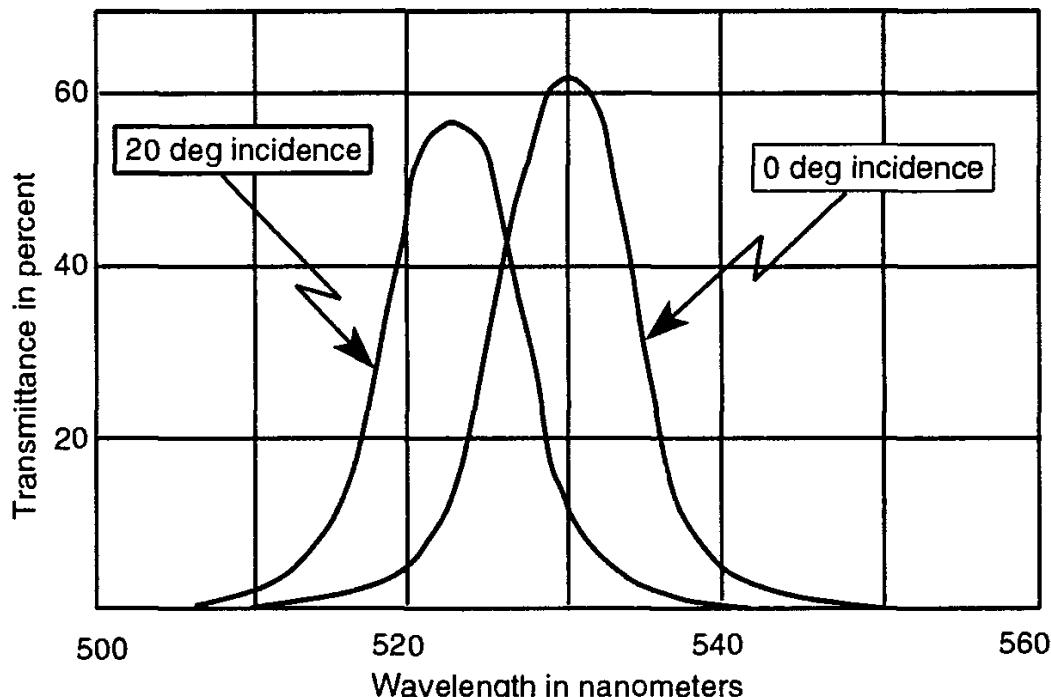


Figure 8.32 Change of interference filter transmittance with angle of incidence. (Courtesy Oriel Corporation.)

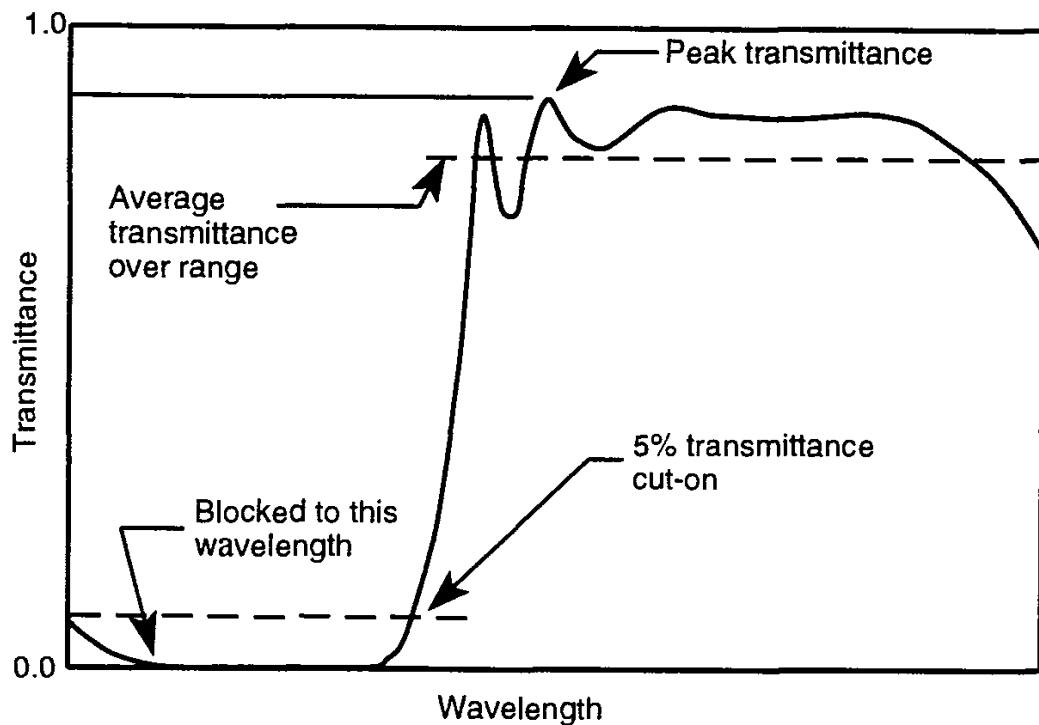


Figure 8.33 Typical long-pass interference filter spectral transmittance. (Courtesy Oriel Corporation.)

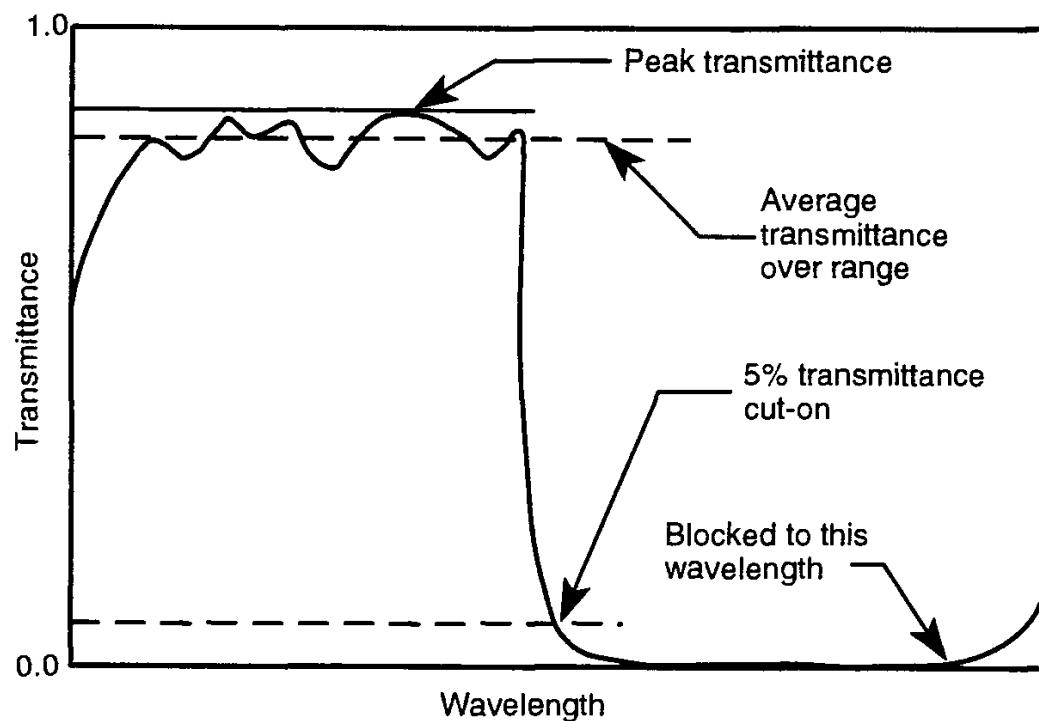


Figure 8.34 Typical short-pass interference filter spectral transmittance. (Courtesy Oriel Corporation.)

for blocking monochromatic radiation from a laser while passing fluorescence radiation or other radiation stimulated by the laser beam outside the blocking bandwidth of the filter.

One of the most remarkable extensions of interference filter technology is the *circularly* or *linearly variable interference filter*. With these devices, the spectral passband varies in wavelength with angle around a center point in the former case or linearly over distance across the filter in the latter case. These devices can therefore form the basis for an inexpensive scanning monochromator having a not very narrow spectral passband. A small beam to be rendered monochromatic is passed through the filter and the transmitted wavelength is made to scan through a range by rotating or translating the filter relative to the incident beam.

Many optical elements, including interference filters, are sensitive to long-term exposure to UV radiation. Generally, the shorter the wavelength of the incident radiation, the greater the sensitivity. Permanent changes in the optical properties of spectral filters can result from UV exposure, even to those filters designed to operate in the UV portion of the spectrum.

8.12.2 Scanning Monochromators

If greater spectral resolution (narrower spectral passband) with scanning range flexibility is needed, then more powerful means must be employed than are available with multilayer interference filters. A detailed description of the design and operation of all the various kinds of scanning monochromators is beyond the scope of this text, so only a limited introduction can be provided here. For more information the reader is directed to modern textbooks on optics, photonics, optical engineering, and optical instruments. Volume V of Kingslake's series on Applied Optics and Optical Engineering has substantial sections on scanning monochromators [41].

Most scanning monochromators employ some optical element that disperses an incident beam of radiation into separate beams, one for each wavelength present, spreading these beams angularly. A limiting aperture is then placed in the dispersed region, selecting a limited wavelength range for passage through the aperture and out of the monochromator. By changing the location of this aperture relative to the spectrum (or vice versa), the wavelengths permitted to leave the monochromator can be made to vary.

8.12.2.1 Prism Monochromator

The prism monochromator accomplishes this feat by virtue of the dispersion in refractive index of the transparent material from which the prism is made. Typically,

the beam to be scanned spectrally is focused onto a very small entrance aperture, either a slit or a pinhole, as illustrated in Figure 8.35. After passing through this aperture into the instrument, the beam diverges to collimating lens L_1 or mirror. The collimated beam then enters the prism, where it is refracted toward the normal to the entering face of the prism. The angle of refraction depends upon wavelength. The chromatically spreading beam then passes through the second face of the prism, back out into the air again, where further angular spreading takes place.

The beam then enters another lens L_2 or mirror, usually a duplicate of the first one, which focuses the beam onto an exit plane, a photographic plate or detector array, an aperture, a slit, or a pinhole, permitting the spectral range to be recorded, or a selected portion of it to emerge from the spectrometer.

Due to the spreading over angles of different wavelength replicas of the beam emerging from the prism, a spectrum of colors (in the visible) is spread across the exit plane. What would normally be a single point or line image of the entrance aperture is now a series of such images spread laterally across the exit plane, one image for each wavelength present in the source.

Because of this spreading, only a very narrow range of wavelengths in the original beam is allowed to pass through an exit aperture and out of the spectrometer. In this mode of operation, the device has therefore produced an essentially monochromatic beam of diverging radiation to be either detected or sent on to another optical system. For example, the emerging monochromatic flux can be

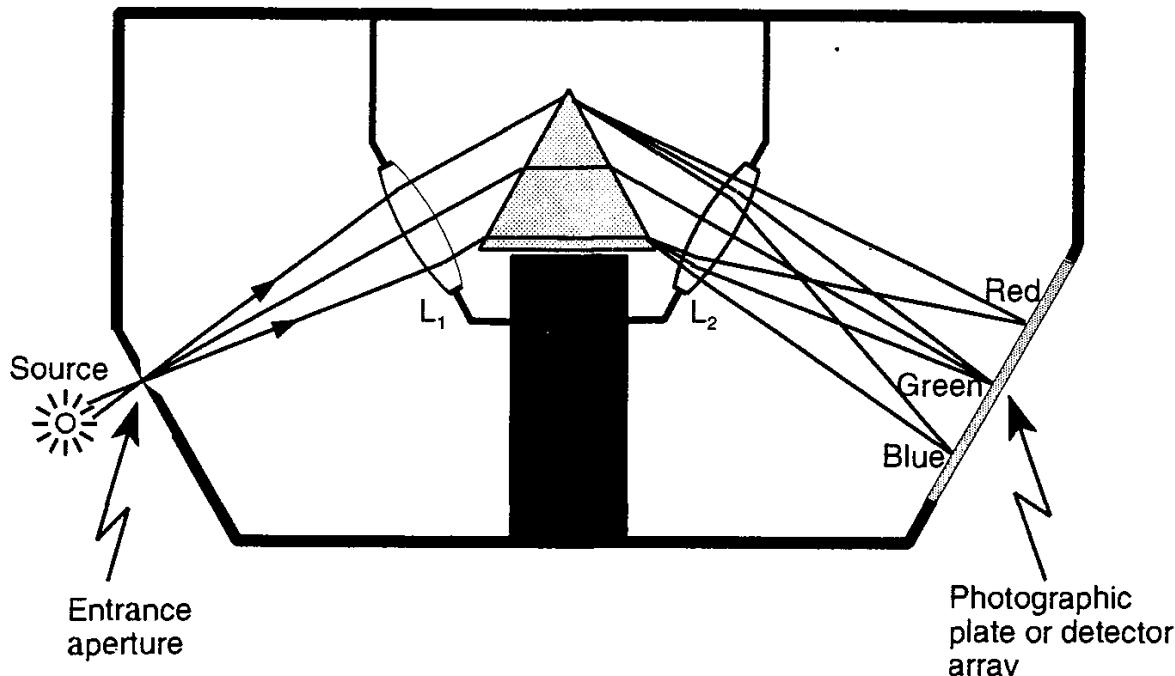


Figure 8.35 Illustration of the principle of operation of a prism spectrometer.

made to illuminate a sample in an integrating sphere transmissometer. By changing the relative positions of the prism, the exit aperture, and the focusing lens, the spectrum can be made to move laterally across the exit aperture, thus changing the wavelength emerging from the instrument.

Because refractive dispersion is not *linearly* proportional to wavelength, the correspondence between wavelength and angle is of course not a linear one. Use of a prism spectrometer requires a careful wavelength calibration process in which radiation from a source containing emissions at specific and known narrow wavelengths is sent through the spectrometer and the angular (or linear) positions of the known wavelengths are recorded. These values are then fit to a curve and the curve is subsequently used to convert angles of emergence from the prism (or positions across the plane) into wavelengths. An advantage of the prism spectrometer is its high efficiency. Another is the fact that only one wavelength can emerge from the monochromator for each of its angular settings, an advantage not shared by the grating monochromator, described in the next section. The version shown in Figure 8.35 is for what is called a *spectrograph*, the name resulting from the fact that a photographic plate or detector array is used in the exit plane to record the full spectrum all at once.

8.12.2.2 Grating Monochromator

The grating monochromator uses a ruled diffraction grating to accomplish the dispersion of beams incident upon it. For more information on the construction and theory of diffraction gratings, consult any textbook on physical optics. Gratings can be constructed by physically marking parallel grooves in the surface of a polished mirror, by depositing opaque parallel lines on a transparent substrate, or by exposing a photosensitive medium to a laser interference pattern and then developing the medium. The latter fabrication process produces gratings that are sometimes called “holographic diffraction gratings.”

The operating equation for a grating is

$$m\lambda = d \sin \theta \quad m = 1, 2, 3, \dots \quad (8.20)$$

where d is the spacing between the groove centers or openings in the diffraction grating and θ is the angle of diffraction, the angle to the m th diffraction maximum for wavelength λ . The fact that m can be any of the positive integers may be confusing at first. It results from the interference effects upon which the basic operation of the grating depends. The integer m gives the number of path differences between adjacent grooves measured in wavelengths λ . The letter m is called the *order of diffraction*. For a given wavelength, the first order will be diffracted

at the angle given by solving (8.20) for θ using $m = 1$ and an appropriate value for d . For equality to be maintained, if m is increased to 2 and d and λ remain the same, then the sine of the angle of diffraction must be double what it was before. Thus, the grating produces first, second, and third-order spectra, each displaced from the former in angle. The design of a monochromator based on this principle is illustrated in Figure 8.36.

From elementary trigonometry, the distance y across the plane of the exit aperture, called here the spectrum plane, is equal to $L \tan \theta$, with L being the distance from the grating to the exit plane. It can be shown that the $\sin \theta$ in (8.20) can be replaced with $y/\sqrt{L^2 + y^2}$, so that (8.20) becomes

$$m\lambda = \frac{yd}{\sqrt{L^2 + y^2}} \quad (8.21)$$

For relatively small angles θ of diffraction, y^2 will be much less than L^2 in (8.21), and we can write

$$m\lambda \approx dy \quad m = 1, 2, 3, \dots \quad (8.22)$$

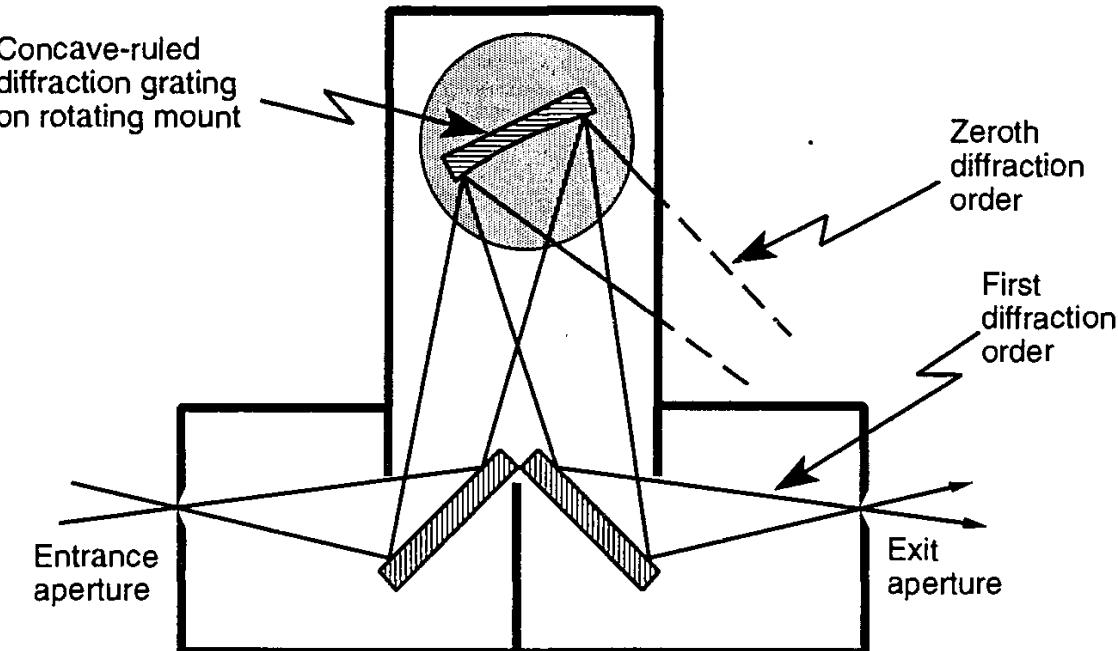


Figure 8.36 Illustration of a diffraction grating monochromator.

showing the approximately linear relationship between wavelength λ and distance y for a given order of diffraction.

Most commonly available grating monochromators employ a wavelength scanning mechanism that compensates for the $\sin \theta$ relationship shown in (8.20) between the angle θ of diffraction and the wavelength λ . With these monochromators, the rotation of the screw driving the grating rotation is made to be linearly proportional to the sine of twice the angle of rotation of the grating, and, hence also to the wavelength. This linear relationship between drive screw rotation and wavelength makes wavelength calibration much simpler for this type of monochromator.

When $m = 0$ in (8.20), the incident beam is reflected without dispersion according to the law of reflection. This $m = 0$ reflected beam is called the central maximum because $\theta = 0$ and it is generally brighter than any of the higher orders of diffraction. It is possible to redistribute the flux, putting more flux into the first order on one side than into any other order. This is accomplished with a process called *blazing* of the grating. The grooves that are normally parallel to the substrate of a reflection grating are shaped into a saw-tooth triangular pattern. By controlling the tilt of the pattern, radiation can be reflected toward any of the orders of diffraction, limited only by the geometry and the fabrication method.

If the longest of the wavelengths in the beam incident on a grating is no greater than twice the length of the shortest wavelength present, then the first and second orders of diffraction will be separated from one another in angle and two separate replicas of the incident spectrum will be spread over the plane of the exit aperture. The grating can be rotated to place any wavelength in either of the orders on the exit aperture. It turns out that wavelengths in the second order are spread twice as much as those in the first order. There is twice as much dispersion in the second order as in the first. The problem is that the reflectance of the grating is generally less in the second order than in the first. Blazing the grating for the second order can partially overcome this problem. Greater spectral resolution can be achieved for the same exit aperture width in the second order than in the first, but less radiation will emerge from an aperture of fixed size.

Whenever the range of wavelengths present in a beam of radiation entering a grating monochromator exceeds a factor of two (i.e., when the longest wavelength present is greater than twice the shortest wavelength), a problem peculiar to grating instruments is encountered. From (8.20) it is seen that radiation at wavelength λ in the first order is diffracted at the same angle θ as radiation at wavelength $\lambda/2$ in the second order. Special filters, called *order-sorting filters*, are employed in this case to block the range of unwanted wavelengths from entering the monochromator.

Gratings are widely used as the only or primary dispersing element in modern monochromators, because greater angular dispersion can be achieved with grating instruments than with prism instruments. Frequently two grating monochromators

are used in series, the first serving as a prefilter for the second, for stray-light rejection purposes.

8.12.2.3 Spectrometers

Suppose a monochromator is placed in an optical system with a source and/or a radiation-gathering subsystem, and the emerging flux is directed to photographic film, to a detector array in the spectrum plane, or to an eyepiece for visual observation. The resulting instrument is called a *spectrometer* since it can be used to measure (or “meter”) the spectral positions of features of interest in the spectrum over a spectral range of interest. A spectrometer can be made to display the whole spectrum all at once or to scan it in time, recording flux levels one wavelength at a time.

Historically, spectrometers were designed primarily to determine the presence or absence (and relative strengths) of flux from gaseous sources stimulated electrically to emit light at discrete wavelengths in the spectrum. The initial emphasis was primarily on determining the wavelength of these discrete emissions, called *spectral lines*, not so much on the radiometric properties of the emissions. They were called spectral lines because vertical slits were used for the entrance aperture and images of the entrance slit were displayed in the spectrum plane for each wavelength present in the source radiation. Since these images looked like narrow lines for monochromatic emissions, they were called spectral lines. A photographic film could record a complete range of the spectrum at one exposure. Finding the wavelength positions of spectral lines yielded information about the atoms and molecules present in the emitting gases. A spectrometer designed to record a whole spectrum at one time is called a *spectrograph*. The detector in such cases is typically a photographic emulsion or a linear array of photodiodes. One advantage of the spectrograph is that for weak sources of radiation, such as distant stars, the exposure time can be lengthened and the quantity of radiant energy absorbed by the emulsion or accumulated by the detector array at each wavelength can be increased.

A spectrometer designed and calibrated primarily for measurements of spectral radiant flux is called a *spectroradiometer* since it is intended to serve as a radiometer for each wavelength of interest in the spectrum, normally performing its measurements using only one wavelength at a time. Spectroradiometers can be handheld instruments, with the wavelength of measurement set by hand, or versions that automatically scan a wavelength range. Complete measuring instruments of this type are discussed in the next chapter.

It is customary to use the term *spectrophotometer* to indicate a laboratory scanning spectrometer designed to measure one or more spectral optical properties

of gaseous, liquid, or solid samples.¹ The CIE has standardized this terminology as follows:

A spectroradiometer is an instrument for measurement of radiometric quantities in narrow wavelength intervals over a given spectral region.

A spectrophotometer is an instrument for measurement of the ratio of two values of a radiometric quantity at the same wavelength.

Spectrometers, spectrophotometers, and spectroradiometers can use a prism or a grating as a dispersing element. Some use circular variable interference filters as the monochromator. It was mentioned previously that, with grating monochromators, incident wavelength ranges covering more than a factor of two will produce some overlap of the first and second orders. For example, radiation at 600 nm in order 1 will be diffracted at the same angle θ as radiation at 300 nm in order 2. To prevent measuring radiation at these two wavelengths simultaneously, the instrument must be equipped with some other dispersing or filtering element that can eliminate the 300-nm radiation when the 600-nm radiation is being measured and vice versa. Order-sorting filters are commonly used to accomplish this feat. Another technique is to predisperse the incident spectrum with a prism at low spectral resolving power and then to follow this with a grating instrument to further disperse the portion of the spectrum of interest.

Other types of spectrum dispersing devices have been developed, including a variety of interferometers. For more information on this subject the reader is referred elsewhere [42].

8.13 WINDOWS

Clear windows of optical glass are frequently needed to protect optical components and/or detection systems from exposure to the outside air and humidity or to contain liquid or gaseous samples. The spectral transmittance of such windows is important

¹Calling a spectrometer that measures at each of a number of discrete wavelengths a spectrophotometer would seem to be a misnomer, since most spectrophotometers have nothing to do with the human eye response. I prefer the terms "spectrotransmissometer" or "spectroreflectometer" for spectrometer-based instruments, and a complete abandonment of the term spectrophotometer. However, the current usage of this term for laboratory spectrotransmissometers and spectroreflectometers is so widespread that the name is unlikely to be discontinued in the foreseeable future. An unfortunate consequence of this usage is that it leads some spectrophotometer manufacturers to speak of "photometric accuracy and reproducibility" when spectral absorptance or transmittance at one or more specified wavelengths both within and outside the visible spectrum is intended. Such misuse of the term "photometric" should be avoided.

in the design of radiometric and photometric instruments. The ultimate objective is to have the window produce negligible alteration of the beam of flux passing through it. In some cases it is possible to place a thin layer (generally of MgF₂) on the window to greatly lower its reflectance over a range of wavelengths, increasing the overall transmittance of the window. This is called an *antireflection coating*. Common glasses used for windows (and lenses) include Optical Crown, BoroSilicate Crown (BSC), BK-7, Pyrex, and IR and UV-grade fused silica (quartz). More exotic window materials, needed for extended spectral ranges, include calcium fluoride, sodium fluoride, zinc selenide, sapphire, and diamond. These materials are characterized by constant spectral transmittance between two widely separated wavelengths and either sharp or rounded fall-off in transmittance beyond these wavelength limits. UV-grade fused silica, for example, has useful transmittance (above 60%) extending from 170 nm to 2.2 μm and flat transmittance from 200 nm to 1.2 μm . Magnesium fluoride has the amazing characteristic of offering essentially uniform transmission from 200 nm to 4 μm . Usable transmittance (above 60%) extends down to 150 nm and out to 7 μm for this material. Spectral transmittance curves for these and other optical materials can be found in the Oriel Volume III catalog of optics and filters [39] and in the catalogs of other optical component manufacturers.

8.14 SOURCES

A brief survey of the spectral distributions of various sources of electromagnetic radiation is presented in Section 3.9. There are many different kinds of sources, including incandescent filament lamps of many kinds, short arc lamps, gaseous and solid-state lasers, fluorescent lamps, high-intensity gaseous discharge lamps (including those using low and high-pressure sodium and mercury), electroluminescent lamps, light emitting diodes (LEDs), carbon arc lamps, nuclear sources, and a variety of naturally occurring sources. The latter include solar radiation and sky radiation, radiation from other planets and distant stars, solar radiation reflected from the moon, and radiation emitted by several plant and animal species.

The important characteristics of sources for use in radiometric and photometric systems are the spectral distribution of emitted flux, the size and shape of the emitting element or elements, the spectral transmittance and other properties of any transparent envelope that may be used, the electrical characteristics of the lamps and of other electrically powered sources (including the requirements needed for electrical power supplies to provide a stable supply of electricity to the lamp), estimates of the degradation in output over time (aging), and the directional distribution of the emitted radiation. There are requirements to be considered for the proper operation of radiation sources. They include forced air or liquid cooling,

human eye and skin protection, and means for keeping the glass envelope free of contaminants such as oil and grease.

In 1992, the CIE published a summary of the state of the art of electric light sources [43], primarily for use in illumination engineering. For more information concerning these and other sources of electromagnetic radiation, the reader is referred to an Oriel Corporation catalog [44] and the IES Lighting Handbook Reference Volume [45], both of which contain excellent discussions of the characteristics of a wide variety of visible, ultraviolet, and infrared sources.

8.15 GONIOMETERS

A goniometer is a device placed on a circular track or made to follow a circular arc path in order to "look" over a range of angles at an object located at the center of the circle upon which the track is based. A *gonioradiometer* is a radiometer incorporated into a goniometer for measuring the flux emanating from a light source or reflector placed at the center of the circle. A goniophotometer is a photometric version of a gonioradiometer. Goniophotometers are frequently used to determine the angular distribution of luminous intensity from a light source such as an electric lamp designed for use as a street light. A goniophotometer is illustrated generically in Figure 8.37. Sometimes goniometers are made to track over an entire hemisphere. The radius of the arc of a goniometer can be small or large, depending on the size of the sample being examined. With some large goniophotometers, the photometer is fixed and a rotating mirror is used to direct light from a circular path surrounding the source to the photometer. After one complete revolution,

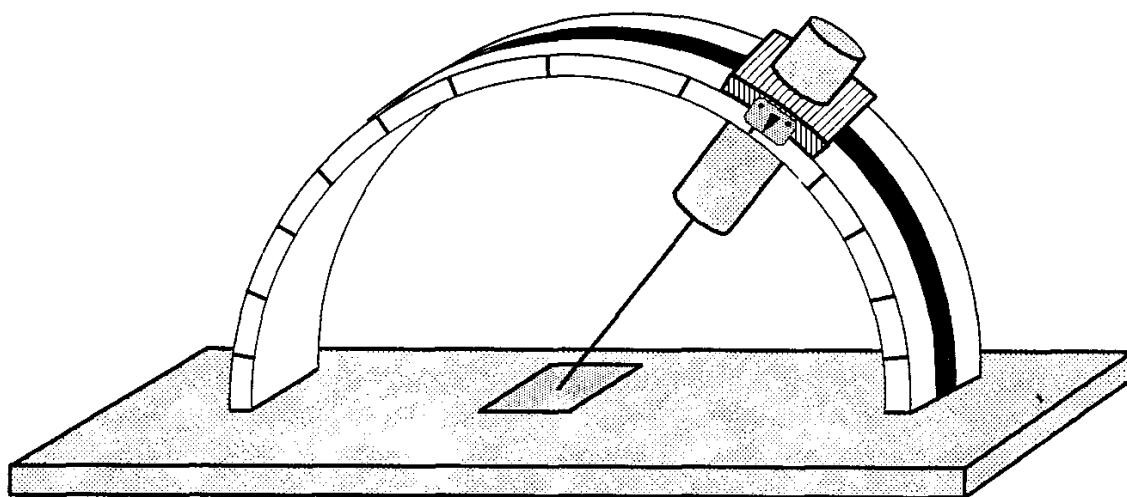


Figure 8.37 Schematic illustration of the operation of a goniometer for radiometry and photometry.

the angle of the mirror is changed, the position of the circle is shifted, and another circular scan is made. A sequence of these measurements is used to determine the full directional distribution of source intensity over a nearly spherical solid angle.

8.16 TRANSMISSOMETERS/REFLECTOMETERS

The considerable range of transmittance and reflectance measurement geometries was surveyed in Section 6.6 for parallel plate samples of modest thickness. The use of integrating spheres for measuring diffuse transmittance and reflectance was discussed in Section 8.11.2. The variety of instruments for measuring transmittance and reflectance is as great as the variety of measurement geometries. Sometimes it is only the normal transmittance and near-normal reflectance that are of interest, and these include specular and nonspecular (biconical or conical-hemispherical) cases. In other cases, the angular variation of the properties is desired and the measuring instrument has to be designed very carefully to provide accurate measurements over a wide range of angles while maintaining good calibration in the process.

In some cases, very thick samples of the material whose transmittance is being measured are needed, such as the case when the transmittance of fairly clear natural water is desired. It is not uncommon to have an optical path in water of one meter in length. Great care must be taken with such instruments to avoid the collection of multiply scattered radiation or to account for this radiation in the design and/or calibration of the instrument [46].

8.17 SCATTERING METERS, NEPHELOMETERS, TURBIDIMETERS, AND HAZE METERS

Sometimes it is important to characterize the quantity or angular distribution of radiation scattered out of a specular beam on transmission or reflection through or from a sample. The sample can be gas, liquid, or solid. When it is a gas or liquid, it is usually the scattered flux per unit volume of the material that is important, without reference to the effects of the material's boundaries. Scattering effects are discussed in Sections 6.4 and 6.5.2. Instruments to perform measurements of the scattering properties of materials are now described.

Nephelometer is the name given to scattering meters that measure the volume scattering function β for a particular angle of scattering, for a specified range of such angles, or for each of a number of angles over a range. Generally the scattering medium is larger than the incident beam. The scattering volume is the intersection of the beam illuminating the medium with the detection beam, the solid angular field of view into the medium of the detector. A baffled tube or telescopic imaging system is placed in front of the detector to define a view geometry [47]. If the angle

of scattering varies, any changes in the scattering volume must be accounted for in the calibration procedure.

Turbidimeters are devices for determining the relative concentration of scattering particles in different media from measurements of light scattered from those different media. They generally measure the flux scattered at right angles to the incident beam. Different turbidimeters measure different optical properties. A survey of water turbidity definitions and instrumental methods [46] showed great variety in this field. It is recommended that the term *turbidity* be reserved as a subjective or relative descriptor of the level of perceived cloudiness in a suspension and that turbidimeters report the specific optical property being measured. Converting a given value for the volume scattering function at a specific angle of scattering or for the transmittance or extinction coefficient of a scattering medium to an estimate of the particulate concentration generally involves independent knowledge of the size distribution, shapes, and optical properties of the suspended particles causing most of the scattering [48–51]. This point of view was reflected in the report of a meeting sponsored by NOAA's National Oceanographic Instrumentation Center on water turbidity measurements [52].

Haze meters are instruments intended to determine the degree of nonspecular transmission from a transparent plate of solid material or the nonspecular reflection from a solid mirror. The method generally employed is to illuminate the sample with a collimated beam at normal incidence and then to compare the radiant flux scattered into a hemispherical solid angle from the sample with the flux in the specularly transmitted or reflected beam. The scattered portion is usually measured with an integrating sphere with ports for both the entrance and exit beams. A reading is taken of the flux reaching a detector in the sphere wall, first without the specular beam included, and then another reading is taken with a diffuse reflector of known reflectance placed as a plug in the exit port of the sphere. This latter reading gives both the scattered and the specularly transmitted components. Subtracting the first reading, of the scattered component only, yields an estimate of the specular component. It is then a simple matter to take the ratio of the scattered to the specular components as a measure of the degree of haze exhibited by the transmitting (or reflecting) sample [38,53], generally defined to be the magnitude of this ratio, or of the ratio of the scattered component to the sum of both the specular and the scattered components.

EXAMPLE PROBLEM 8.1

Problem: Design a radiometer to measure the average solar reflectance of the moon, assuming the moon is a flat Lambertian disk facing the earth.

Solution: The “solar constant” is $E = 1353 \text{ W/m}^2$. (See Example Problem 5.1.) This is approximately the solar irradiance incident on the lunar surface when

the moon is full. During a solar eclipse, the moon just covers the solar disk. This means that the moon's angular diameter measured from a point on the earth's surface is approximately the same as the 0.535-deg solar diameter. Using this value, the moon subtends a solid angle Ω of approximately $69 \mu\text{sr}$. Let R be the average solar reflectance of the moon. The reflected radiance of the moon will therefore be approximately $1353 R/\pi \text{ W} \cdot \text{m}^{-2} \cdot \text{sr}^{-1}$ according to (1.19). The flux Φ received by a detector of area A perpendicular to a line from the center of the full moon to the center of the detector will be given by

$$\Phi = \frac{\Omega R A}{\pi} = \frac{69 \times 10^{-6} \times 1353 \times R \times A}{\pi} \quad (8.23)$$

The reflectance of the moon can be determined from this equation by measuring the flux on the detector and solving the equation for R . We would like to have an estimate of the irradiance to be reaching the detector, to assist in the choice of an appropriate detector for our radiometer. The flux level to be measured can be approximated by letting R be, for example, 0.2. Dividing the above equation by A , the irradiance on the detector will be approximately $69 \times 10^{-6} \times 1,353 \times 0.2/\pi = 5.9 \text{ mW/m}^2$.

From Figure 7.21 it is seen that flat-response thermopile or pyroelectric detectors have a D^* value on the order of $2 \times 10^8 \text{ cm} \cdot \text{Hz}^{1/2} \cdot \text{W}^{-1}$. It is desired to solve the D^* defining equation (7.24) for the NEP of the detector, assuming $A = 1 \text{ cm}^2 = 0.0001 \text{ m}^2$ and taking a reasonable value for the bandwidth Δf . In seeking a value for Δf , we must think about the detector output processing electronics. If the detector/electronics system responds to a wide range of frequencies, including the 0-Hz signal to be measured, then the value of Δf will be quite large. This will make NEP correspondingly large and reduce the SNR.

Suppose for the sake of the preliminary design, however, that the post-detector electronic system operates effectively as a low-pass filter, rejecting all frequencies above 10 Hz and passing all below this value. This means that our measurement time constant, the reciprocal of the frequency, will be about 0.1 sec. In this case, the 10-Hz bandwidth yields $\text{NEP} = 3.16/(2 \times 10^8) = 1.6 \times 10^{-8} \text{ W}$. Multiplying the irradiance on the detector by its area yields the incident lunar flux: $5.9 \times 10^{-7} \text{ W}$, a little better than an order of magnitude above the NEP.

It appears from this estimate that a measurable signal can be obtained with a thermopile detector of 1-cm² area if all other radiant flux is excluded from the detector's field of view. This indicates that the needed measurement could be performed by placing such a detector at the bottom of a baffled tube and pointing it directly at the full moon on a clear night, preferably at high altitude. The signal processing electronics need to exclude all frequencies above 10 Hz.

This theoretical SNR can be improved by either reducing the frequency bandwidth or gathering more flux. If the bandwidth is dropped to 1 Hz, for a 1-sec time constant, this will increase the signal-to-noise level by the square root of 10, or 3.16. If the true lunar reflectance is below 0.2, the received signal will be lowered accordingly. In order to provide a more comfortable margin for error, it is desired to increase the flux on the detector using a telescopic system having a focal length such that the lunar image just fills the detector area. Using a 1-cm² detector means a detector radius of 0.56 cm, or 1.12-cm diameter. Dividing this diameter by the angular diameter of the moon in radians ($0.5 \text{ deg} \times \pi / 180 \text{ rad/deg} = 8.7 \text{ mrad}$) gives the desired focal length, 128 cm. The flux on the detector can easily be increased 100-fold by using a telescope with a 100 cm² aperture area A_t , or a diameter of 5.64 cm, giving an *f-number* of $128/5.64$ nearly equal to 23, a reasonable value for astronomical telescopes.

Before finishing this preliminary design, it is noted that typical thermopile sensitivities are on the order of $10^{-5} \text{ V} \cdot \text{W}^{-1} \cdot \text{m}^2$. This means that the signals produced by a bare 1 cm² = 0.0001 m² detector receiving $5.9 \times 10^{-7} \text{ W}$ (corresponding to $5.9 \times 10^{-3} \text{ W/m}^2$) will be on the order of $5.9 \times 10^{-3} \times 10^{-5} = 5.9 \times 10^{-8} \text{ V} = 59 \times 10^{-9} \text{ V}$. This is quite a small signal and a moderately difficult one to measure. It would be better to use the telescopic system and increase the signal a hundredfold, to 5.9×10^{-6} . Microvolt signals are generally easier to measure than nanovolt ones.

This basically completes the problem. The use of a 1.12-cm diameter thermopile detector at the focus of a 128-cm focal length, 5.64-cm diameter telescope (of area $A_t = 100 \text{ cm}^2 = 0.01 \text{ m}^2$) has been specified (filling the area of the detector with an image of the moon), with a voltmeter having a significant response to voltage fluctuations from 0 to no more than 1 Hz. In performing the measurements, care must be taken to make sure that the telescope tracks the full moon steadily for a period of at least several seconds on a clear night without significant contamination from atmosphere-scattered terrestrial light sources, and the voltmeter must be allowed to stabilize on the signal for a period of several seconds before taking readings.

The voltage readings V are then to be reduced by subtracting the dark voltage V_o and then multiplying by the sensitivity of the detector to determine the flux incident on it. This flux is then reduced by the ratio of the telescope aperture area A_t to the detector area A_d and increased by the transmittance T of the telescope's optical system to determine the flux Φ incident upon the known area A_t of the telescope entrance aperture. These values are substituted into equation (8.23) and the result is solved for the average reflectance R of the lunar surface. Let S be the responsivity of the detector, V_o its dark voltage, and V the output signal voltage. The approximate reflectance of the moon is then given by

$$\begin{aligned}
 R &= \frac{\pi S(V - V_o)}{\Omega EA, T} \\
 &= \frac{\pi S(V - V_o)}{690 \times 10^{-6} \times 1353 \times 0.01 \times T}
 \end{aligned} \tag{8.24}$$

Once the telescope transmittance T , the detector responsivity S , and dark and signal voltages are known, this expression can be evaluated.

REFERENCES

- [1] Born, M., and E. Wolf. *Principles of Optics*, 3rd ed., Pergamon Press: New York, NY, 1965, p. 151.
- [2] Welford, W. T. *Aberrations of Optical Systems*, Institute of Physics Publishing: Bristol, UK and Philadelphia, PA, 1986.
- [3] Kingslake, R. *Lens Design Fundamentals*, Academic Press: New York, NY, 1978.
- [4] Kingslake, R. *Optical System Design*, Academic Press: New York, NY, 1983, p. 84.
- [5] Kingslake, R. *Optical System Design*, Academic Press: New York, NY, 1983.
- [6] Kingslake, R. "Basic Geometrical Optics," in *Applied Optics and Optical Engineering Vol. I, Light: Its Generation and Modification*, Academic Press: New York, NY, 1965.
- [7] Williams, C. S., and O. A. Becklund. *Optics: A Short Course for Engineers and Scientists*, Wiley-Interscience: New York, NY, 1972.
- [8] Kingslake, R. *Lens Design Fundamentals*, Academic Press: New York, NY, 1978. See also modern textbooks on geometrical or classical optics and on optical engineering.
- [9] Jenkins, F. A., and H. E. White. *Fundamentals of Optics*, 3rd ed., McGraw-Hill Book Co.: New York, NY, 1957, Eq. 4a, p. 46.
- [10] Strong, J. *Concepts of Classical Optics*, W. H. Freeman and Company: San Francisco, CA, 1958.
- [11] Smith, W. J. *Modern Optical Engineering*, McGraw Hill Book Co.: New York, NY, 1966.
- [12] Kingslake, R. "Lens Design," in *Applied Optics and Optical Engineering*, Vol. III Optical Components, ed. by R. Kingslake, Academic Press: New York, NY, 1965.
- [13] Born, M., and E. Wolf. *Principles of Optics*, 3rd ed., Pergamon Press: New York, NY, 1965.
- [14] Kingslake, R. *Optical System Design*, Academic Press: New York, NY, 1983, p. 21.
- [15] Jenkins, F. A., and H. A. White. *Fundamentals of Optics*, 3rd ed., McGraw-Hill Book Co.: New York, NY, 1957, p. 302.
- [16] Pearson, J. E., R. H. Freeman, and H. C. Reynolds. "Adaptive Optical Techniques for Wave-Front Correction," Chapter 8 in *Applied Optics and Optical Engineering*, Vol. VII, ed. by R. R. Shannon and J. C. Wyant, Academic Press: New York, NY, 1979.
- [17] Kingslake, R., ed. *Applied Optics and Optical Engineering*, Vol. III: Optical Components, Volumes IV and V: Optical Instruments, Academic Press: New York, NY, 1967.
- [18] Jacobs, D. H. *Fundamentals of Optical Engineering*, McGraw-Hill Book Co.: New York, NY, 1943.
- [19] Welford, W. T., and R. Winston. *The Optics of Nonimaging Concentrators*, Academic Press: New York, NY, 1978.
- [20] Nicodemus, F. E. "Self-Study Manual on Optical Radiation Measurements: Part I—Concepts, Chapters 1 to 3," NBS Technical Note 910-1, National Institute for Standards and Technology, Gaithersburg, MD, March 1976.
- [21] Steel, W. H. "Luminosity, Throughput, or Etendue?," *Applied Optics*, Vol. 13, 1974, pp. 704–705.

- [22] Goebel, D. G. "Generalized Integrating-Sphere Theory," *Applied Optics*, Vol. 6, 1967, pp. 125-128.
- [23] Jacques, J. A., and H. F. Kuppenheim. *J. Opt. Soc. Am.*, Vol. 45, 1955, p. 460.
- [24] Hisdal, B. J. *J. Opt. Soc. Am.*, Vol. 55, 1965, pp. 1122 and 1255.
- [25] Halon, available from Allied Chemical Co., P.O. Box 697, Pottsville, PA 17901. Eastman White Reflectance Paint, available from Kodak Laboratory and Specialty Chemicals, Eastman Kodak Co., 343 State St., Rochester, NY 14650.
- [26] Weidner, V., and J. J. Hsia. "Reflection Properties of Pressed Polytetrafluoroethylene Powder," *J. Opt. Soc. Am.*, Vol. 71, 1981, pp. 856-861.
- [27] Goebel, D. G. *J. Opt. Soc. Am.*, Vol. 56, 1966, p. 783.
- [28] ANSI/ASHRAE Standard 74-1988. "Method of Measuring Solar-Optical Properties of Materials," Amer. Soc. Heat. Refr. & Air-cond. Eng., Atlanta, GA, 1988.
- [29] Goebel, D. G., B. P. Caldwell, and H. K. Hammond. "Use of an Auxiliary Sphere with a Spectroreflectometer to Obtain Absolute Reflectance," *J. Opt. Soc. Am.*, Vol. 56, 1966, pp. 783-788.
- [30] ASTM. "Standard Method for Absolute Calibration of Reflectance Standards," Standard E 306-71, American Society for Testing and Materials, and ANSI Standard Z172.4-1973.
- [31] Labsphere. "Celebrating 100 Years of Integrating Sphere Technology," undated publication from Labsphere, P. O. Box 70, North Sutton, NH 03260.
- [32] Nicodemus, F. E., et al. "Geometrical Considerations and Nomenclature for Reflectance," NBS Monograph 160, National Bureau of Standards (now NIST), Gaithersburg, MD, 1977.
- [33] Weidner, V. R., and J. J. Hsia. "NBS Measurement Services: Spectral Reflectance," NBS Special Publication 250-8, National Institute of Standards and Technology, Gaithersburg, 1987.
- [34] Kessel, J., "Transmittance Measurements in the Integrating Sphere," *Applied Optics*, Vol. 25, 1986, p. 2752.
- [35] Zerlaut, G. A., and T. E. Anderson. "Multiple-Integrating Sphere Spectrophotometer for Measuring Absolute Spectral Reflectance and Transmittance," *Applied Optics*, Vol. 20, 1981, p. 3797.
- [36] Hisdal, B. J. "Reflectances of Perfect Diffuse and Specular Samples in the Integrating Sphere," *J. Opt. Soc. Am.*, Vol. 55, 1965, p. 1122.
- [37] Hisdal, B. J. "Reflectances of Nonperfect Surfaces in the Integrating Sphere," *J. Opt. Soc. Am.*, Vol. 55, 1965, p. 1255.
- [38] ASTM. "Standard Test Method for Haze and Luminous Transmittance of Transparent Plastics," D 1003-92, Amer. Soc. Testing & Matl., Philadelphia, PA.
- [39] Oriel Corporation. "Optics & Filters," Vol. III, Oriel Corporation, 250 Long Beach Blvd., P. O. Box 872, Stratford, CT 06497, 1990.
- [40] Lubkin, G. B., ed. "Tenth Annual Physics Today Buyers' Guide," Part 2 of the August 1993 issue of *Physics Today*, 335 East 45 St., New York, 10017.
- [41] Kingslake, R. *Applied Optics and Optical Engineering, Vol. V: Optical Instruments-Part II*, Academic Press: New York, NY, 1969.
- [42] Girard, A. and P. Jacquinot, "Principles of Instrumental Methods in Spectroscopy," and Baird, K. M., "Interferometry: Some Modern Techniques," in A.C.S. van Heel, ed., *Advanced Optical Techniques*, John Wiley & Sons, New York, 1967.
- [43] CIE. "Electric Light Sources State of the Art-1991," 1st ed. 1992, Technical Report, Publ. No. CIE 96, CIE Central Bureau, Kegelgasse 27, A-1030, Vienna, Austria, 1992.
- [44] Oriel Corporation. "Light Sources, Monochromators & Spectrographs, Detectors & Detection Systems, Fiber Optics," Volume II, Oriel Corporation, 250 Long Beach Blvd., P. O. Box 872, Stratford, CT 06497, 1994.
- [45] Kaufman, J. E., ed. IES Lighting Handbook - 1984 Reference Volume, IESNA, 345 E. 45 St., New York, 1984. See also: *IES Lighting Handbook: Reference and Application*. 8th ed. 1993.
- [46] McCluney, R. "Radiometry of Water Turbidity Measurements," *J. Water Poll. Contr. Fed.*, Vol. 47, No. 2, 1975, pp. 252-266.

- [47] McCluney, R. "Two-Channel Laboratory Scattering Meter," *Rev. Sci. Instr.*, Vol. 46, No. 9, 1975, pp. 1231–1236.
- [48] Kerker, M. *The Scattering of Light and Other Electromagnetic Radiation*, Academic Press: New York, NY, 1969.
- [49] Harris, F. S. *Tellus*, Vol. 21, 1969, p. 223.
- [50] Latimer, P. *Coll. Interface Sci.*, Vol. 39, 1972, p. 497.
- [51] Westwater, E. R., and A. Cohen. *Applied Optics*, Vol. 12, 1973, p. 1340.
- [52] NOAA. "Conclusions and Recommendations," NOAA/NOIC Turbidity Workshop, Arlington, VA, May 6–7 1974.
- [53] Billmeyer, F. W., and Y. Chen. "On the Measurement of Haze," *Color Research and Application*, Vol. 10, 1985, pp. 219–224.

Chapter 9

Radiometers and Photometers

9.1 INTRODUCTION

Radiometer is the term given to an instrument designed to measure radiant flux. Some radiometers measure the radiant flux contained in a beam having a known solid angle and cross-sectional area. Others measure the flux over a large range of solid angles. Some radiometers measure over a wide wavelength range simultaneously, while others perform measurements only over a narrow spectral interval. When the shape of the spectral response of a broadband radiometer is made to match the human spectral photopic efficiency function, the *V-lambda* curve, it is called a *photometer*. Some narrow spectral interval radiometers are made to scan the position of their narrow spectral interval across the spectrum. These are generally called *spectroradiometers*.

Radiometers are divided into radiance and irradiance subclasses in this chapter. Instruments with intentionally broad spectral coverage, called *broadband radiometers*, are discussed in Section 9.3. Narrow spectral band irradiance meters are dealt with in Section 9.4, along with a short discussion of radiance meters. Photometers are similarly divided into luminance and illuminance versions. These are discussed in Section 9.5. Spectroradiometers are discussed in Section 9.6 and the chapter is concluded in Section 9.7 with a discussion of instrument calibration. Calibration is a very important part of the field of radiometry and photometry. It is a truism that radiometric and photometric measurements are only as good as their calibrations.

The spectral responses of various detector types and their uses with spectral filters and other flux conditioning systems are discussed in Chapter 7. The use of cosine-response correcting optical systems in irradiance and illuminance meters is also discussed there. Chapter 8 contains descriptions of optical system components found in many radiometers. With the addition of calibration equipment and calibration procedures, as well as post-detector signal conditioning electronics, these optical systems can be made into fully functioning radiometers and photometers, the primary focus of this chapter.

Example Problem 1.2 in Chapter 1 illustrates the principles involved in the design of a radiance sensor. Section 5.9 describes the use of a lens to improve the flux-gathering performance of radiance/luminance meters. Chapter 8 expands on this idea with more information about lenses and mirrors, and their use in imaging and nonimaging optical systems. The current chapter provides an overview of the design, operation, and calibration of radiometers and photometers.

In the practical use of radiance (and irradiance) meters, it is especially important to be cognizant of the spectral limitations of the meter and to include these limits when reporting measurement results [1–4]. Flux entering the radiance (or irradiance) meter having wavelengths outside the range of sensitivity of the meter will not be measured. In such cases, the measurement results will only be a sampling of a portion of the incident flux and should be so reported. Well-built photometers do not share this characteristic. If the spectral response of a photometer strictly matches the shape of the V -lambda curve, then flux outside the visible wavelength range should not be measured, will not be recorded, and cannot be reported. Furthermore, any spectral distribution of radiation incident on the photometer will be measured correctly without relative spectral response errors.

On the other hand, if a photometer's spectral response does not quite match the shape of the photopic (V -lambda) curve, the errors can be small or large depending upon the spectral distribution of flux from the source over the spectral region of the departure from $V(\lambda)$ response. Consider, for example, the case of a measurement of the illuminance from a He Ne laser beam at wavelength 632.8 nm. This wavelength is at the red edge of the visible spectrum and a relatively small error in the V -lambda correction of a photometer at this wavelength can yield a large error in the measurement of illuminance from this source.

The kinds of radiometers, photometers, and spectroradiometers needed for general engineering work as well as for scientific research in many fields are extremely varied. In some cases the market for such devices is large enough to support their commercial development and continuous sales. An important example is the need in the field of illuminating engineering for a variety of photometers to determine if light levels are adequate at different places in buildings where people live and work. In many other cases, such as in research and other narrow applications where custom design and fabrication are required, the market is small. This would be the case, for example, if one needed a radiance meter having a specific, spectral response and solid-angle field of view, capable of scanning a hemispherical range and making rapid measurements at a specified set of directions over that range. (An example of such an instrument is one designed to scan the directional distribution of sky luminance.)

Several investigators have reported on custom designs in the open literature. The journal *Reviews of Scientific Instruments* can be consulted for information about previously designed custom radiometers or photometers. The technical journals of the Optical Society of America and *The Society of Photo-Optical Instrumentation Engineers* also contain many articles on the design and use of various types of radiometers and photometers.

mentation Engineers (SPIE) both frequently publish papers on the design and use of radiometers and photometers. Similar sources are available in other countries.

9.2 GENERAL DESIGN FACTORS

Three factors are crucial in the design of radiometers. The first consideration results from the great width of the electromagnetic spectrum (Figure 1.1). The designer must decide how much of this spectrum needs to be (or can be) covered by the radiometer. Few materials have the desired optical properties over large portions of this spectrum. The more restricted the spectral range, the easier the design task will be. The spectral distributions of common sources are plotted in Chapter 3. Most likely the choice of the source will be dictated by the application. If radiation from a given source is to be measured, then one must determine whether all electromagnetic radiation from this source is of interest or if hopefully a more restricted range of wavelengths is acceptable. For example, if a specified source such as a tungsten filament lamp, is to be used, to illuminate some sample whose transmitted radiance is to be measured, then the spectral range of significant flux output by the source, the spectral range of significant response from the sample, and the spectral range desired by the user of the radiometer will all play a role in the selection of the desired spectral range to be covered.

Almost every radiometer has windows, lenses, and/or diffusers, which inevitably impose their limitations on radiation passing through them. The spectral transmittances of common window and lens materials are described in Section 8.13. Also, many radiometric systems contain mirrors to guide or focus the flux being measured. The spectral reflectances of common mirror materials are plotted in Figure 6.1. Some radiometers and all photometers employ spectral filters. The spectral transmittances of some spectral filters are described and plotted in Chapter 8. Chapter 7 describes and provides plots of the spectral sensitivities of a number of common detectors of optical radiation. The most sensitive detectors generally have a limited spectral range of good sensitivity. This is an example of the old adage, "you can't have your cake and eat it too." High sensitivity generally comes at the expense of poor spectral response and vice versa. Other detector characteristics are also important in the selection process. The designer of a radiometer, whether it be a broad or narrowband irradiance or radiance meter, has to choose a combination of detector sensitivity, detector spectral response, the source output spectrum, and the spectral range of interest to the user. Sometimes an optimum combination of these constraints is just not available, as when very broad spectral ranges must be covered with flat spectral response and very high detector sensitivity.

Next in importance in radiometer and photometer design is the choice of the solid angle range of the field of view of the meter. Is it to look at a narrow range of angles, a broader range, or the whole hemisphere?

The third design factor of importance is the selection of a calibration method. The choice of the calibration method can in some cases overcome serious design problems.

Occasionally one can find special characteristics of a planned source that simplify the design problem. An example would be the case in which the shape of the source spectral distribution remains fairly constant while the overall magnitude varies. In this case, flat spectral response is not needed nor is full coverage of the spectral range of interest. If the *shape* of the source spectrum really does remain constant, all that is needed is measurement over a relatively narrow portion of the spectrum.

In other situations, one can limit the angular field of view and avoid dealing with cosine response correction problems. Other simplifications are possible to aid the design process. These issues are also discussed in Section 7.2.

The number of commercially available instruments that can be called radiometers or photometers is very large. Adding custom instruments to the list increases the number still further. Sections 9.3 through 9.6 are devoted, *not* to providing a comprehensive listing of the available designs and commercially available products, but to introducing the reader to examples in each of several categories. It is likely that some specialized applications will not be well represented in the summaries to follow. Some radiometer/photometer manufacturers offer a "mix-and-match" set of attachment components for their basic radiometers, thereby allowing the user to customize the instrument to some extent. One could, for example, start with a basic radiometric/photometric signal processing unit and add to it a detection (or reception) device with the specific optics and detection combination desired. Adding a flat-response correction filter to a photovoltaic detector and a cosine correction receptor on top would result in an irradiance meter whose range of relatively constant spectral response would be known. One could calibrate the resulting instrument oneself or ask the vendor to calibrate it prior to shipping. A list of photometer vendors is provided in Section 9.5. Many vendors offer comprehensive lines of radiometers as well as calibration sources and systems.

9.3 BROADBAND IRRADIANCE AND RADIANCE METERS

There are few commercially available broadband irradiance meters, and almost no broadband radiance meters. There are commercially available infrared radiometers designed to view a small solid angle and infer the temperature of the source from the measured radiance. Since the ultimate goal of these instruments is the measurement of temperature, they aren't considered radiance meters. One could purchase a basic radiometric measurement unit, add a detector/filter combination that provides the desired spectral response, and then add a beam-limiting tube or telescope to make the device a radiance sensor. Special care would have to be exercised

in this case to make sure that the device was properly calibrated and used, but the result could be considered a commercially available radiance meter.

One reason for the lack of commercially available instruments for some categories of radiometers is that the applications are so varied and the markets are so small that it doesn't make much business sense to invest in the development of such meters. This is especially true of radiance meters, which can be expensive to develop and have relatively limited applications. In many cases, therefore, investigators design and develop their own specialized radiometers for custom fabrication and limited use. Of course, one could take a calibrated irradiance meter and turn it into a radiance meter with the addition of a properly designed and baffled Gershun tube. The sensitivity of the resulting instrument, however, might be sacrificed, due to losses in the cosine correction subassembly that is not needed for radiance measurements using a small solid angular field of view.

Broadband radiometers are distinguished by the spectral range covered and their angular fields of view. In general, if high sensitivity is desired, a limited spectral range over which the sensor has flat response generally has to be tolerated. One way of achieving flat spectral response in the visible part of the spectrum is to use a specially fabricated spectral—sensitivity—compensating filter together with, for example, a photovoltaic diode detector. A reasonably flat spectral response can be obtained with this combination over the range from about 400 to about 1,100 nm. As long as the radiation to be measured falls within this range, the radiometer can be calibrated to read directly in watts of incident radiant flux, or watts per unit area of irradiance, regardless of spectral distribution within the wavelength limits. An example of the spectral response curve for such an instrument is illustrated in Figure 9.1. The curve in this figure is plotted on a logarithmic scale, which makes departures from spectral flatness or constancy appear to be less than they are. It should be noted that the spectral response for the particular detector/filter combination shown in Figure 9.1 varies considerably over the spectral range from 500 to 1,000 nm. It is therefore not a very good example of a "flat response" detector/filter combination.

When the source of radiation is quite strong, so that high flux levels are involved, detectors like the thermopile and the pyroelectric detectors described in Chapter 7 can be used to extend the spectral range of flat response even though their sensitivities are generally low.

If one plans to use a radiometer with varying, nonuniform spectral response, special care must be taken in making measurements and reporting results. The relative spectral distributions of the source and of the radiometer must be known for good results to be obtained.

Most broadband radiometers are irradiance meters, designed to accept radiation either from a hemispherical solid angle (ideally with good cosine response) or from a restricted solid angle. In the latter case, the radiation to be measured is generally assumed to be limited to a relatively narrow angular range and the field

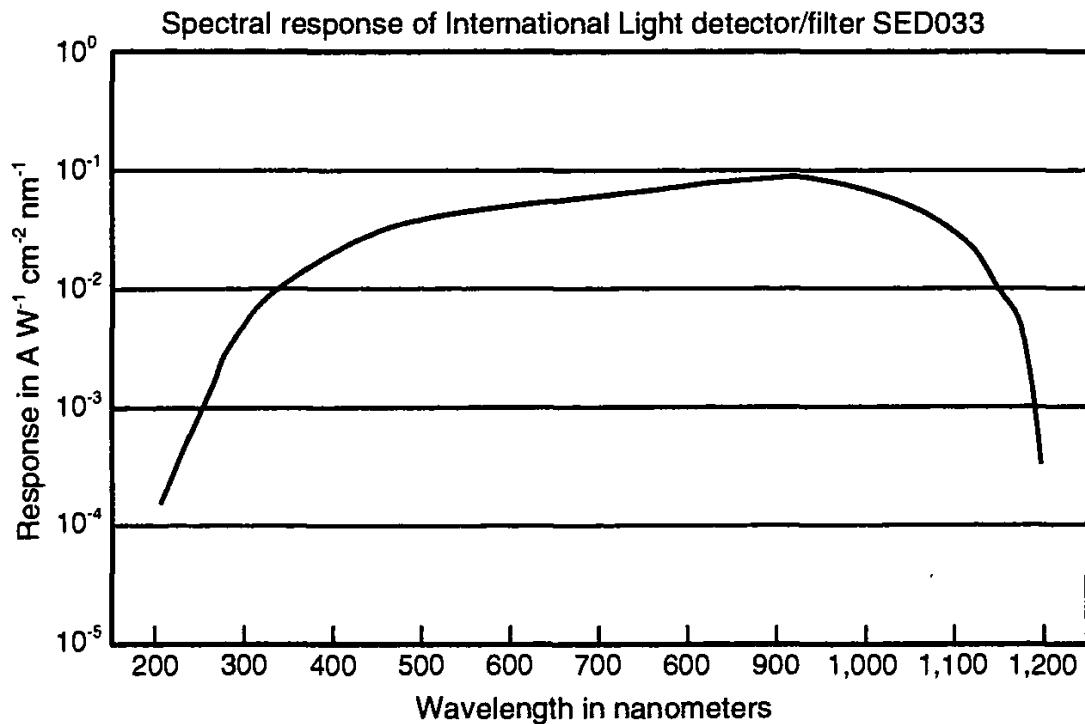


Figure 9.1 Spectral response of a detector/filter combination devised to provide somewhat uniform response over the spectral range from 400 to 1,100 nm.

of view of such irradiance meters is therefore made to be only a little bit larger than the incident beam. These radiometers are calibrated to provide measurements of the flux per unit area entering their apertures and/or incident upon their sensitive surfaces.

One area where broadband radiometers are widely available is the case of outdoor measurements of solar radiation. Most of the radiometers used for this measurement have flat spectral response from about 340 to 4,000 nm. Since the solar radiant flux is normally quite strong, a relatively insensitive detector can be used. The U.S. Weather Service once measured the global (total hemispherical) irradiance from sun and sky at a number of weather stations, thereby providing a significant market for radiometers (called pyranometers for this application) meeting these requirements and they have been on the market for some time as a result. With the collapse of the large measurement network operated by the U.S. Weather Service, the market declined somewhat, but there is a growing number of networks outside the United States. In addition, some private networks inside the U. S. measure solar radiation. The National Renewable Energy Laboratory in Golden, Colorado operates a small network and has plans for an expanded solar measurements program in the U.S. These activities continue to support the global market for solar radiometers.

Although the U.S. solar radiation measurement network is not as vast as it once was, there continue to be governmental and nongovernmental stations needing

rugged outdoor pyranometers. Other countries have substantial solar radiation measurement networks, and in 1991 a multiyear International Daylight (and solar radiation) Measurement Programme (IDMP) was begun under the sponsorship of the CIE with the participation of the World Meteorological Organization (WMO) of the United Nations. The International Energy Agency (IEA) also has an interest in the program. The need for measurement instruments for the participating stations provides a modest contribution to the demand for commercially available radiometers and photometers in this field.

The Eppley Laboratory (12 Sheffield Avenue, Newport, Rhode Island 02840), is a major supplier to the U.S. market. Their model PSP "precision pyranometer" irradiance meters incorporate a thermopile having typical sensitivity to solar radiation of $9 \mu\text{V}$ per (watt · m⁻²) with a 1-cm² sensitive area coated with Parsons' Black optical lacquer and are especially popular in the U.S. The sensitive surface of these radiometers is placed beneath two concentric hemispheres of optical glass and the housing is provided with a white shade ring to reduce temperature dependence. The sensing element is a circular multijunction Eppley thermopile of the plated, copper-constantin, wire-wound type that is temperature compensated to render the sensitivity essentially independent of the ambient temperatures normally encountered in outdoor use. More information on the design of thermopile detectors is provided in Section 7.3.1.

Eppley also offers radiometers with other spectral sensitivities, including one called a pyrgeometer that measures the flux at the far infrared wavelengths emitted by the atmosphere; ultraviolet radiometers, normal incidence irradiance meters (called pyrheliometers) that track the sun; and a precision tracking cavity radiometer for very precise measurements of direct beam solar radiation.

Similar radiometers are made in Europe by Kipp and Zonen. They are marketed in the United States by Kipp and Zonen (Div. of Enraf-Nonius Co., 390 Central Ave., Bohemia, NY 11716). Included are models CM 21 and CM 11, satisfying the WMO standard for secondary radiometers, and CM 6B and CM 5 at lower prices. The company also markets actinometers, albedometers, shadow-band radiometers for measuring diffuse sky radiation while excluding direct beam solar radiation, as well as ultraviolet sensors having various spectral sensitivities.

Additional radiometers for this market have become available from EKO Instruments Trading Company (21-8, Hatagaya, 1 Chome Shibuya-ku, Tokyo, 151 Japan). This company offers a comprehensive line of radiometers for solar energy applications and daylight photometers and radiometers, including a directionally scanning luminance meter (sky scanner) and a sunshine recorder. Miyake has reviewed the principles of operation of many of these instruments [5].

PRC Krochmann GMBH (Genestrasse 6, D-1000, Berlin 62, Germany) also offers a comprehensive line of radiometers, photometers, luminance scanners, colorimeters, and calibration services.

Another class of "broadband radiometers" is really designed for measurement of flux levels of monochromatic, collimated beams of laser radiation. The job of

these *laser power meters* is to accept radiation in a highly collimated beam and to determine the total flux in watts contained in that beam, but to do so accurately for any selected wavelength over a wide spectral range.

A general purpose broadband radiometer can be envisioned for this application that would use one of the broad spectral range detectors described in Chapter 7 to accept radiation from any one of a number of powerful lasers having emission wavelengths from the ultraviolet to the infrared portions of the spectrum, and calibrated to provide accurate power (flux) measurements for any wavelength within the design range of the instrument. For example, the Oriel Corporation (250 Long Beach Blvd., Stratford, CT 06497), offers its model 70160 Radiant Power Meter with model 70161 Thermopile Probe that is calibrated to measure monochromatic or broadband power, dosage, or energy, over a flux range from $30 \mu\text{W}$ to 2W and a wavelength range from 190 nm to 20 μm . Oriel offers a range of detector/input optics combinations for use with several electronic readout instruments. Other instrument companies also offer similar "mix-and-match" services.

9.4 RESTRICTED SPECTRAL BAND IRRADIANCE METERS FOR THE ULTRAVIOLET THROUGH THE INFRARED

For a variety of applications, it is sufficient to measure radiation only over a restricted spectral region. For such applications, detectors with greater sensitivity but non-uniform spectral response can often be used. In such applications, it is sometimes beneficial to precede the detector with a narrow bandpass filter to exclude radiation outside the spectral range of interest. Narrow passband spectral filters are described in Section 8.12.1.

With a combination of spectrally selective detector response and a spectral filter, the combined detector/filter combination often can be given the precise spectral response required by the application. Precisely known spectral response is especially important in the ultraviolet region of the spectrum, where many sources (such as the sun) have an output that varies rapidly with wavelength. It is important for the spectral pass band of the radiometer to be very narrow and very constant. In many cases, such as the solar ultraviolet one just mentioned, the out-of-band rejection capabilities of the filter will be very important. Even a 1% transmittance in the wings of such a filter can transmit out-of-band flux comparable to the in-band transmitted flux. If the detector sensitivity is much greater out-of-band than in-band, this effect is even more pronounced.

If the shape of the source spectrum remains constant, it is possible (as described in Section 7.2) to use unfiltered detectors whose spectral responses vary over a portion of the source spectrum, provided these detectors are calibrated properly for the spectral shape of the source they are intended to measure.

Occasionally one encounters design cases for which accurate measurements of radiation sources of relatively low flux having high spectral variation over a wide spectral range are needed. The low flux rules out flat response broadband sensors. (Unless one can concentrate the flux from the source onto the detector to achieve higher and more measurable flux levels. See Example Problem 8.1.) The strong spectral variations lead to a need to measure the spectrum one wavelength at a time with a highly sensitive detector and to integrate the resulting spectrum to obtain the total radiant flux, irradiance, or radiance.

9.5 ILLUMINANCE AND LUMINANCE METERS

Since illuminance meters measure the illumination needed by humans to see, they are very common. For a number of years, technical societies concerned with lighting have offered recommendations about illuminance levels needed for various visual tasks [6–9]. These recommendations are discussed briefly in Section 2.7. In order to determine compliance with these recommendations, it is important to be able to measure directly the illumination levels produced. The two kinds of illumination measurement of interest are illuminance and luminance.

Illuminance meters are generally hand-held detectors equipped with cosine correcting optical components of various kinds, photopic spectral correction filters that modify the filter/detector spectral response so that it matches the human photopic visibility function tabulated in Chapter 2, and signal conditioning and displaying electronics to output the measurement results.

Luminance meters are almost as important as illuminance meters, since they measure a quantity that is close to what the eye sees. It is the reflected and emitted luminance from objects that we see and which luminance meters measure. Some objects are self-luminous, such as light sources, video screens, cockpit displays, and the sun and sky, while others reflect light from other sources. Luminance meters are used to measure the light approaching the eye from such objects.

Many luminance meters are handheld devices, equipped with field-of-view delimiting optical components and sighting optics to make it easy to determine the specific location on an emitting or reflecting object where the luminance is being measured. Typically one looks through an eyepiece and depresses a switch to take (or hold) a reading. The luminance meter accepts luminous flux leaving points over a defined area (the field of view) of a surface, within a solid angle defined by the aperture area of the meter and a point in the surface, and effectively averages this flux over the defined area of the source and the solid angle. The result is a measurement of luminous flux per unit area and per unit solid angle leaving the source, the luminance of the source. The methodology of this measurement is described in greater detail in Section 5.9.

Table 9.1 lists U.S. manufacturers of illuminance and luminance meters. Most of these make other radiometers and photometers as well. Some offer a very wide range of radiometers, photometers, spectroradiometers, and calibration services for a variety of applications.

9.6 SPECTRORADIOMETERS

Monochromators are discussed in Section 8.12. A *scanning spectroradiometer* is an optical system containing viewing or flux reception optics, some form of monochromator that scans a range of wavelengths over time, and a detection and signal processing system.

A *spectrograph* presents a dispersed spectrum containing a whole wavelength range to the output plane all at once. A spectrograph can be thought of as a *polychromator*, an instrument that spreads an incident beam of radiation into its various wavelengths so that they can be detected simultaneously. Historically, the radiation so distributed was recorded on photographic plates and it was the presence or absence of the radiation that was more important than its strength at each wavelength. Today spectrographs are available with many-element detector arrays that can record the magnitudes of the different wavelength components of the radiation essentially simultaneously.

There are two basic kinds of spectroradiometers. The first kind measures radiation incident on it from outside of the instrument and it is calibrated to quantify the spectral irradiance or radiance of the incident beam. Such a device is properly called a *spectroradiometer*. The second kind measures radiation from a source contained within the instrument after it passes through or reflects from a material sample placed in a sample chamber or liquid cell inside the instrument. This is normally called a *spectrophotometer*. (See Section 8.12.2.3.)

Spectral reflectance can be measured with most spectrophotometers through the use of special attachments. Normally these devices measure the specular transmittance of a sample (the biconical transmittance with the angle of transmittance equal to the angle of incidence). Integrating sphere attachments are available with some of them to permit measurements of conical-hemispherical transmittance and reflectance of samples at normal or near-normal incidence. Some manufacturers offer additional attachments for measurement of reflectance and transmittance over a range of incident angles. U.S. vendors of spectrophotometers include Perkin Elmer Corp., Analytical Instruments, Main Ave., MS-12, Norwalk, CT 06856; Beckman Instruments, Inc., 2500 Harbor Blvd., Fullerton, CA 92634; Varian Analytical Instruments, 505 Julie Rivers Rd., Suite 150, Sugar Land, TX 77478; Hitachi Instruments, Inc., 13 Riverside Road, Suite 102, Weston, MA 02193; and Optronics Laboratories, 4470 35th St., Orlando, FL 32811. Additional vendors are listed in the *Physics Today* and other science buyers' guides [10–14].

Table 9.1
Photometer Manufacturers

<i>Vendor</i>	<i>Illuminance Meters</i>	<i>Luminance Meters</i>
Tektronix, Inc. P. O. Box 4600, Beaverton, OR 97076-4600	Model J17 photometer with J1811 probe (head)	Model J17 photometer with J1803 head (8 deg) and J1823 head (1 deg)
Li-Cor, Inc., 4421 Superior St., P. O. Box 4425, Lincoln, NB 68504	LI-210SA photometric sensor with LI-1000 readout instrument/data logger	N/A
Minolta Corporation, 101 Williams Drive, Ramsey, NJ 07446-1293	T-1, T-1H, and T-1M handheld autoranging illuminance meters	LS-100 and LS-110 handheld autoranging luminance meters
EKO Instruments Trading Co., Ltd., 1-21-8 Hatagaya, Shibuya-ku, Tokyo Japan	N/A	MS-300 Sky Luminance scanner
Optronic Laboratories, Inc. 4470 35th Street, Orlando, FL 32811	OL730, OL730A, and OL730HV radiometer/photometers with cosine receptors	OL730 series radiometer/photometer with luminance/telescopic receptors
The Cooke Corporation 900 Hertel Ave., P. O. Box 209 Buffalo, NY 14207-2866	A variety of luxmeters, including the cal-Light, a calibrated, hand-held illuminance meter.	Universal Photometer S2, with both illuminance and luminance capabilities.
PRC Krochmann, GmbH Geneststrasse 6 D-10829 Berlin, Germany	Full line of photometers and specially designed daylight measuring illuminance meters	Sky luminance scanner and other luminance meters
Photo Research, Div. of Kollmorgen Corp., 9330 DeSoto Ave., Chatsworth, CA 91311-4926	Litemate/Spotmate System 500 handheld photometer with illuminance attachment	Litemate/Spotmate System 500 handheld photometer with illuminance attachment
EG&G Gamma Scientific, 8581 Aero Drive, San Diego, CA 92123	Full line of digital radiometers and photometers with a variety of detection systems	Photometer with model 700-2A photometric telescope attachment
International Light, 17 Graf Rd., Newburyport, MA 01950-4092	IL1400A handheld and IL1700 research radiometer/photometer with cosine corrected photopic detectors	IL1400A handheld and IL1700 research photometer with photopic correction and input optic barrels

Many different forms of scanning monochromators are available, each with its own strengths and weaknesses. Normally, some form of wavelength dispersing device is placed between an entrance and an exit aperture. Frequently the entrance and exit apertures have the shape of a narrow slit. The entrance slit selects the portion of the incident beam to be measured and the exit slit selects the range of wavelengths permitted to emerge from the monochromator. Some monochromators

utilize a refracting prism as the dispersing device (see Section 8.12.2.1). Other monochromators use a diffraction grating to disperse the wavelengths. (See 8.12.2.2.)

Interference filters (discussed in Section 8.12.1) are multilayer coatings of dielectric and other materials which have strongly spectrally selective transmittance. The narrow passband version of the interference filter can be used as a monochromator. If only a few wavelengths are of interest, a wheel of filters or some other mechanism can be constructed to permit rapid changing from one filter to another, producing a fairly simple form of scanning spectroradiometer. In this case, however, the "scanning" is discontinuous and "filter wheel radiometer" would be a better term for the device. Filter wheel radiometers have been sold commercially in the past.

It is possible to manufacture what is called a *circularly variable interference filter*, an interference filter in the shape of a wheel or disk mounted on a rotating axle through its center. The wavelength of the center of the spectral passband of such a device varies continuously with angle around the axle. The circularly variable interference filter makes an excellent monochromator for a scanning spectroradiometer. Its primary limitations are that the width of the spectral passband is generally greater than that achievable with a grating monochromator and the spectral range of coverage is limited. Several companies offer both circular and linearly variable interference filters. One U.S. manufacturer has based an entire line of spectroradiometers on these monochromators [15].

9.7 CALIBRATION OF RADIOMETERS AND PHOTOMETERS¹

Radiometers and photometers involve a number of components all contributing to the overall sensitivity of the instrument to incident radiant flux. Although one could in principle determine the contribution of each individual component to the overall calibration of the instrument, in practice this procedure is seldom used. Instead, the complete instrument is calibrated all at once.

Calibration is usually a two-step process. First one determines the mathematical transformation needed to convert an output electrical signal into an estimate of the input flux in the units desired for the quantity being measured. Second, one ensures the accuracy of this transformation over time as the characteristics of the components making up the radiometer or photometer change or drift of their own accord or as a result of outside stimuli.

There are two approaches to calibrating or recalibrating a radiometer/photometer. In the first case, one uses the radiometer/photometer to measure flux from a standard source whose emitted flux is known accurately in the desired units

¹Some of the information contained in this section was provided by William E. Scheider and David Goebel, and they are thanked for permission to reproduce this material.

and then applies a suitable transformation to convert the output signal from the device being calibrated to the proper magnitude and units of the standard input. For this procedure to work, it is critical that the overall response of the radiometer/photometer be quite constant over the period of time between calibrations.

In the second case, one measures the flux from an uncalibrated source first with the device to be calibrated and then with an already calibrated standard radiometer/photometer having identical field of view and spectral response. The output of the device is then calibrated so that it is identical to the measured result and units obtained with the standard radiometer/photometer. Once the calibration transformation is known, it can be applied to subsequent measurements and the device is thereby said to be calibrated. Both of these approaches are described subsequently.

For radiometers and photometers whose calibration is known to drift slowly over time, one can calibrate the device when it is first fabricated and can then recalibrate it periodically over some acceptable time period. For most accurate results, it is advisable to recalibrate frequently at first, and to increase the time interval between recalibrations only after a history of relatively minor drift has been established. For most precise radiometry and photometry, a *working standard* or *transfer standard* is used to make frequent calibration checks between (or even during) measurements to account for the effects of small residual drifts in the calibration of a radiometer or photometer. (See Section 9.7.1.)

Historically, the focus of calibration was on the preparation of standard sources, most notably standard lamps, which produce a known and constant quantity of flux giving a known irradiance at a fixed distance from the emitting element. A typical measurement configuration is illustrated schematically in Figure 9.2(a). However, there is now growing interest in the use of calibrated *detection standards*; that is, detectors whose responsivity is sufficiently constant and reproducible over time to make possible the calibration of other detectors or radiometers based on these standard detectors.

Standard lamps generally produce a fixed output flux distribution with wavelength. Because of the possibility of nonlinear response effects, radiometers and photometers should be calibrated only over their ranges of linearity, within which a standard lamp can be found. One common method of equalizing standard lamp output to the intended measurement range of the radiometer is to take advantage of the inverse square law shown in (4.7) to reduce the irradiance from an essentially point source standard lamp, from its calibrated value E_o at the specified distance R_o from the lamp to the reduced value E at the greater distance R :

$$E = \frac{E_o R_o^2}{R^2} \quad (9.1)$$

There are limits to the use of this technique, however. The distance R cannot

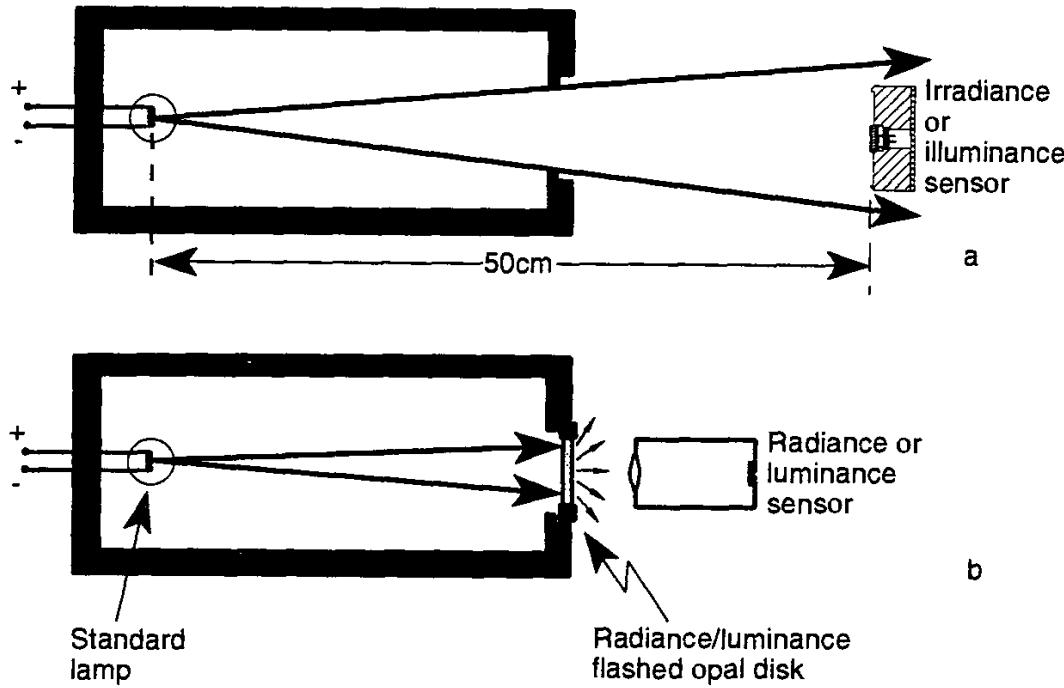


Figure 9.2 Illustration of the use of a standard lamp (a) to calibrate an irradiance or illuminance sensor and (b) to calibrate a radiance or luminance sensor.

practically be made many orders of magnitude greater than R_o . For example, one cannot hope to use a standard whose spectral irradiance is in the $1\text{-W}\cdot\text{m}^{-2}\cdot\text{nm}^{-1}$ range at 50 cm from the source to calibrate a device intended for measurements many decades lower than this value and achieve meaningful results. The signal-to-noise ratio would be too low. Special-purpose standards have been developed that can be used at radiation levels substantially different from the normal standard sources used widely.

An example is the case of radiometers for measuring the radiant flux from distant stars. The irradiance levels in these cases are significantly lower than those produced by most standard lamps, even at large distances from the lamps. Special standard sources have been devised that produce standard irradiance distributions at the very low irradiance levels measured by star radiometers.

9.7.1 Transfer Standards

Many nations of the world maintain *primary standards* for radiometry and photometry in national laboratories dedicated to this purpose. In the United States, such standards are maintained by the National Institute of Standards and Technology in Gaithersburg, Maryland. From these are derived *secondary standards*

(also called *transfer standards*) that can be maintained at private laboratories or by other organizations for the purpose of calibrating and recalibrating commercially and custom produced radiometers and photometers. *Working standards* are standards derived from secondary standards but which are designed and intended for easy and repeated use to check the calibration of a radiometric or photometric system periodically during or between measurements. For example, scattering meters and reflectance meters generally derive their calibration from the replacement of the sample being measured with a working standard whose scattering or reflectance properties are well known. Spectroradiometers designed for measurement of unknown spectral distributions emanating from external objects, surfaces, or media generally can compare the flux received from the target medium with that received from a working standard source of known flux at each wavelength of measurement.

An example of the latter is found in the LANDSAT remote sensing satellite. A multispectral scanner on this satellite has a rotating mirror that scans the earth below the satellite, laterally to the satellite's motion across the earth. As the mirror rotates, sending radiation from the earth's surface to a detection and measurement system, the satellite advances forward, so that each scan or sweep of the earth is displaced slightly from the previous sweep. Each time the mirror sweeps upward, away from the earth, in its rotation cycle, it "looks" at a working standard source of known spectral radiance above the mirror inside the satellite. As a result, the data stream from the satellite contains frequent signals from the working standard so that a nearly continuous recalibration process can be instituted when the signals are received on the ground.

Another example of what might be called a working standard is the use of an integrating sphere or an opal-like white diffusing glass disk to receive radiation from an incandescent filament lamp and redistribute it in direction. The diffuse radiation leaving the disk or sphere can be made into a standard of radiance or luminance if properly calibrated. This is illustrated in Figure 9.2(b). This arrangement can provide a finite-sized source of spectral radiance that is relatively uniform spatially over the exit aperture of the sphere or the surface of the diffusing disk and relatively uniform in directional radiance over a defined solid angular range. A radiance (or luminance) meter can be calibrated with such a device as long as the field of view of the radiance or luminance meter is smaller than the spatially uniform area of the source and the meter "looks" at the source from within the solid angle over which the source calibration is valid.

Beginning in the early 1950s, the National Bureau of Standards (now NIST) provided two-inch squares of flashed opal glass as transmitting luminance standards. Originally they were used in specifying and procuring instrument-panel lighting. Over the years they found wider application in the calibration of luminance meters, telephotometers, cathode ray tube screens, and so forth. The NBS calibrations were performed by a substitution method against a set of three master standards.

The original calibration of the master standards was not well documented however, and opal glass is far from ideal as a transfer standard for general use in luminance meter calibration. The glass substrate is somewhat spectrally selective, the opal flashing on the glass is not a perfect diffuser, and with a narrow field of view the glass exhibits a forward peak in the transmittance function that can lead to calibration errors with narrow field of view luminance or radiance meters [16]. Caution is therefore recommended in the use of opal glass squares or disks as radiance/luminance standards. If they are used together with irradiance/illuminance standard sources, as illustrated in Figure 9.2(b), it is recommended that they be calibrated for radiance/luminance independently of the radiance transmittance calibrations furnished by NBS and that care be taken to use them only at normal incidence. The calibration of these opal glass radiance/luminance standards should be performed with the same angular field of view for which they will be used.

Calibrated silicon photodetectors are now available as transfer or working standards that are based on the NIST detector response transfer intercomparison package (DRTIP) [17], which consists of a carefully designed silicon detector radiometer whose calibration is based on the NIST absolute radiant power scale. The calibration transfer instruments are known as detector response transfer packages (DRTP), but they are commonly referred to as DRIPs [18]. Spectral response values for the DRTPs are given at selected wavelengths over its useful range.

9.7.2 Broadband Irradiance Standard Sources

Calibrations using a standard lamp frequently utilize specially designed tungsten filament lamps whose emitting characteristics are known to be quite constant over a period of time if the lamp is not frequently used. Such lamps must be operated with precisely the same electrical current through the filament as when their calibrations were initially set. Specially designed power supplies are made for use with such lamps. These power supplies ensure the constancy of this filament current and also keep track of how many hours the filament has been operated since initial calibration.

Until about 1912, the only radiometric standards in existence were crude oil lamps or carefully trimmed and sized candles. In 1913, W. W. Coblenz set up the carbon filament lamp as a practical working standard for irradiance [19]. These lamps were calibrated relative to the exitance of a blackbody radiator, as given by the Stefan-Boltzmann law, (3.1). The total (meaning integrated over wavelength) irradiance produced by these lamps for calibration was on the order of 1 W/m^2 at a specified distance from the lamp and the spectral distribution was such that a relatively low color temperature was produced, around 1,900K. At this low color temperature most of the emitted flux was in the infrared portion of the spectrum.

Koblenz's carbon filament lamps received wide acceptance in scientific research throughout the world and were used extensively until the establishment of tungsten filament lamps as more stable and higher output standards a number of years later. Calibration of the new tungsten lamps was based on the computed spectral radiance $L_{bb\lambda}$ of a blackbody, using Planck's radiation law, (3.3). The lamps were used as irradiance standards, with the standard value of irradiance being quoted for a plane normal to the lamp filament and assumed to be a blackbody at a specified distance from the filament. Over the years, the U.S. National Bureau of Standards [20], now NIST, has worked to develop improved irradiance standards.

Today, one can obtain irradiance standard lamps commercially and use them for the calibration of broadband irradiance sensors. They must be operated according to manufacturer specifications and care must be taken to avoid stray light from the source reflecting from adjacent objects and into the radiometer being calibrated.

If the radiometer has a nonuniform spectral response over the range of sensitivity and/or over the range of emission from the standard source, then extra care must be taken in its use. In such cases, accurate results can be obtained from using a standard lamp whose spectral irradiance or radiance distribution is known, so that the correct quantity of radiation within the wavelength range of sensitivity of the radiometer will be known. Calibrating with such a source, however, cannot be expected to provide accurate results when the radiometer is used with sources whose spectral distributions are shaped differently. For further discussion of these effects, see Section 7.2. The discussion there addresses detectors independent of any attached optical system, but it is applicable to complete radiometric measuring systems as well.

9.7.3 Standard Sources for Spectral Irradiance and Spectral Radiance

With the exception of a blackbody radiator (Section 3.6), until around 1960 there was no acceptable and convenient standard for the spectral distribution of either radiance or irradiance. The blackbody was and is the primary standard of spectral radiance for most infrared calibrations. However, its use in the ultraviolet (UV), visible (VIS), and near infrared (NIR) parts of the spectrum is very limited. For some time laboratories used published values for the emittance ϵ of tungsten filaments, observed the color temperature or brightness temperature of the heated filament, and inferred the spectral radiance of the filament from Planck's radiation law. This procedure was based on the doubtful assumption that all samples of tungsten have the same emittance. No account was taken of the effects of impurities present, of the size and shape of the filament, or of its crystalline structure. All of these properties affect the spectral and total exitance and radiance. In addition, interreflections within the lamp's glass envelope affect the spectral radiant flux

emanating from a particular tungsten lamp. In order to obtain the correct spectral radiance or irradiance from such a lamp, it is important to calibrate each particular lamp against a blackbody. Today, accurate approximations of a blackbody radiator are used to establish a standard scale for irradiance. The design and construction of a practical blackbody radiator is discussed in Section 3.8.

Researchers at the National Bureau of Standards (NBS, now National Institute of Standards and Technology, or NIST) have worked to develop improved standards of spectral radiance and irradiance for the ultraviolet, visible, and near infrared portions of the spectrum. In 1963, the NBS developed a quartz halogen lamp standard of spectral irradiance [21]. A GE 200-watt quartz iodine lamp was found to have acceptable characteristics for use as a standard. It is a rugged lamp in a small quartz envelope of relatively good optical quality such that the intensity usually varies little over a considerable solid angle centered normal to the designated axis of the lamp. The filament is a compact coiled coil with overall dimensions approximating 3 mm × 13 mm. The small size of the lamp envelope (12 mm × 50 mm) together with the small area of the filament permits placing the lamp within 50 cm of the slit of a spectrophotometer.

Because of its high operating temperature and UV-transmitting quartz envelope, the quartz halogen lamp emits a greater percentage of radiation in the UV than conventional tungsten incandescent lamps. The so-called "halogen cycle" results in the return of evaporated tungsten from the glass envelope back to the lamp filament, thereby keeping the envelope clean and prolonging the useful life of the lamp. The design life of this lamp when operated at 6.6A is 500 hours. For calibration as a standard, the current is set precisely at 6.5A, giving a filament temperature of around 3,000K. A somewhat larger 1,000W standard lamp was set up in 1965 that produced an output approximately five times that of the 200W quartz iodine lamp.

The uncertainties assigned to the 200 and 1,000W lamp standards varied from ±8% at 250 nm to ±3% in the visible. The widespread use of these standards and the need for a higher accuracy standard led NBS to establish a newer, higher accuracy scale of spectral irradiance in 1973 [22]. The new scale had estimated uncertainties which were about 1/3 that of the older 1963 scale. The new irradiance standards were calibrated over the wavelength region from 250 to 1,600 nm. In 1975, NBS switched to another lamp that was similar in filament and envelope, but had a different base, permitting quick removal and replacement.

In 1977, NBS added still another new standard of spectral irradiance for the wavelength range 200–350 nm. It is based on a deuterium lamp that utilizes an argon miniarc as a standard of spectral radiance. By modifying the calibration technique, this new lamp can be used as an irradiance standard. The spectral irradiance at 50 cm from this lamp is 100 times that of the quartz halogen lamp at a wavelength of 200 nm, they are equal at about 280 nm, and the quartz-halogen exceeds the deuterium lamp by 30 times at 350 nm.

Since the shapes of the two spectra are so different, a critical test of a narrow spectral band radiometer's calibration would be to measure the output of the two lamps at the same wavelength separately. If the ratio of the signals from the radiometer for the two lamps is not equal to the ratio of their calibrated radiances or irradiances at the given wavelength, it would follow that a systematic source of error is present in the measuring system. The source of the error could be scattered light, second-order radiation from a grating, poor out-of-band rejection characteristics, or detector nonlinear response.

In 1980, several problems were reported for the deuterium standards, mainly due to a shift in absolute output over time. The NBS then recommended use of a tungsten quartz halogen-type FEL (ANSI designation) lamp standard of spectral irradiance to normalize the reported deuterium (D_2) spectral irradiance values to the FEL values at the time of measurement. As a result of this experience, the NBS switched to deuterium lamps from a different manufacturer that exhibited much less drift. However, normalization to an FEL lamp was still recommended [23].

According to a 1988 NBS publication [24], "Modified type FEL lamps are routinely calibrated from 250 to 2,400 nm. Deuterium lamps are routinely calibrated from 200 to 350 nm. The spectral irradiance values transferred to the deuterium lamps in the spectral range 200 to 250 nm are based on the hydrogen and blackbody line arcs developed primarily for use in the vacuum ultraviolet." The deuterium lamps are described by Klose and Bridges [25]. From 250 to 350 nm, the reported spectral irradiance values are transferred from the modified type FEL lamps. The equipment used for the deuterium lamp calibrations is identical to that used for the modified type FEL lamp calibrations, and the measurement procedures are similar [18].

The FEL lamps issued by the NIST presently are calibrated as standards of spectral irradiance at 31 selected wavelengths from 250 to 2400 nm. The deuterium lamp standards are calibrated at 16 wavelengths from 200 to 350 nm. The uncertainties in these calibrations are estimated at the level of three standard deviations from the mean to range from 2.2% at 250 nm to 1.0% at 654.6 nm to 6.5% at 2,400 nm for the FEL lamps and 7.5% at 200 and 210 nm and 5% from 220 nm to 250 nm for the deuterium lamps. The deuterium lamps are used mainly in the range from 200 to 250 nm because of the increased output flux compared with the FEL lamps. These standards are described in more detail by Walker, et al [26].

Standards of spectral irradiance have found wide use in the calibration of spectroradiometric and other equipment. For many types of spectroradiometric calibrations, dealing for example with radiant emissions from electrical plasmas, furnaces, or incandescent surfaces, a standard of spectral radiance is very useful. The NIST issues lamp standards of spectral radiance calibrated at 34 wavelengths from 225 to 2400 nm. The NIST scale of spectral radiance is based upon a blackbody standard. In 1960, Stair developed tungsten flat, ribbon-shaped filament lamp

standards of spectral radiance for 250 to 2,600 nm through comparison with black-body sources, with accuracies ranging from 8% to 3% [27]. In 1965, Kostkowski developed more accurate lamp standards in the 210 to 850-nm range, using a spectroradiometer also used as a photoelectric pyrometer, relying on a variable temperature blackbody as a comparison source [28]. The present NIST spectral radiance scale derives from these procedures and spectroradiometric instrumentation but has undergone many refinements, detailed in the relevant NIST special publication [29]:

Determination of the spectroradiometer response linearity allows more accurate comparison of unequal sources. The temperature of the variable temperature blackbody is determined by nearly direct comparison to the gold point blackbody. The wavelength range has been extended to 2,400 nm. Accuracy estimates have been more rigorously evaluated. The instrument has been incorporated into an automated calibration facility, utilizing a computer for operational control, data acquisition, and data reduction.

9.7.4 Absolute Radiometry

The ultimate accuracy of the total irradiance scales described above depends upon the accuracy with which the thermodynamic temperature scale can be realized and upon the validity of the Stephan-Boltzmann and Planck radiation laws. There is another approach, however, that does not appeal to these scales and laws. It is termed *absolute radiometry* because special detectors are made that are based on very well known physical principles allowing one to know their absolute spectral responses without appeal to the use of a calibrated source. This new development in detector technology leads to the second method of calibration mentioned in Section 9.11, calibration by comparison with already calibrated or self-calibrated radiometers or photometers. Before discussing this, however, a brief description of absolute radiometry is offered.

The most common form of absolute radiometer is based on the fact that temperature, electrical current, and voltage can be measured very precisely. Since the electrical power consumed by a resistor is equal to the product of the voltage across the resistor and the current through it, one can heat a resistive load to some temperature and know the power needed to keep it at that temperature precisely. One places a resistive heater on the back of a black, flat spectral response, metal disk of high conductivity, and then exposes the surface to radiation allowing it to heat up until it reaches thermal equilibrium. The detector element is placed inside a metal cavity enclosure of high mass. A schematic diagram of the arrangement called an *absolute cavity radiometer*, is shown in Figure 9.3.

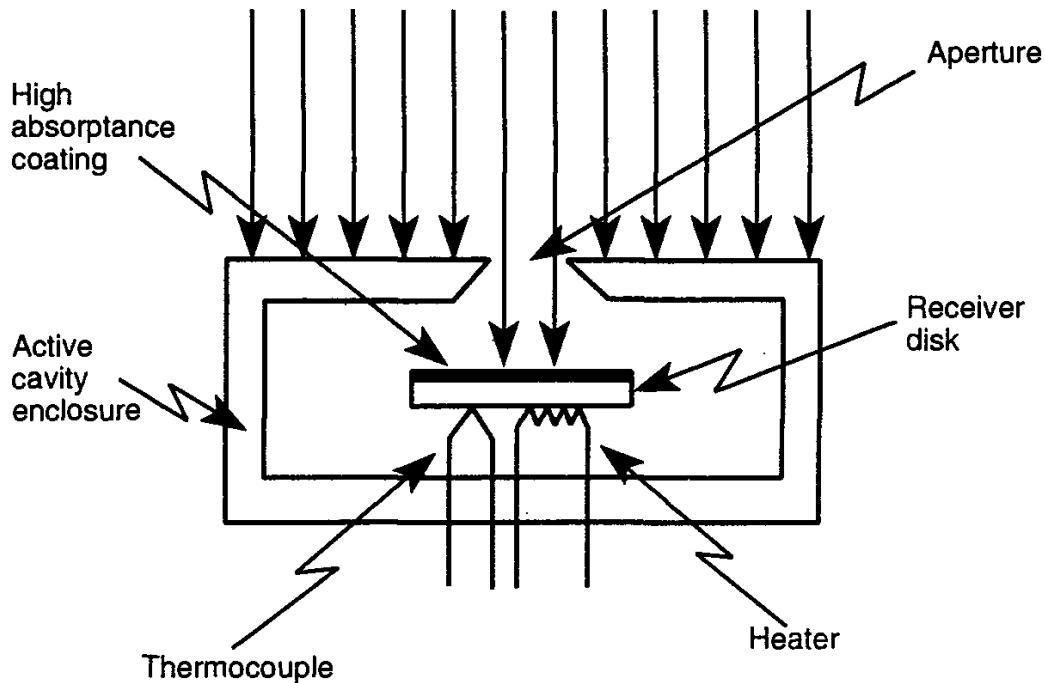


Figure 9.3 Illustration of the principles of operation of an absolute cavity radiometer.

The temperature of the surface at equilibrium is measured and recorded. Then the metal disk is shielded (through closure of the shutter) from the radiation source in such a manner that the shield itself introduces negligible radiation onto the metal surface and does not alter the means of escape of heat from the metal (or alters it in a precisely known way). Electrical power is then applied to the resistor attached to the back of the metal and the current is varied until the same equilibrium temperature is reached as before. The electrical power is then determined from the current and voltage through the resistor and this power is approximately equal to the radiant flux that was received by the detector when it was irradiated with the radiant flux being measured. There are a number of known systematic errors in this process. Applying corrections for each of these to the measured power heating the resistor converts this value to one which is very close to the radiant flux reaching the absorbing surface with the shutter open, without reference to any standard source or other radiometer.

The spectral response of an absolute radiometer depends upon the spectral absorptance of the coating for the metal disk inside the cavity, which receives the radiation. As described in Chapter 7, materials such as Parson's Black have been found with very high and constant spectral absorptance over a wide wavelength range.

This method of absolute radiometry is used for very precise measurements of direct beam solar radiation, both terrestrially and extraterrestrially. For the most

precise measurements, the design is improved in several ways. First, the absorbing disk is converted into a conically shaped light trap or cavity. The interior of the cavity is coated with a specular black paint that yields an effective cavity absorptance of 0.999999, with an uncertainty of ± 0.000003 , according to Iqbal [30]. Next, an additional secondary, inverted, and opposing conical cavity is attached beneath the primary one for thermal compensation purposes. Both cavities have identical absorbing surfaces and the temperature-measuring thermocouples have their cold junctions attached to a heat sink mass that surrounds the cavities. A wire electrical heater is wound around both cavity cones. The rear compensating cavity views a blackbody at the temperature of the heat sink [31]. The switch from flat (diffusing) black paint to a gloss (specularly reflecting) coating was made to avoid errors due to departure of the flat black paints from being truly diffuse (Lambertian). Another problem resulted from this switch to gloss coatings. When a conical cavity is used, a meniscus forms at the apex of the cavity when the paint is applied, causing a retroreflectance back out along the axis of the cone. Various approaches including inverted cones and more complicated geometries have been tried in an attempt to avert the meniscus. One of the best is to terminate the cone in a curved tube through which the paint can be sucked out before drying [32].

These absolute cavity radiometers can be operated in either a passive or an active mode. In the former, the procedure is as described previously: The heater is adjusted with the shutter closed to produce the same signal as when the shutter is open and the heater is completely off. The irradiance is equal to the electrical power applied divided by the area of a precise beam delimiting stop placed just in front of the primary cavity. An additional small correction factor, typically on the order of 1.005, is applied to this result to account for the small residual systematic errors that are known to be present in the measurement. In this mode, there is a typically long period of operation with the shutter open to allow the instrument to stabilize and to ensure precise alignment with the source. The shutter is then closed and the heater turned on and adjusted until precisely the same temperature is achieved for the cavity. The electrical heating rate is determined for this condition and the shutter-open measurement is repeated to account for any residual drift in solar irradiance between open and closed readings and in other factors affecting the measurement.

In the *active cavity* mode of operation, the input radiation is mechanically chopped and electronic (servo) control circuits are used to maintain the same signal from the front detector when it is being electrically heated during dark periods as when it is heated only by the incoming radiation. The power needed for this is continuously monitored and output. The thermal time constant of the cavity should be as short as possible for this measurement, to permit relatively short chopping intervals.

The uncertainty for solar irradiance measurements with these devices is claimed to be $\pm 0.3\%$ or better. To achieve such accuracies, absolute cavity radi-

ometers require careful characterization of their physical and optical properties in order to correct for the small but significant differences between the electrical and solar radiant heating mechanisms employed. These differences arise from defects in the effective absorptances of the cavities, error in knowledge of the primary aperture area, the presence of stray light, voltage drops across the wires leading to the heating elements and thermocouples, and other effects. For more information about absolute cavity solar radiometers, the reader is directed to [30] and [31].

Another method of absolute radiometry was developed at NIST in the early 1980s. It is based on the physics of silicon junction diodes and has resulted in a self-calibration procedure described by Zalewski and Geist [33], whose paper can be consulted for the details of the technique and the principles involved. The method was developed when investigating the effect of interface recombination on oxide-coated junction diodes. In combination with conventional reverse-bias measurements, this technique can be used to determine the absolute quantum efficiency of one type of shallow junction, oxide-coated photodiode. An uncertainty of less than $\pm 0.05\%$ was indicated in experiments with a 632.8-nm He Ne laser.

This new self-calibration capability represents a significant step forward in detector calibrations. When done correctly, the procedure yields a calibration that is more accurate than that of any other reasonably available radiometric standard [34]. An experimental scale of spectral irradiance based on an absolute silicon photodetector, was established at NBS in 1987 for the wavelength range 400 to 700 nm. The new scale made use of interference filters and a 100% quantum efficient detector invented by researchers at NBS and United Detector Technology, Inc. The new scale is independent of either of the traditional scales based on the thermal physics of blackbodies or absolute cavity radiometers [35].

In 1988, the radiometric physics division of NBS explained observed discrepancies between radiometric scales based on blackbody radiators and those based on absolute detectors. The experimenters began by measuring the temperature of freezing gold radiometrically, using absolute detectors. The experiment compared the two scales directly, eliminating the possibility of transfer errors. An integrating sphere irradiated by a dye laser was used as a radiance source to simulate a blackbody source at one wavelength. The sphere then became a non-polarized Lambertian radiator having the same size aperture as the gold point blackbody. The radiance of the sphere was compared to the radiance of the blackbody. Knowing the radiance of the gold point blackbody during the freeze, the temperature was computed using Planck's radiation law. The experiment showed that the temperature of freezing gold (1,337.58K) should be lowered by 0.2K.

More recent work [36] shows that, in principle, a single inversion layer photodiode can be used to establish an accurate scale of radiometry for the visible and near-ultraviolet using commonly available laboratory equipment. Detailed measurements show that the electrically based absolute cavity radiometer scale, the

absolute quantum efficiency detector scale, and the Planckian radiator scale all agree within the estimated uncertainties (approximately 1%).

Currently, NIST offers a photodetector spectral response calibration transfer program. The instruments used to transfer the calibration from NIST detectors to the user's secondary standard detectors are rented from NIST rather than purchased and they include sensitive and stable current-input amplification circuits. The user needs to have a tunable monochromatic source, such as a monochromator or a set of narrowband spectral filters combined with a suitable source of radiation. The linearity of the supplied photodetectors and amplifiers is calibrated over the radiant flux range from 10^{-3} to 10^{-7}W . The standard photodetectors are calibrated in units of A/W in 10-nm intervals from 250 to 960 nm. They are also calibrated at two additional spectral lines in the infrared: 1,014 and 1,064 nm. A calibrated 0.5-cm² aperture stop is included to enable proper transfer of the calibration. A more detailed description of the calibration transfer program offered by NIST can be found in the relevant NIST publication [18].

9.7.5 Standard Illuminance and Luminance Sources

Standard sources of radiance and irradiance that emit usable quantities of radiation over the visible portion of the spectrum can be used as standards for the calibration of photometers if the photometric outputs of these sources is known. Commercial radiometric and photometric standards laboratories generally can supply photometric calibrations for their radiometric sources for modest additional cost. The most common source is the incandescent filament lamp, with its characteristic spectral output distribution (shown in Figures 3.9 and 6.19). If the primary use of the photometer being calibrated is to measure light levels derived from sources with similar spectral distributions, and if the V-lambda correction of the photometer is good, then use of tungsten filament standard lamps is an acceptable means of calibration. If the photometer is intended for measurement of radiation with substantially different spectral distribution and the V-lambda correction is not good, then significant measurement errors can result from calibration using tungsten sources. Fortunately, other standard spectral distributions have been defined. They are based on phases of daylight (primarily for colorimetric applications). Sources exhibiting approximations of these distributions have been developed [37]. For cases of imperfect V-lambda correction, it is recommended that calibration sources be used that more closely match the distributions to be measured with the photometer.

For an extensive discussion of the characterization of radiometer and photometer performance, including spectral and angular variations and departures from ideal response, the reader is referred to Publication 53 of the CIE [38] on this subject.

NIST presently offers a set of photometric standards and calibration services that includes standards for luminous intensity, luminous flux, and color temperature. Inside-frosted lamp standards of luminous intensity are calibrated at NIST by a substitution method on a calibrated horizontal bench photometer. The bench photometer is calibrated at the time of the measurements against a source having approximately the same color temperature as CIE-standard illuminant A (described in Section 11.8.1). The luminous intensity scale is derived from the NIST spectral irradiance scale described in Section 9.7.2 and is maintained in a group of nine 500W inside-frosted lamps. The luminous intensity of these lamps was determined by measuring the spectral irradiance of each lamp on an automated high-accuracy spectroradiometer that serves as the transfer instrument, and then calculating the luminous intensity according to (1.26) and (2.5). The inside-frosted lamps are calibrated in three sizes, consuming 100, 500, and 1000W of electrical power, and having corresponding luminous intensities of approximately 90, 700, and 1400 candelas, respectively.

The total luminous flux emitted by a source is determined with a goniometric method at NIST. The standard source calibrated in this way is then used in a substitution method to transfer the calibration to another source using an integrating sphere. In 1980, a group of opal bulb lamps that served as the NIST primary reference standard for luminous flux was recalibrated in accordance with the 1979 international redefinition of photometric units [39] using the goniometric distribution of the lamps. This calibration was then transferred, using the substitution method, to three working groups, consisting of lamps ranging from 100 to 500W in electrical power consumption. Lamps calibrated for issuance to the public are measured against the working group in a manner similar to the calibration of the working groups against the NIST primary reference standards.

NIST incandescent lamp color temperature standards are calibrated using a red-blue ratio substitution method. The basic instrumentation consists of a 20-cm diameter integrating sphere and a commercially available telephotometer.

For more information on luminous intensity, luminous flux, and color temperature calibrations at NIST, the relevant NIST publication can be consulted [40].

9.7.6 Radiometer/Photometer Calibration Using Standard Sources

Figure 9.2 illustrates a typical arrangement for the calibration of an irradiance or illuminance meter using a standard source. It is important that radiation from the standard source reach the radiometer directly, without any stray light from multiple reflections or other sources reaching it. The configuration in Figure 9.2(a) is for irradiance or illuminance calibration, and the configuration in Figure 9.2(b), using a flashed opal glass insert, is used for radiance or luminance calibrations. Of course, there are variations on this basic arrangement, including the one illustrated in

Figure 9.4 using an integrating sphere and a series of internal lamps to provide a constant and uniform radiance or luminance over a large area.

It is important to note that the incandescent lamp spectrum changes fairly rapidly with wavelength over the visible range from 360 to 780 nm. Thus, any errors in the V-lambda correcting filter in the photometer being calibrated could be magnified. A photometer calibrated with an incandescent lamp spectrum might not read correctly when measuring sources such as sky light having different spectral distributions. For this reason, it is generally best to calibrate photometers having imperfect V-lambda correction with a spectral distribution somewhat close to the distributions most likely to be encountered in unknown sources to be measured with them.

9.7.7 Spectroradiometer Calibration

There are two calibrations needed for spectroradiometers. First is the wavelength calibration. This is most often accomplished through the use of special low-pressure gas-discharge lamps that emit radiation at very specific and narrow wavelengths. The spectroradiometer wavelength scanning mechanism is adjusted so that it reads the wavelengths of these spectral lines correctly.

Second is the radiometric calibration. The spectral responses of most spectroradiometers vary strongly over their wavelength range of intended use. They are calibrated by recording their output signals while they measure the spectral radiant flux emanating from a standard source of spectral radiance or irradiance.

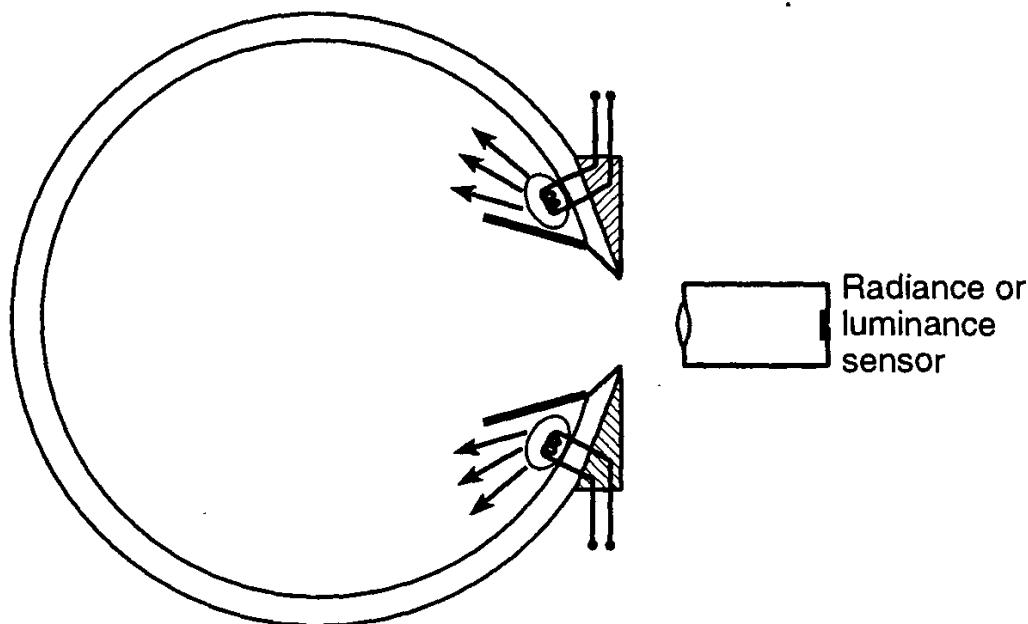


Figure 9.4 Illustration of the use of several sources inside an integrating sphere to provide a source of constant radiance and luminance.

The output signals are then corrected at each wavelength so that these instruments read the correct values of the input spectral flux distribution. If a spectroradiometer is susceptible to drift in sensitivity, or when highly accurate results are needed, it is common to perform a calibrating scan using a working standard source before and after a scan of a test source. With this procedure, the strong variations in system spectral response are compensated for in the final results. As with broadband measurements, best results are obtained when the spectral distribution of the standard source has approximately the same shape as that of the test source being measured.

9.7.8 National Standards Laboratories

It is clear that anyone concerned with calibration of radiometers and photometers can benefit greatly from the work of the radiometric physics division of the National Institute of Standards and Technology and its counterparts in other countries. It is beyond the scope of this book to list all the standards laboratories around the world. Fortunately, the Institute of Physics has recently published a compilation of National Standards Laboratories offering services in what it calls “electromagnetic metrology” [41]. Readers outside the United States are referred to their own national standards laboratory for further information on standards services available in their countries.

EXAMPLE PROBLEM 9.1

Problem: You have a light meter whose electronic features you like, but you test its cosine response and find that it is very poor for angles above 45 degrees. You wish to improve this for certain measurements and consider putting an integrating sphere over the sensitive area of your meter for these measurements to improve its cosine response. Estimate the loss of signal that will be produced by the integrating sphere.

Solution: The cosine corrector on your light meter measures 20-mm in diameter and it rises 3-mm above the housing surrounding it. You decide to make both entrance and exit ports in the integrating sphere 21-mm in diameter so they can be used interchangeably, and you choose to use pressed TFE powder as a coating. The ports will have an area of about 3.46 cm^2 . You know that spun aluminum spheres in the size range from around 50 mm to 200-mm diameter are made of aluminum having a 1-mm thickness. You plan to have a sphere coating thickness of 8 mm and to taper this down to 2 mm at the two ports. The reflectance of TFE powder is 0.990 (see Section 8.11). You think that a sphere diameter of 10 cm would be about right, giving a sphere area of $4\pi 5^2 = 314 \text{ cm}^2$ so that f_2 in (8.19) is 0.011. Substituting this value and 0.99 for ρ_w into (8.19) yields a sphere efficiency

of 0.343. This indicates approximately the loss you can expect when you place the integrating sphere over the existing cosine receptor of your light meter.

EXAMPLE PROBLEM 9.2

Problem: A light meter is used to measure the average illuminance of a laser beam incident directly and normally on it. The dark reading is 8 lux and the laser-illuminated reading is 990 lux. The laser beam wavelength is 670 nm and it has a diameter of 4 mm. The sensitive area of the light meter has a diameter of 10 mm. Assume that the V-lambda correction of the meter is perfect. Convert the illuminance reading from the meter into luminous and radiant flux.

Solution: Since the laser beam underfills the detector area, the total flux in the laser beam will be given by the measured illuminance multiplied by the sensitive area of the light meter (the area for which it is calibrated). This area is $\pi (0.5 \text{ cm})^2 = 7.85 \times 10^{-5} \text{ m}^2$. Multiplying the difference ($990 - 8 \text{ lm/m}^2$) in readings by this area yields the luminous flux 0.077 lumens in the laser beam. Only a single wavelength is present in this flux. V-lambda is known (from Table 2.1) to be 0.032 at this wavelength. For monochromatic radiation, the conversion from spectral radiant flux Φ_λ to luminous flux Φ_v is given (2.1) by

$$\Phi_v = 683 \times K_\lambda \times \Phi_\lambda \times \Delta\lambda = 683 \times 0.032 \times \Phi_e$$

which we solve for the spectral radiant flux $\Phi_e = 0.077/(683 \times 0.032) = 3.5 \text{ mW}$. Checking the literature supplied with the laser we find that the manufacturer has rated its output power at greater than 3.0 mW.

REFERENCES

- [1] Palmer, J. M., and M. G. Tomasko. "Broadband Radiometry with Spectrally Selective Detectors," *Optics Letters*, Vol. 5, 1980, pp. 208–210.
- [2] Nicodemus, F. E. *Appl. Opt.*, Vol. 12, 1973, p. 2960.
- [3] *Ibid.*, p. 1649.
- [4] Kelton, G., et al. *Infrared Physics*, Vol. 3, 1963, p. 139.
- [5] Miyake, I. "Recent Topics in the Measurement of Solar Radiation," *Int. J. Solar Energy*, Vol. 9, 1990, pp. 47–56 (1990). Translated from the Japanese, *Journal of Japan Solar Energy Society*, Vol. 14(5), 1988, pp. 34–40.
- [6] Kaufman, J. E., and J. F. Christensen. *IES Lighting Handbook, Reference Volume*, IESNA: New York, NY, 1984.
- [7] *Lighting Handbook*, Publ. No. A-9950, North American Philips Lighting Corporation, 1984.
- [8] Boaz, J. N., ed. *Architectural Graphic Standards*, 5th ed., John Wiley & Sons: New York, NY, 1970; 8th ed., ed. by J. R. Hoke, 1988.
- [9] Murdoch, J. B. *Illumination Engineering—From Edison's Lamp to the Laser*, Macmillan: New York, NY, 1985.

- [10] *Physics Today Buyers' Guide*, 335 East 45 St., New York, NY 10017.
- [11] *American Laboratory Buyers' Guide*, P. O. Box 870, Shelton CT 06484.
- [12] *Analytical Chemistry Lab Guide*, Amer. Chem. Soc., 1155 16th St. NW, Washington, DC 20036.
- [13] *ISA Directory of Instrumentation*, P. O. Box 12277, Research Triangle Park, NC 27709.
- [14] *The Photonics Directory*, Books 1-4, 40th ed., Pittsfield, MA: Laurin Publishing Co., 1994.
- [15] CI Systems, Inc., 5137 Claretton Dr., Suite 220, Agoura Hills, CA 91301.
- [16] McSparron, D. A. "The Future of Luminance Standards," *Optical Radiation News*, No. 37, Radiometric Physics Division, National Bureau of Standards, Gaithersburg, MD, Oct. 1981.
- [17] Lind, M. A., E. F. Zalewski, and J. B. Fowler. NBS Tech. Note 594-13, 1977.
- [18] Zalewski, E. F. "NBS Measurement Services: The NBS Photodetector Spectral Response Calibration Program," NBS Special Publ. 250-17, March 1988.
- [19] Coblenz, W. W. Bull. Bur. Stds., 11, 87, (1914).
- [20] Stair, R., W. E. Schneider, and W. E. Fussell. *Applied Optics*, Vol. 55, 1965, p. 1567.
- [21] Stair, R., W. E. Schneider, and J. K. Jackson. *Applied Optics*, Vol. 2, 1963, p. 1151.
- [22] Saunders, R. D., and J. B. Shumaker. NBS Tech. Note, 594-13, 1977.
- [23] McSparron, D. A. "New D₂ Lamp Standards of Spectral Irradiance," *Optical Radiation News*, Radiometric Physics Division, National Bureau of Standards, Gaithersburg, MD, No. 32, July 1980.
- [24] Walker, J. H., et al. "The NBS Scale of Spectral Irradiance," *J. Res. of the NBS*, Vol. 93, No. 1, Jan.-Feb. 1988, pp. 7-20.
- [25] Klose, J. Z., and J. M. Bridges. "NBS Measurement Services: Radiometric Standards in the Vacuum Ultraviolet," Natl. Bur. Stand. (U. S.) Spec. Publ. 250-3, June 1987.
- [26] Walker, J. H., et al. "NBS Measurement Services: Spectral Irradiance Calibrations," NBS (now NIST) Spec. Publ. 250-20, Gaithersburg, MD, Sept. 1987.
- [27] Stair, R., R. G. Johnston, and E. W. Halbach. "Standard of Spectral Radiance for the Region of 0.25 to 2.6 Microns," *J. Res. Nat. Bur. Stand. (U.S.)*, 64A(4), July-Aug. 1960, pp. 291-296.
- [28] Kostkowski, H. J., D. E. Erminy, and A. T. Hattenburg. "High Accuracy Spectral Radiance Calibration of Tungsten-Strip Lamps," from *Advances in Geophysics*, Vol. 14, Academic Press: New York, NY, 1970, pp. 111-127.
- [29] Walker, J. H., R. D. Saunders, and A. T. Hattenburg. "NBS Measurement Services: Spectral Radiance Calibrations," NBS (now NIST) Spec. Publ. 250-1, Jan. 1987.
- [30] Iqbal, M. *An Introduction to Solar Radiation*, Academic Press: New York, NY, 1983, p. 346.
- [31] Zerlaut, G. "Solar Radiation Instrumentation," Ch. 5 of *Solar Resources*, ed. by R. L. Hulstrom, The MIT Press: Cambridge, MA, 1989.
- [32] Zalewski, G. F. "New Reflectometer for Measuring Very Low Reflectance Losses from Electrically Calibrated Detectors," *Optical Radiation News*, Radiometric Physics Div., National Bureau of Standards, No. 30, Jan. 1980.
- [33] Zalewski, E. F., and J. Geist. "Silicon Photodiode Absolute Spectral Response Self Calibration," *Appl. Opt.*, Vol. 19, 1980, pp. 1214-1216. Note that this issue of *Applied Optics* happens to contain a number of papers on retroreflection measurements, many by authors at NBS. One by Rennilson, "Chromaticity Measurements of Retroreflective Materials Under Nighttime Geometry," pp. 1260-1267, includes examples of the use of a spectroradiometric system, a goniometer, and comparisons with different standard source spectral distributions.
- [34] Zalewski, E. F. "Absolute Spectral Response Intercomparison," *Optical Radiation News*, No. 39, April 1982. Radiometric Physics Division, National Bureau of Standards, Gaithersburg, MD.
- [35] Hughes, C. G. "A Spectral Irradiance Scale Using Silicon Detectors," *Optical Radiation News*, No. 47, Spring 1987, Council for Optical Radiation Measurements, Baltimore, MD.
- [36] Booker, R. L., and J. L. Geist. *Appl. Opt.*, Vol. 23, 1984, p. 1940.
- [37] CIE. "A Method for Assessing the Quality of Daylight Simulators for Colorimetry," Pub. No. 51, 1981. *Commission Internationale de l'Eclairage*, CIE Central Bureau, A-1030 Vienna, Kegel-

- gasse 27, Austria. Available in the U.S. from USNC/CIE Publications Office, c/o TLA Lighting Consultants, 7 Pond Street, Salem, MA 01970.
- [38] CIE. "Methods of Characterizing the Performance of Radiometers and Photometers," Publ. No. 53, 1982. *Commission Internationale de l'Eclairage*, CIE Central Bureau, A-1030 Vienna, Kegelgasse 27, Austria. Available in the U.S. from USNC/CIE Publications Office, c/o TLA Lighting Consultants, 7 Pond Street, Salem, MA 01970.
- [39] Giacomo, P. "News from the BIPM," *Metrologia*, Vol. 16, 1980, pp. 55-61.
- [40] Booker, R. L., and D. A. McSparron. "NBS Measurement Services: Photometric Calibrations," NBS (Now NIST) Spec. Publ. 250-15, Oct. 1987.
- [41] Bailey, A. E., ed. *URSI Register of National Standards Laboratories for Electromagnetic Metrology*, IOP Publishing, Inc.: Public Ledger Bldg., Suite 1035, Independence Square, Philadelphia, PA 19106, 1990. Readers outside the U.S. can contact IOP Publishing, Ltd., Techno House, Redcliff Way, Bristol BS16NX, UK.

Chapter 10

Metric Primer and Additional Radiometric and Photometric Quantities and Units

10.1 INTRODUCTION

This text adheres almost exclusively to the SI (Système Internationale) system of units known as the metric system. Although there are several alternative systems of units, most notably the inch-pound (IP), also called the U.S. customary system of units, it is thought that many of the concepts introduced in this book and used in solving problems in radiometry and photometry should not be burdened with having to deal with a conflicting and inconsistent set of units and terminology. Discussion of the alternative units is therefore saved for this chapter. Fortunately, for the last few decades there has been a growing trend toward complete metrification in radiometry and photometry, and the watt, lumen, lux, and candle are now almost universally recognized by workers in those fields.

On the other hand, many previously published books and technical papers are filled with a seemingly hodgepodge pattern of changing symbols, units, and nomenclature. The interested reader needs only to scan through papers by Meyer-Arendt [1] and Biberman [2] or old technical dictionaries, textbooks, and reference volumes to get a glimpse of the difficulties faced by previous students of radiometry and photometry. In one field, solar radiation was measured and reported in langleys/s or langleys per min, in another it was in $\text{cal} \cdot \text{s}^{-1} \cdot \text{cm}^{-2}$, and in still another talbots/m² were used. Meyer-Arendt points out that many of the old units—such as the nox, phot, glim, skot, and scot (identical), as well as the bril and brill (different)—are so antiquated as to be of interest only to the historian.

Due to the current trend toward metrification and standardization, as represented by ANSI/IES standard RP-16 [3] and the *CIE/CEI International Lighting Vocabulary* [4–5], this chapter should not be needed for current work in radiometry and photometry. Since the IP system is still in widespread use in the United States, all that would be needed are short sections at the appropriate places providing the simple conversions from SI to IP. However, adherence to the ANSI/IES and CIE/

CEI standards is by no means complete in current technical literature and publications prior to the recent trend toward improved standardization. For access to this literature, more material is needed to describe the relationships and conversion factors amongst the various units. The purposes of this chapter are therefore to

- Describe the metric system and summarize metric terminology for the most frequently encountered quantities in radiometry and photometry;
- Provide definitions and conversion factors for the most commonly found radiometric and photometric quantities in the IP system of units;
- Provide a listing with conversion factors to SI units for a variety of old and discontinued units.

10.2 THE SI SYSTEM OF UNITS

10.2.1 Basic Metric Principles

The last two centuries have seen many versions of the metric system [6]. The current modernized version receives its name from Le System International d'Units, established in 1960 by international agreement. The Bureau International des Poids et Mesures (BIPM) regularly publishes a document containing revisions and new recommendations of the international conferences and committees dealing with terminology and units. The International Standards Organization (ISO) publishes standards on the practical uses of the SI system in a variety of fields. The United States National Institute of Standards and Technology (NIST) periodically publishes an English language translation of the latest BIPM publication *Le Systeme International d'Units*. The 1991 edition [7] was used in the preparation of this chapter.

One difference of the current metric system from older ones is the adherence to standard prefixes for standard orders of magnitude. These are listed in Table 10.1, where we discover immediately that distance units such as the angstrom, micron, and millimicron, previously used for wavelength, are no longer permitted. The micron should be replaced by the micrometer (μm), the millimicron by nanometer (nm) and the Angstrom by 0.1 nm. Another example is the cc or cubic centimeter which should be replaced with cm^3 .

The following simple rules govern the use of the prefixes given in Table 10.1:

- Prefix symbols are printed in roman type without spacing between the prefix symbol and the unit symbol: cm^2 , nm, μm , klx.
- The grouped unit symbol plus its prefix is inseparable but may be raised to a positive or negative power and combined with other unit symbols: $1 \text{ cm}^2 = (10^{-2} \text{ m})^2 = 10^{-4} \text{ m}^2$.
- No more than one prefix can be used at a time: 1 nm is acceptable but 1 $\mu\mu\text{m}$ is not.
- A prefix should never be used alone, except in descriptions of systems of units.

Table 10.1
SI Prefixes

Factor	Prefix	Symbol	Factor	Prefix	Symbol
10^{24}	yotta	Y	10^{-1}	deci	d
10^{21}	zetta	Z	10^{-2}	centi	c
10^{18}	exa	E	10^{-3}	milli	m
10^{15}	peta	P	10^{-6}	micro	μ
10^{12}	tera	T	10^{-9}	nano	n
10^9	giga	G	10^{-12}	pico	p
10^6	mega	M	10^{-15}	femto	f
10^3	kilo	k	10^{-18}	atto	a
10^2	hecto	h	10^{-214}	zepto	z
10^1	decka, deca	da	10^{-24}	yocto	y

There are three classes of units in the SI system:

- Base units and symbols: meter (m), kilogram (kg), second (s), ampere (A), kelvin (K), mole (mol), and candela (cd);
- Derived units: such as joule ($= \text{kg} \cdot \text{m}^2 \cdot \text{s}^{-2} = \text{N} \cdot \text{m}$), watt ($= \text{J} \cdot \text{s}^{-1}$), lumen ($= \text{cd} \cdot \text{sr}$), and lux ($= \text{lm} \cdot \text{m}^{-2}$);
- Supplementary units: combinations of the above units and including the SI units for plane and solid angle to be described below

Although this division is somewhat arbitrary, the international body decided that setting up seven base units regarded as dimensionally independent would simplify and improve international relations, teaching, and scientific and engineering work. The derived units can be formed by combining base units according to algebraic relations linking the corresponding physical quantities. The derived units are all therefore combinations of the base units, determined by the laws of physics and chemistry. To further simplify terminology, some combinations of units are given specific names such as the lumen, which stands for $\text{cd} \cdot \text{sr}$ and the lux, which stands for $\text{cd} \cdot \text{sr} \cdot \text{m}^{-2}$.

The three classes form what is called a *coherent set of units*, a system of units mutually related by rules of multiplication and division *without any numerical factor required*. This is a remarkable characteristic of the SI system of units. Conversion factors are not needed to relate different units in this system to each other.

It is emphasized [7] that each physical quantity has only one SI unit, even if the name of it can be expressed in different forms, using the prefixes listed in Table 10.1 for example. The inverse is not true: the same SI unit name can correspond to several different quantities. An example of the latter is the newton · meter, which is the unit of both torque and energy ($1 \text{ N} \cdot \text{m} = 1 \text{ joule}$). Throughout this chapter, the symbol · is used to indicate multiplication and to separate units used in combinations. The international standard permits this symbol to be replaced with a space. However, standard ANSI/IEEE 268-1982 Metric Practice states that in

United States practice only the raised dot (·) is to be commonly used. This practice is followed throughout this book.

In 1969, the International Committee for Weights and Measures recognized that there are several additional non-SI units which may be acceptably used with SI units. They are the minute, hour, and day of time, the degree, minute, and second of angle, the liter (10^{-3}m^3), and the tonne (10^3 kg). These should be used judiciously in order not to lose the coherence that makes the SI system so attractive.

10.2.2 Metric Units for Radiometry and Photometry

Units for the fundamental quantities of radiometry and photometry are defined in terms of their base units in Table 10.2. A number of additional quantities in radiometry and photometry are defined in the *International Lighting Vocabulary* but are not used in this book. Most are straightforward combinations of the quantities listed in Table 10.2. Examples include *spherical irradiance* and *illuminance* (also called *radiant* and *luminous fluence rates*), defined to be the radiant or luminous flux of all radiation incident on the outer surface of an infinitesimally small sphere centered at the point of definition divided by the diametrical cross-sectional area of that sphere. The units are $\text{W}\cdot\text{m}^{-2}$ and lux, respectively. Similar definitions can be found for cylindrical irradiance and radiant cylindrical and spherical exposure. There are also definitions for photon exitance and photon exposure. The symbol for exposure is H and the usual subscripts e , v , and p are used to indicate *radiant*, *luminous*, and *photon* exposure, respectively.

10.3 THE I-P SYSTEM OF UNITS

The most prominent alternative to the metric system, is the inch-pound or the so-called “English” system of units. In this system the foot and British thermal unit (Btu) are units for length and energy, respectively. Table 10.3 gives the correspondence between radiometric and photometric quantities based upon these units.

10.4 PHOTON FLUX UNITS

Section 1.9 dealt with photon flux. According to LI-COR, Inc. [8] the *mole of photons* and the *einstein* are equivalent units used to designate Avogadro’s number of photons, 6.022×10^{23} photons.

The einstein has been used in plant science but according to LiCor, most of the relevant societies now recommend the use of the mole (mol) since this is an SI unit. In some other fields, the einstein has been used to indicate the quantity of radiant energy in Avagadro’s number of photons [9].

Table 10.2
SI Units for Radiometry and Photometry

Quantity	Name	Symbol	Expression in Terms of Other Units	Expression in SI Base Units
Radiant energy	joule	J	N·m	$\text{m}^2 \cdot \text{kg} \cdot \text{s}^{-2}$
Radiant flux	watt	W	J/s	$\text{m}^2 \cdot \text{kg} \cdot \text{s}^{-3}$
Plane angle	radian	rad		$\text{m} \cdot \text{m}^{-1}$
Solid angle	steradian	sr		$\text{m}^2 \cdot \text{m}^{-2}$
Radiant intensity	watt·sr ⁻¹	W/sr		$\text{m}^2 \cdot \text{kg} \cdot \text{s}^{-3} \cdot \text{sr}^{-1}$
Irradiance, Exitance	watt·m ⁻²	W/m ²		$\text{kg} \cdot \text{s}^{-3}$
Radiant exposure	joule·m ⁻²	j/m ²	W·s·m ⁻²	$\text{kg} \cdot \text{s}^{-2}$
Radiance	watt·m ⁻² ·sr ⁻¹	W·m ⁻² ·sr ⁻¹		$\text{kg} \cdot \text{s}^{-3} \cdot \text{sr}^{-1}$
Luminous intensity	candela	cd		cd
Luminous flux	lumen	lm		cd·sr
Quantity of light	lumen·sec	lm·s		cd·sr·s
Illuminance	lux	lx	lm/m ²	cd·sr·m ⁻²
Luminous exposure	lux·sec	lx·s	lumen·sec·m ⁻²	cd·sr·s·m ⁻²
Luminance	candela per square meter	cd/m ²	lm·m ⁻² ·sr ⁻¹	cd/m ²
Photon flux	number per second	s ⁻¹	quanta/s	s ⁻¹
Photon intensity	number per second per steradian	s ⁻¹ ·sr ⁻¹		s ⁻¹ ·sr ⁻¹
Photon irradiance	number per second per square meter	s ⁻¹ ·m ⁻²		s ⁻¹ ·m ⁻²
Photon radiance	number per second per square meter and per steradian	s ⁻¹ ·m ⁻² ·sr ⁻¹		s ⁻¹ ·m ⁻² ·sr ⁻¹

Note: N is the symbol for the Newton, the metric unit for force.

The Crop Science Society of America, Committee on Terminology [10] and other societies [11] deal with photon flux-related quantities that incorporate in some way the *activity* of different wavelengths; that is, the relative responsivity of different plant species to different wavelengths, in a quantity called by limnologists and oceanographers the *quantum scalar irradiance* or *photon spherical irradiance*. LI-COR has introduced the term *photosynthetic photon flux fluence rate* (PPFFR) [7]. A LI-COR publication [12] shows a plot of relative action versus wavelength for the range 300 to 1,100 nm. An approximate relationship over the range from 400 to 700 nm is indicated between a typical photosynthetic action spectrum and what is called the ideal quantum response, defined to be λ/hc . It therefore says that an ideal quantum sensor (for photosynthetic studies) has uniform quantum response over this wavelength range (and zero outside this range).

Table 10.3
SI to IP Conversion Factors for Radiometry and Photometry

Quantity	Units	Symbol	Conversion Equation
Radiant flux	British thermal unit per hour	Btu/h	$1 \text{ Btu/h} = 0.2928 \text{ W}$
	Watt	W	$1 \text{ W} = 1 \text{ J/s}$
	Erg per sec	erg/s	$1 \text{ erg/s} = 10^{-7} \text{ J/s} = 10^{-7} \text{ W}$
Radiant intensity	Btu per hour and per steradian	$\text{Btu} \cdot \text{h}^{-1} \cdot \text{sr}^{-1}$	$1 \text{ Btu} \cdot \text{h}^{-1} \cdot \text{sr}^{-1} = 0.2928 \text{ W} \cdot \text{sr}^{-1}$
	Watt per steradian	W/sr	$1 \text{ W/sr} = 1 \text{ J} \cdot \text{s}^{-1} \cdot \text{sr}^{-1}$
Irradiance	Btu per hour and per square foot	$\text{Btu} \cdot \text{h}^{-1} \cdot \text{ft}^{-2}$	$1 \text{ Btu} \cdot \text{h}^{-1} \cdot \text{ft}^{-2} = 3.1503 \text{ W/m}^2$
	Watt per square foot	W/ft ²	$1 \text{ W/ft}^2 = 10.76 \text{ W/m}^2$
Radiance	Btu per hour and per square foot and per steradian	$\text{Btu} \cdot \text{h}^{-1} \cdot \text{ft}^{-2} \cdot \text{sr}^{-1}$	$1 \text{ Btu} \cdot \text{h}^{-1} \cdot \text{ft}^{-2} \cdot \text{sr}^{-1} = 3.1503 \text{ W} \cdot \text{m}^{-2} \cdot \text{sr}^{-1}$
	Watt per square foot and per steradian	$\text{W} \cdot \text{ft}^{-2} \cdot \text{sr}^{-1}$	$1 \text{ W} \cdot \text{ft}^{-2} \cdot \text{sr}^{-1} = 10.76 \text{ W} \cdot \text{m}^{-2} \cdot \text{sr}^{-1}$
	Lumen	lm	$1 \text{ lm} = 1 \text{ cd} \cdot \text{sr}$
Luminous intensity	candela	cd	$1 \text{ cd} = 1 \text{ cd}$
Illuminance	Lumen per square foot	lm/ft ²	$1 \text{ lm/ft}^2 = 10.76 \text{ lm/m}^2 = 10.76 \text{ lx}$
Luminance	Foot candle	fc = lm/ft ²	$1 \text{ fc} = 10.76 \text{ lux}$
	candela per sq. foot	cd/ft ²	$1 \text{ cd/ft}^2 = 10.76 \text{ cd/m}^2$
	Foot Lambert*	fl = fc/(π sr)	$1 \text{ fl} = 10.76 \text{ lx}/(\pi \text{ sr}) = 1/\pi \text{ lm/ft}^2 = 3.425 \text{ cd/m}^2$

*The foot Lambert is the unit of luminance from a Lambertian surface in terms of the exitance of the surface divided by π . This is a very inappropriate term that invites misuse and is to be deprecated. Unfortunately, it is still used in some fields.

10.5 OTHER QUANTITIES AND UNITS

There are a number of additional units that are or have been used and which should not be used in the future. Many are listed in Table 10.4.

Terms such as foot candle, cm candle, ft lambert, and meter lambert are most unfortunate. They seem to indicate the multiplication of length by intensity or luminance, leading to ridiculous combinations of units without physical significance. All of the units in the second column of Table 10.4 should be avoided. They are given here only to facilitate their conversion to the appropriate SI units. Kingslake points out [13] that the Lambert is applicable only to Lambertian surfaces. Its use in non-Lambertian situations can lead to errors and misconceptions.

Some additional units can be related to SI units in photometry. One is the *Troland*. It is used by many vision scientists as a unit of retinal illuminance and is

Table 10.4
Conversion Factors for Deprecated Units

<i>Quantity</i>	<i>Units</i>	<i>Symbol</i>	<i>Conversion Equation</i>
Radiant flux	Calorie/second	cal/s	$1 \text{ cal/sec} = 4.184 \text{ J/sec}$ $1 \text{ cal/sec} = 4.184 \text{ W}$
	Calorie/min	cal/min	$1 \text{ cal/min} = 251 \text{ W}$
	$\text{Cal} \cdot \text{s}^{-1} \cdot \text{m}^{-2}$	$\text{cal}/(\text{s m}^2)$	$1 \text{ cal s}^{-1} \cdot \text{m}^{-2} = 4.184 \text{ W/m}^2$
Irradiance	$\text{Langley/s} = \text{g-cal} \cdot \text{cm}^{-2} \cdot \text{s}^{-1}$	Ly/s	$1 \text{ Ly/s} = 4.1868 \text{ W/cm}^2$ $1 \text{ Ly/s} = 4.1868 \times 10^4 \cdot \text{W/m}^2$ $1 \text{ Ly/s} = 41.868 \text{ kW/m}^2$
	Talbot		$1 \text{ Talbot} = 1 \text{ lm} \cdot \text{s}$
	Lumen-hour	lm · hr	$1 \text{ lm} \cdot \text{hr} = 3600 \text{ lm} \cdot \text{s}$
Luminous energy Also called “quantity of light”	candle, unit of luminous intensity in the U.S., France, and G.B. from 1948 to 1979		$1 \text{ candle} = 1 \text{ candela}$
Illuminance	candle-power	cp	$1 \text{ candle power} = 1 \text{ cd}$
	Foot candle = lm/ft ²	fc or ft cd	$10.764 \text{ lux} = 1 \text{ FC}$
	Phot = cm-candle = lm/cm ²	Phot	$1 \text{ Phot} = 1 \text{ lm/cm}^2 = 10^{-4} \text{ lx}$ 0.1 milliphot = 1 lx 1 phot = 929 FC 1 milliphot = 0.929 FC.
Luminance	Nox = millilux	Nox	$1 \text{ Nox} = 10^{-3} \text{ lux} = \text{mlx}$
	m-candle = lux	m-cd	$1 \text{ m-cd} = 1 \text{ lm/m}^2 = 1 \text{ lx}$
	nit	nit	$1 \text{ nit} = 1 \text{ cd/m}^2 = 1 \text{ lm} \cdot \text{m}^{-2} \cdot \text{sr}^{-1}$
Luminance	stilb	stilb	$1 \text{ stilb} = 1 \text{ cd/cm}^2 = 929 \text{ cd/ft}^2 = \pi \text{ Lamberts}$
	Lambert	Lambert	$1 \text{ Lambert} = 1/\pi \text{ cd/cm}^2 = 3.18 \text{ stilb} = 929 \text{ ft lambert}$
	ft-Lambert	fL	$1/\pi \text{ cd/ft}^2 = 10.76/\pi \text{ cd/m}^2$
Luminance	International apostilb = 10^{-4} Lambert, also called blondel	apostilb	$1 \text{ apostilb} = 1/\pi \text{ cd/m}^2 = 0.1 \text{ millilambert} = 1/\pi \times 10^{-4} \text{ stilb}$
	German Hefner apostilb = 0.9×10^{-4} Lambert	apostilb	$1 \text{ apostilb} = 0.09 \text{ millilambert}$
	skot	skot	$1 \text{ skot} = 1/\pi 10^{-3} \text{ cd/m}^2 = 1 \text{ milliapostilb}$
Luminance	glim	glim	$1 \text{ glim} = 10^{-3} \text{ ft-Lambert}$
	cd/ft ²		$10.764 \text{ cd/ft}^2 = 1 \text{ cd/m}^2$

defined as luminance times pupillary area, and the units are given as $\text{cd} \cdot \text{m}^{-2} \cdot \text{mm}^2$ where the pupillary area is the area of the iris entrance pupil of the eye [14]. The quantity is only proportional to retinal illuminance and the term is used only when relative amounts of light are important. An older term for this quantity is the *luxon*.

Another interesting unit is the *bougie*. Bougie is the French word for candle. One bougie is a former unit of luminous intensity equal to 0.96 international standard candle [15]. The *international standard candle* is the candela, so 1 bougie = 0.96 cd. The term *millibougie* would refer to 0.00096 cd and a millibougie · cm⁻² would be 0.96 mcd · cm⁻². According to de Brichambaut [16], translated [17] from the French: "In the old French standards, luminance was called *brillance* and was defined by a certain candle (Hefner type) whose apparent area was 2 cm² and the intensity approximately 0.9 times the 'international candle'. Its luminance was therefore $L = 0.9/2 = 0.45 \text{ cd/cm}^2 = 4500 \text{ cd/m}^2$." According to this, the bougie is equivalent to the old Hefner candle. Comparing this passage and the entry for German apostilb in Table 10.4 with the definition in Ref. 15, there seems to be some uncertainty as to whether the conversion factor from bougie to candela is 0.90 or 0.96, possibly due to difficulties in determining the true intensity of the Hefner candle in modern units or to an error in Ref. 15.

In spite of the desire stated several times in this text for uniformity in both terminology and units, it is important to retain some degree of flexibility. For example the term "intensity" is widely used for irradiance. In many of these discussions, other quantities of radiometry or photometry are seldom mentioned. Authors are always free to define their terms and then to use them consistently. A consistent use of the word "intensity" for irradiance, where the definition is clearly stated at the outset, can be tolerated in applications where the use of internationally recognized and standardized terminology is not essential.

Along these lines, the concepts of *luminosity* and *radiosity* can be introduced. These terms appear in specialized areas, frequently with different meanings. Luminosity is by far the most common of the two terms, and was used in the past to designate the human photopic spectral luminous efficiency function [18], $V(\lambda)$. It is also used in interferometry to designate [19] "the ratio of the output flux to the flux per unit area and solid angle of the source lying within the spectral range of width $\delta\sigma$. It is given in terms of the parameters of the instrument by $L = S\Omega\tau$. $S\Omega$ is the product of the area times solid angle taken at the position in the optical system giving the smallest value, and τ is the transmission at the center of the passband." Comparison with (8.15) reveals that this version of luminosity is probably equivalent to throughput or *etendu*.

In astronomy, luminosity is defined [20] as "the rate at which energy of all types is radiated by a star in all directions. A star's luminosity depends on its size and temperature, varying approximately as the square of the radius and T^4 . The sun is a medium-sized star with a luminosity of 3.8×10^{33} ergs/sec . . .".

REFERENCES

- [1] Meyer-Arendt, J. R. "Radiometry and Photometry: Units and Conversion Factors," *Appl. Opt.*, Vol. 7, 1968, pp. 2081-2084.

-
- [2] Biberman, L. M. "Apples, Oranges, and UnLumens," *Appl. Opt.*, Vol. 6, 1967, p. 1127.
 - [3] Illuminating Engineering Society of North America. "Nomenclature and Definitions for Illuminating Engineering," ANSI/IES RP-16-1986, Approved Oct. 18 1985 as a transaction of the IESNA and June 23 1986 by the American National Standards Institute; published by IESNA in 1987.
 - [4] Commission Internationale de l'Eclairage. *International Lighting Vocabulary*, 4th ed., Publication No. 17.4 (joint IEC/CIE publication), 1987.
 - [5] Commission Internationale de l'Eclairage. *International Lighting Vocabulary*, 3rd ed., Publication CIE No. 17 (E-1.1.) 1970.
 - [6] American National Metric Council. "Metric Guide for Educational Materials", ANMC, 1625 Massachusetts Ave., N. W., Washington, DC 20036, 1977.
 - [7] Taylor, B. N. "The International System of Units," NIST Special Publication 330, 1991 ed., NIST, Gaithersburg, MD 20899, Aug. 1991.
 - [8] LI-COR, Inc., "Radiation Measurements," RMR2-1084, Li-Cor, Inc., Box 4425, Lincoln, NB 68504, phone 402/467-3576, 1979.
 - [9] Incoll, L. D., S. P. Long, and M. A. Ashmore. "SI Units in Publications in Plant Science," in *Commentaries in Plant Science*, Vol. 2, Pergamon Press: Oxford, UK, 1981, pp. 83-96.
 - [10] Shibles, R. "Committee Report. Terminology Pertaining to Photosynthesis," *Crop Sci.*, Vol. 16, 1976, pp. 437-439.
 - [11] Thimijan, R. W., and R. D. Heins. "Photometric, Radiometric, and Quantum Light Units of Measure: A Review of Procedures for Interconversion," *Hort. Sci.*, Vol. 18, 1983, pp. 818-822.
 - [12] LI-COR, Inc. "Radiation Measurements and Instrumentation," Publ. 8208-LM, LI-COR, Inc., 4421 Superior Street, Lincoln, NB 68504, Aug. 1982.
 - [13] Kingslake, R. "The Photometry of Optical Systems," unpublished manuscript distributed by the Optical Design Department of Hawk-eye Works of Eastman Kodak Company, Rochester, New York, Sept. 1946, retyped Dec. 1958.
 - [14] Lewis, Alan L. College of Optometry, Ferris State University, Big Rapids, MI, private communication, 1993.
 - [15] *McGraw-Hill Dictionary of Scientific and Technical Terms*, 2nd ed., McGraw-Hill Book Co.: New York, NY, 1978.
 - [16] de Brichambaut, C. P. *Rayonnement Solaire et Echanges Radiatifs Naturels*, Gauthier-Villars: Paris, 1963.
 - [17] Gueymard, C. Florida Solar Energy Center, Cape Canaveral, FL, private communication, 1993.
 - [18] Fry, G. A. "The Eye and Vision," in *Applied Optics and Optical Engineering, Vol. II, The Detection of Light and Infrared Radiation*, ed. by R. Kingslake, Academic Press: New York, NY, 1965.
 - [19] Baird, K. M., and G. R. Haines. "Interferometers," in *Applied Optics and Optical Engineering, Vol. IV, Optical Instruments*, ed. by R. Kingslake, Academic Press: New York, NY, 1967.
 - [20] Harris, W. H., and J. S. Levey, eds. *The New Columbia Encyclopedia*, Columbia University Press: New York, NY, 1975.

Chapter 11

Basic Concepts of Color Science

11.1 INTRODUCTION

The perception of color is a psychophysical process. It involves both the physical processes of presenting a spectral distribution of radiant flux to the eye (including its passage into the eye to the retina), and the psychobiological processes involved in its reception and interpretation by the eye and the brain. The perception of color is a subjective phenomenon that varies somewhat from person to person. Fortunately, however, there is sufficient constancy in color perception across most individuals that it is possible to specify color analytically with reasonably good results for most practical applications.

The physical stimulus that produces color sensation, spectral radiant flux (primarily as spectral radiance entering the eye), is easily described and measured using the methods described in this book. Accurately predicting the human color sensation resulting from the stimulus, however, is not as simple or straightforward. There has been more than a century of investigation into the science of color. Even so, many questions still remain unanswered. This chapter focuses mostly on the photometric and radiometric aspects of color. In preparation, it begins with important terminology and some of the more subjective aspects of the phenomenon.

The colors of monochromatic beams of light at several wavelengths were listed in Section 2.1. They “cover the spectrum” from violet through red. These monochromatic colors are “pure” in the sense that they are composed of only one wavelength. The colors of most light beams and the colors present in light reflected by objects are generally mixtures of many wavelengths. Other color sensations besides those produced by monochromatic beams are therefore both possible and common.

Developing the means of specifying and measuring color, including all the possible mixtures of monochromatic colors, has been understandably difficult. Indeed, the history of color science is filled with a number of milestones and changes in the common practice of color measurement and specification. A brief account is provided in the CIE’s publication on Colorimetry [1], which is an excellent basic reference text covering the core concepts and tabulated data used in this chapter.

A brief survey of current colorimetry methods was provided by Miller and Schneider [2] and by Robertson [3].

Although it is not the purpose here to survey the history of color measurement, a certain amount of historical understanding is needed. Several systems of color specification have been developed at different times in this century. Many of these (in somewhat modified form) are still in use today. Some are numerical in nature while others rely on words to name and describe a perceived color. The choice of a system of color specification depends upon the application and the intended goal of the work.

This chapter seeks to introduce the basic concepts of colorimetry and color science, sufficient in scope to enable the reader to see how these concepts relate to photometry and radiometry. It is hoped that this will enable easier access to more detailed treatments of color science. Comprehensive discussions of the subject can be found in excellent textbooks on colorimetry [4–5] and in a publication of the IESNA on color and illumination [6].

11.2 BASIC CONCEPTS AND DEFINITIONS

When we think of color, it is the color of *objects* that usually comes to mind first. In actuality, it is the color of the *light* from these objects that we see and interpret as color. Thus it is important to distinguish between the color of light and the *apparent* color of objects (see Figure 11.1). The apparent color of an object can be altered by changing the spectral distribution of the light used to illuminate it. For example, illuminating a yellow surface with only monochromatic red light from a He Ne laser will make that object appear red. The apparent color of a reflecting object depends strongly on how it is illuminated. Thus we are led to the idea of defining a standard source for use in determining the inherent color of objects. The perceived color of an object depends upon *both* the spectral distribution of the source used to illuminate it *and* the spectral reflectance of the object in the direction of the observer. In using a standard source spectral distribution and by examining the color of the reflected light, one might hopefully be able to specify something like the inherent color of an object.

In colorimetry, the term *illuminant* is used to indicate the shape of the spectral distribution of flux from a source. The CIE defines it this way: “Radiation with a relative spectral power distribution defined over the wavelength range that influences object colour perception.” It is emphasized that an illuminant is a numerically defined spectral distribution of radiant flux. (It may be spectral irradiance or spectral radiance.) It is not any particular source of light. *Standard* illuminants are important in colorimetry. Descriptions of several are provided in a Section 11.8.

Color means different things to different people. The chemist treats it one way, the optical physicist another, and the psychologist still another. The *Inter-*

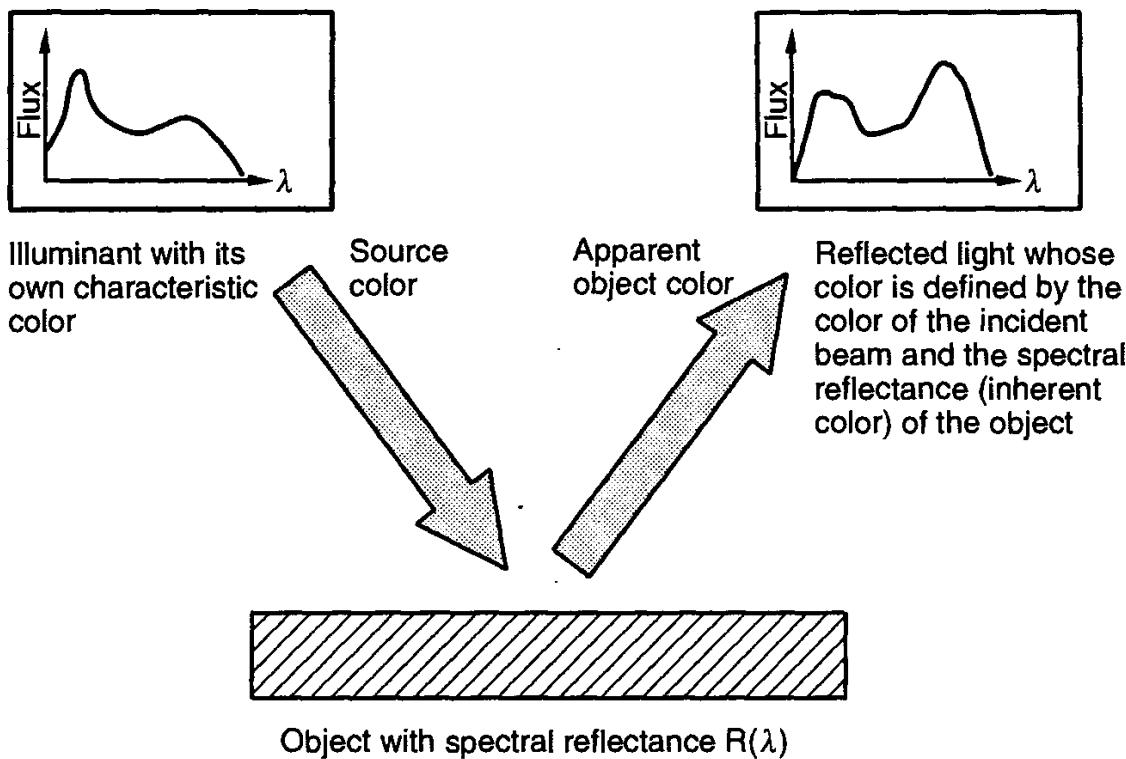


Figure 11.1 Distinguishing between the color stimulus of an illuminant and that reflected from a colored object.

national Lighting Vocabulary of the CIE [7] in its section on light and color offers definitions of the following ten concepts of color:

- (Perceived) color;
- Object color;
- Surface color;
- Aperture color;
- Luminous (perceived) color;
- Nonluminous (perceived) color;
- Related (perceived) color;
- Unrelated (perceived) color;
- Achromatic (perceived) color;
- Chromatic (perceived) color.

The phenomenon of color vision begins when light of various wavelengths is absorbed by the cones in the retina. Upon absorption of the spectral radiant flux incident upon them, the cones produce nerve impulses that are passed on to later stages of the human visual system. The process is described briefly in Section 2.1. There are three different types of cone receptors in the retina, each with its own different spectral sensitivity, shown approximately in Figure 11.2. Since these spec-

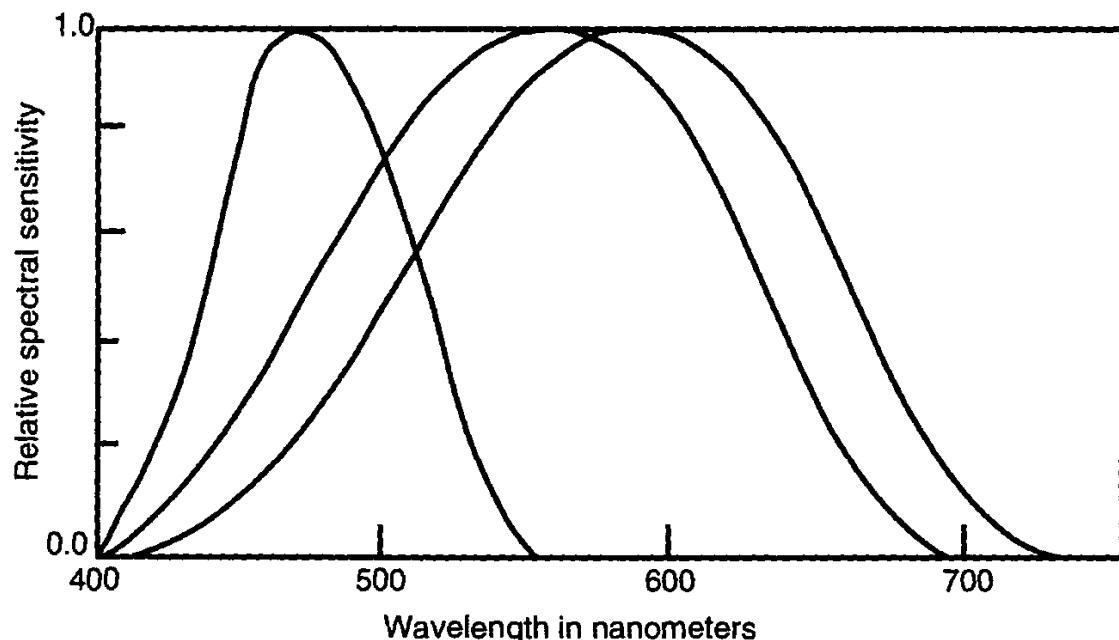


Figure 11.2 Relative spectral sensitivities of the three types of cone receptors in the human eye. These curves are only approximate [6].

tral responses are so broad and overlap so extensively, the curves shown in Figure 11.2 are not very useful in developing a system of color specification. The perception of color involves a complicated neural processing of the signals from all three of these receptors.

It has been found experimentally that if one devises three different *primary sources* of light having carefully chosen spectral flux (radiance or irradiance) distributions, and if these three sources are mixed together in carefully chosen proportions, the perceived color of almost any other light can be matched [2, 4, 5] uniquely. Balancing these sources carefully, the color white can be produced. Many different primary source spectral distributions can be found that are capable of producing a wide range of perceived colors (including white) through proper mixing. Indeed, even monochromatic sources, such as lasers, can produce a wide range of colors, if their wavelengths are chosen properly. The degree of coverage of all perceivable colors depends upon the choice of the primary light sources. Some source spectral distributions can be mixed to form nearly any color while others provide a very limited range of colors that can be generated. Although any of the primary sources could be considered an illuminant, they are not used this way. An illuminant generally refers to the spectral distribution of flux (usually white or nearly white in color) used to illuminate an object.

We are led to the concept of having a system of three standard color spectral distributions, (each with a different perceived color), with which it is possible to specify any arbitrary color as some combination of the three standard source spec-

tra. Such a system is referred to as a *trichromatic system*. It is based upon matching an arbitrary color by an additive mixture of the suitably chosen reference color spectral distributions. Trichromatic systems have found wide use in scientific and engineering applications of color science.

For example, let $\langle D \rangle$, $\langle E \rangle$, and $\langle F \rangle$ represent three different spectral distributions forming the basis of a trichromatic system. These are called the *primaries* or *primary stimuli* of the system. In choosing this basis set of spectral distributions, which can be thought of as an orthogonal set of unit vectors in a three-dimensional vector space, it is important that they be mutually independent; that is, that none of the three can be matched in color by any mixture of the other two [8]. Accordingly, in principle, any arbitrary color $\langle C \rangle$ occupying a point in the three-dimensional space, can be specified by the sum

$$\langle C \rangle = D\langle D \rangle + E\langle E \rangle + F\langle F \rangle \quad (11.1)$$

for some values of the coefficients D , E , and F , which play the roles of the coordinates of the point in the three-dimensional vector space. This approach to color specification has been in use for many years. In the early days, the color-matching primaries $\langle D \rangle$, $\langle E \rangle$, and $\langle F \rangle$ were chosen for experimental convenience to be monochromatic sources, usually centered near to 700 nm, 546 nm, and 436 nm [8]. Having monochromatic primaries would be wonderful from a mathematical perspective. However, it has been found that such primaries simply cannot fully cover the range of colors the eye and brain can perceive. Accordingly, the use of monochromatic primaries has been abandoned in favor of broad spectral distributions. Letting $\langle R \rangle$, $\langle G \rangle$, and $\langle B \rangle$ symbolize the (unit vector) spectral distributions serving as the primaries of a red-green-blue color system, one can rewrite (11.1) as

$$\langle C \rangle = R\langle R \rangle + G\langle G \rangle + B\langle B \rangle \quad (11.2)$$

Instead of using the terms “source spectral distribution,” “spectral distribution of radiant flux,” or “color spectral distribution,” the term *color stimulus* or just *stimulus* is used. It is defined as “visible radiation entering the eye and producing a sensation of color, either chromatic or achromatic.” (The term *achromatic* means “without color.”) An achromatic lens, for example, is one devoid of chromatic aberration.) The CIE prefers to call a spectral radiant distribution by the name *color stimulus function* ($\phi_\lambda(\lambda)$), the spectral concentration of a radiometric quantity, such as radiance or radiant power, as a function of wavelength [7].

It is noted in the CIE Vocabulary [7] that color stimuli are either real color stimuli or theoretical stimuli, which are defined by linear combinations of real color stimuli. The magnitude of each of the three primary color stimuli used to produce a color $\langle C \rangle$ is expressed in terms of $\phi_\lambda(\lambda)$.

more commonly by specifying the ratios of their magnitudes. The definition of the *tristimulus values* of a color stimulus then follows: "amounts of the three primary color stimuli, in a given trichromatic system, required to match the color of the stimulus considered." The tristimulus values are the components D , E , or F , (or alternatively R , G , and B) in the above equations.

An eleventh definition for color is now introduced. It comes from the *CIE Lighting Vocabulary* [7] section on colorimetry: "A specification of a colour stimulus in terms of operationally defined values, such as three tristimulus values." For precise specification of color, we need a system that permits quantification as well as easy and precise reproduction of colors from a set of color coordinates, such as the three tristimulus values.

At this point it will be convenient to introduce the definitions of some additional terms often found in discussions of color and colorimetry. The term *color stimulus* includes the various gradations from black to white, which are called *achromatic stimuli*, i.e., stimuli devoid of any perception of color. Achromatic sensations are included because they play an important role in characterizing colors, since "pure" or monochromatic colors can be mixed with achromatic ones to produce different color sensations.

There is a very interesting characteristic of human color vision, called *metamerism*. The eye can perceive as identical a large number of different spectral distributions incident on it. *Metameric color stimuli* (also called *metamers*) are spectrally different color stimuli that have the same tristimulus values and therefore appear the same to the human eye. Reference [6] shows three example metamers. One is monochromatic radiation in the yellow portion of the spectrum. Second are two monochromatic stimuli at about 540 and 640 nm in wavelength together stimulating a perception of the same yellow color. Last is a source of constant spectral flux from about 540 to 700 nm. All three of these dramatically different sources appear identically yellow to the normal human observer. Reference [2] offers a more subtle example of two metamers, differently shaped paint spectral reflectance distributions that produce the same apparent color when illuminated by average daylight.

There is another characteristic of human color vision that must be noted. It is called *chromatic adaptation* and it results in shifts in the perceived colors of sources over time in response to the changing average color of other sources over the field of view. According to the IES [6]: "The human visual system adapts rapidly and strongly to large changes of average stimulus, for example, from incandescent light to average daylight. If the eyes are adapted to daylight, an incandescent light appears to be yellow; however, both incandescent and daylight appear white when the eyes fully adapt."

The effects of metamerism, chromatic adaptation, and the influence of the illuminant on the perceived color of objects are the main reasons that the science

of color perception has been so difficult to develop and remains in some state of flux.

11.3 SYSTEMS OF COLOR SPECIFICATION

From human experience it is found that we can describe the appearance of a color stimulus by reference to three characteristics or attributes, *hue*, *saturation*, and *brightness*. Hue is the closeness of a color to one or two of the perceived colors purple (P), blue (B), green (G), yellow (Y), or red (R). Saturation refers to the relative mixture of a chromatic color with white (which is called an achromatic color, in spite of this sounding like an oxymoron, or contradiction in terms). Pink, for example, is a version of red having less saturation than red. Brightness is related to the apparent luminance or flux in the stimulus.

Over the years there has developed a number of different languages for talking about color. This led the Intersociety Color Council to standardize a system of color specification and language for talking about color [9]. The system is used widely around the world by lay persons, artists, printers, manufacturers, and color professionals. It is based on the Munsell Color System [10], which specifies colors by reference to three characteristics of color called *hue*, *chroma*, and *value*, which are the same as hue, saturation, and brightness, respectively. Munsell was an artist. His system of color was originally developed for artistic purposes around the turn of the century.

The Munsell Color System is described in the next section. This is followed by a short description of what is called the $L^*a^*b^*$ system. It is based on the Munsell system, but assigns the colors of the Munsell system to a three-dimensional coordinate system. Next comes a short history of tristimulus colorimetry. Tristimulus colorimetry is based on the approach to colorimetry embodied in (11.1) and (11.2). Sections 11.4 through 11.6 describe the CIE tristimulus systems of color specification in some detail. Readers most interested in the *science* of color as it relates to radiometry and photometry may wish to skip ahead to Section 11.4.

11.3.1 Munsell Color System

In the Munsell system, the monochromatic colors are arranged in a circle with the ends of the spectrum, red and purple, being joined to complete the circle. The geometrical arrangement of the Munsell scale is illustrated in Figure 11.3. *Hue* was described previously. *Chroma* in the Munsell system refers to what was previously called saturation. The most saturated colors (those with high chroma) lie along the circumference of a disk bounded by the hue circle. The least saturated color, grey, lies at the center of the disk. Above and below this central disk, with the most

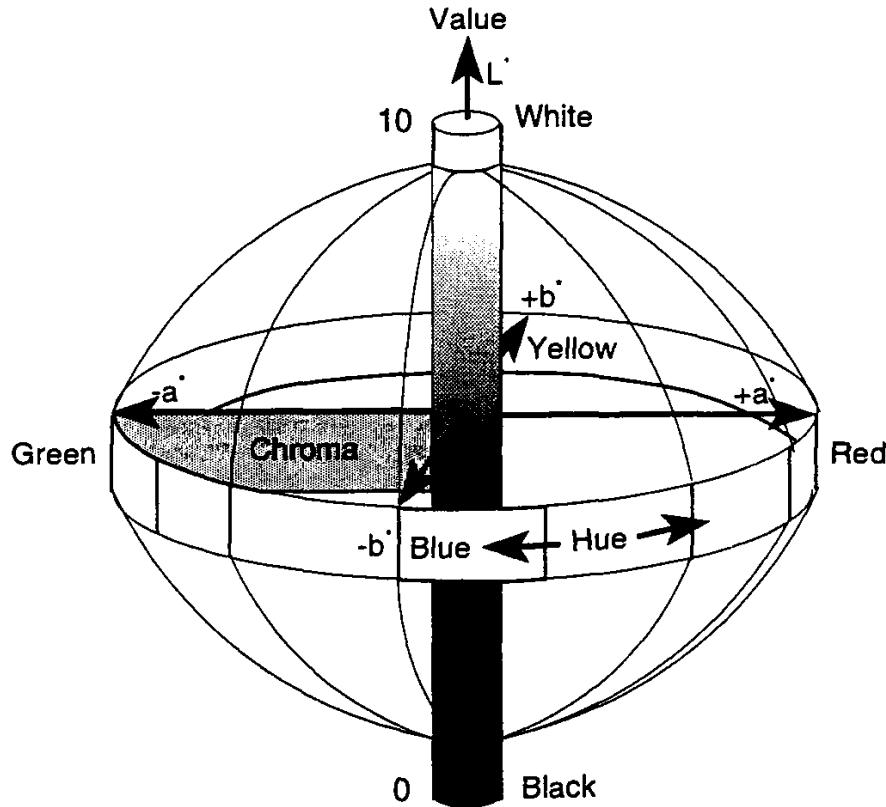


Figure 11.3 Illustration of the Munsell hue, chroma, and value color designation scale adopted by the Intersociety Color Council. (The L^* , a^* , and b^* symbols refer to the CIE 1976 CIELAB color space described in Section 11.3.2.)

saturated colors at its edge, lie additional disks. The ones above the central disk have colors with increasing brightness or, in Munsell terms, *value*. Below this disk, the colors become darker. Fewer colors can be distinguished as they become lighter and darker. Thus, the sizes of the disk decrease upward and downward, until they reach one point at the top, which represents white, and another at the bottom, which represents black.

Numbers have been assigned to the angular positions around the Munsell disks, to the radial distances outward, and to the distance from 0 for black at the bottom to 10 for white at the top. The five basic hues are given letter designations, as are the hues lying in between them, resulting in the sequence

Y GY G BG B PB P RP R YR Y.

Each of these is assigned a value of 5. Numbers from 1 to 10 are assigned to the hue of a given stimulus according to whether the hue lies along a scale from half

way toward the hue to the left (1), to half way toward the hue to the right (10) in the above sequence. For example, hues from B to PB are designated as follows:

B										PB
5	6	7	8	9	10	1	2	3	4	5

A letter and number designation scheme indicates the coordinates on this three-dimensional representation of color in the Munsell system. The format for stating the hue, value, and chroma is $H\ V/C$. For example, 5R 4/14 designates a red hue with a lightness or value of 4 and a strong saturation or chroma of 14. The practical maximum values for chroma are not constant and have been increasing for some hues more than others. The result is a somewhat distorted boundary to the Munsell color space, as indicated in Figure 11.4. The “object” illustrated in Figure 11.4 is referred to as the *color solid*. The *Munsell Book of Color* [11] contains about 1,600 reflective color samples approximately equally spaced in the three attribute scales.

The idea of the equal spacing of colors leads to the concept of color space uniformity. This is especially important in practical applications where the comparison of two or more colors is important.

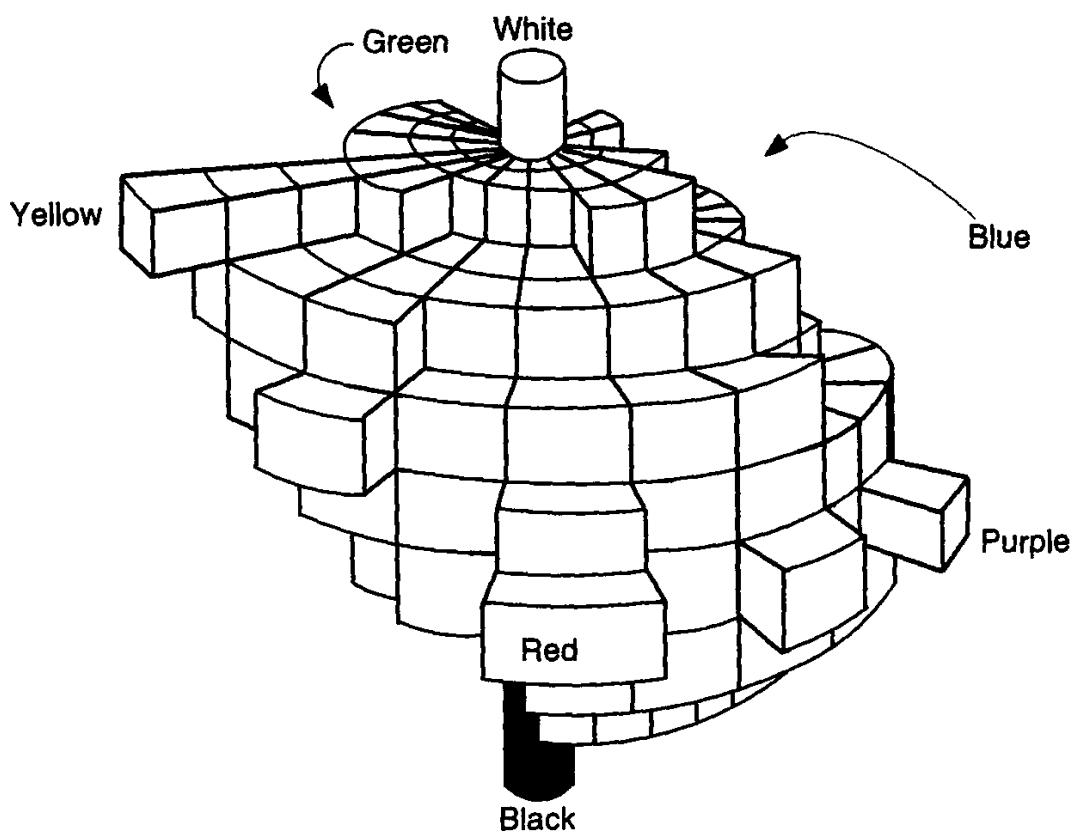


Figure 11.4 Schematic representation of the Munsell color solid.

11.3.2 CIE 1976 ($L^*a^*b^*$) Color Space

Another system of color specification has been developed. Illustrated in Figure 11.3, it is based upon the hue, chroma, and value coordinates of the Munsell color system. The main use of this system today is for improved specification of color *differences*. Called the CIE 1976 $L^*a^*b^*$ system, it provides a three-dimensional color space in which the numerical positions of colors more closely match their perceived relative spacings. In this system, a rectangular coordinate system is placed on the central color disk, the one for zero value, where the hues at the edges of this disk are the most saturated. The positive x axis extends out from the center toward its maximum value of +60 at red. This axis is named the $+a^*$ axis. Perpendicular to this, and going toward the yellow hue is the positive y axis, which is given the name $+b^*$ in this system. The $-a^*$ axis extends toward green and the $-b^*$ towards blue. The value axis, which is the z axis, extends upward, coincident with the value or lightness direction of the Munsell system. The value (or “lightness”) axis is given the name L^* in the resulting CIE 1976 ($L^*a^*b^*$) space, or CIELAB. The CIELAB color space is encountered again in Section 11.6, where its connection to the CIE tristimulus system is described.

11.3.3 Tristimulus Colorimetry

For the tristimulus colorimetry system, an assumption is made that any color can be uniquely matched by a suitable additive combination of three primary stimuli, (generally chosen to be red, green, and blue). Other primary colors can be selected, but with varying abilities to match the wide range of possible colors we see. The most important tristimulus systems of colorimetry are those developed over the last several decades by the CIE. The reasons for this preeminence of the CIE systems are 1. that the systems are readily applicable for engineering applications due to their numerical nature and analyses and 2. they are nearly universally accepted around the world. The latter is explicitly mentioned in the preface to the 1986 CIE publication on colorimetry [1] as follows: “By general consent in all countries the specification of basic standards for use in colorimetry is the province of the Commission Internationale de l’Eclairage (CIE).” Further excerpts from this preface offer a succinct look at the historical development of the tristimulus color systems developed by the CIE and described in the next sections:

The first major recommendations regarding colorimetric standards were made by the CIE in 1931, and these formed the basis of modern colorimetry. The original recommendations of 1931 were reviewed from time to time by the CIE Colorimetry Committee and changes were made when these were considered necessary. New recommendations were

added to supplement the existing ones or to broaden the scope of colorimetry in accordance with developments in practice and science.

The deliberations and recommendations made by the CIE Colorimetry Committee are recorded in the Proceedings . . . Unfortunately the distribution of these has always been rather limited and ready access to them often proves difficult. In addition, much of the material published in the Proceedings is obsolete or inconsistent with current colorimetric practice. . . .

In view of these circumstances the CIE published in 1971 a special Document on Colorimetry to provide a consistent and comprehensive account of basic colorimetric recommendations according to the CIE. . . . The document was issued as CIE Publication No. 15.

Since 1971 it has been necessary to add two supplements to the document [on metamerism and on uniform color spaces and color difference equations]. . . . Agreement reached by the committee in 1982 on a formula for evaluating whiteness would have called for a third supplement, but it was decided that the time had come to issue a second edition of Publication No. 15 incorporating all [the] supplementary material and various other changes of a more detailed nature.

The second edition of the special *Document on Colorimetry* was published in 1986. For more information about the historical development of CIE color systems, the reader is referred to the introduction in the 1986 edition of Publication 15.2. The 1931 system has survived the test of time and is still used widely for color specification. Later modifications of this system by the CIE were made to deal with specific problems that arose subsequent to the system's adoption in 1931. In particular, later experiments showed that the 1931 system was severely lacking in specifying color differences. In consequence, a 1960 uniform color space diagram was developed, employing a linear transformation of the 1931 tristimulus values. This diagram was constructed so that isotherm lines are perpendicular to the line drawn on the curve of the colors of Planckian blackbody radiators over a range of temperatures. This arrangement greatly enhances one's ability to determine the correlated color temperature of a given color. Color temperature is discussed in Sections 3.3 and 11.7. The 1960 uniform color space was superseded by a 1976 one described in Section 11.6, but the use of the 1960 system for the determination of correlated color temperature is retained.

Since both the 1931 and 1960 systems assumed brightness (value) to be held constant, there are a number of colors in the Munsell three-dimensional solid that cannot be properly represented by either of these two-dimensional systems. The color brown, being a version of yellow with a low value parameter, is one example.

At the 18th session of the CIE in 1975, the Colorimetry Committee approved the adoption of two new color spaces and associated color difference formulae. They are known as the CIE 1976 L^{*}u^{*}v^{*} color space (CIELUV) and the CIE 1976 L^{*}a^{*}b^{*} (CIELAB) space. Neither of these two new spaces and associated formulae represents average visual color difference judgments entirely satisfactorily, and work continues to find a better solution to the problems. The relationships among the CIELUV, CIELAB, CIE 1931, and 1964 tristimulus systems are described in Section 11.6.

11.4 CIE 1931 COLOR SYSTEM

The tristimulus system adopted by the CIE is based upon three primary color stimuli, corresponding roughly to the colors red, green, and blue. They are based on many observations with visual colorimeters, enclosures in which a white screen is half-illuminated by an unknown source of colored light and half by light that is a known mixture of red, green, and blue light. The observer adjusts the mixture of the known colors until the colors on the two halves of the screen appear to be identical. According to the IESNA [6]: "By a proper choice of colored lights a wide gamut of unknown lights can be matched. Red, green, and blue primaries provide a wide gamut, but no set of primaries can provide a match for all colors. For some colors, it may be necessary to add one of the primary lights to the unknown light. When this is achieved, it is possible to match all colors."

The *primary color stimuli* chosen by the CIE are three spectral distributions of radiant flux defining what is called the *standard observer*. They are given the symbols (X), (Y), and (Z) and were chosen for reasons of convenience in colorimetric computation. For the specification of color in the CIE system, three functions, $\bar{x}(\lambda)$, $\bar{y}(\lambda)$, and $\bar{z}(\lambda)$ of wavelength λ are defined. They are related to the primary color stimuli and are called *color matching functions*. They are the amounts of each of the primary stimuli (X), (Y), and (Z), mixed together, that are needed for a human observer to perceive the mixture as having the same color as that of monochromatic flux at the wavelength λ , for any wavelength over the visible spectrum [1, 2]. These CIE 1931 color-matching functions are plotted in Figure 11.5.

When the CIE adopted the standard observer, the original colorimeter data were transformed 1) so as to eliminate some negative coefficients present in a previous system, 2) so that the $\bar{y}(\lambda)$ function matches the human photopic spectral luminous efficiency function $V(\lambda)$ given in Table 2.1, and 3) to make the $\bar{x}(\lambda)$ and $\bar{z}(\lambda)$ color matching functions have zero brightness (or value) [2, 4, 5]. More recently, the color-matching functions were changed slightly at three wavelengths from the values defined in 1931. The changes are quite small, but noticeable in one case: $\bar{y}(\lambda)$ at 555 nm, the wavelength of the peak of this function, had its value reduced from the 1.0002 value previously found in most references for V-lambda

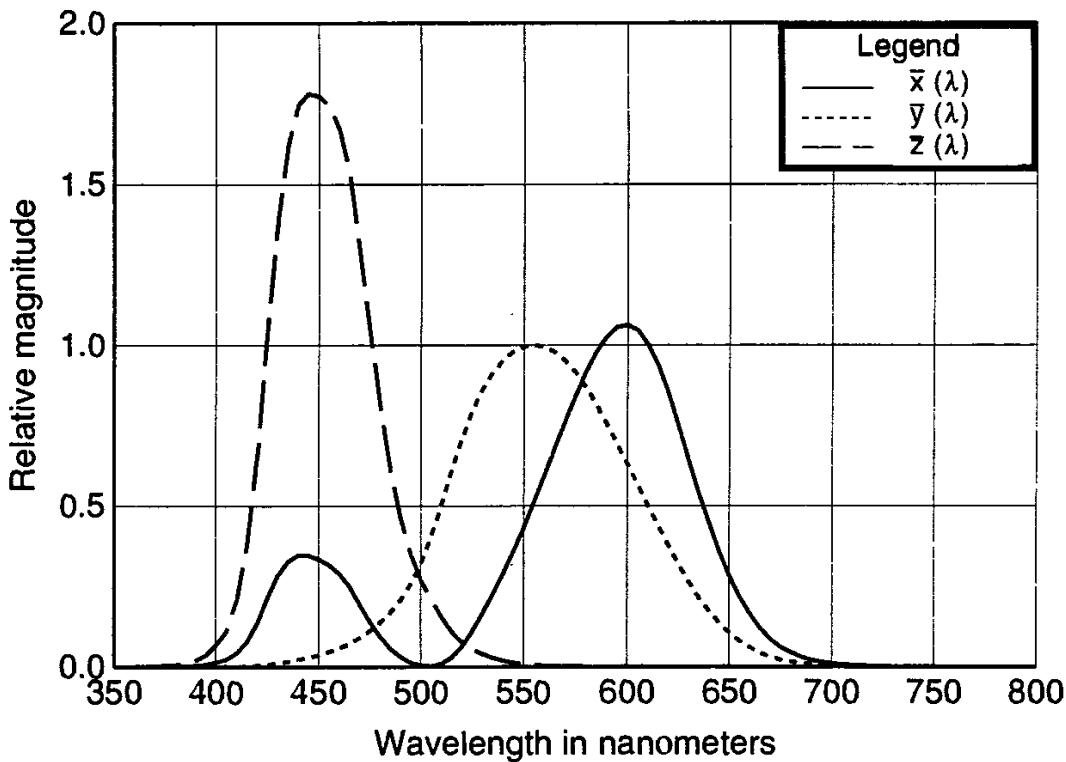


Figure 11.5 Plots of the color-matching functions of the CIE 1931 Colorimetric System. $\bar{x}(\lambda)$ = red, $\bar{y}(\lambda)$ = green, and $\bar{z}(\lambda)$ = blue.

at 555 nm to 1.0000. These color matching functions are used for the specification of color instead of the primary color stimuli for convenience. In 1990 the V-lambda curve was modified slightly, into a new supplementary one, designated by the symbol $V_M(\lambda)$ and named the CIE 1988 Modified 2° Spectral Luminous Efficiency Function for Photopic Vision. It is a *supplement* to, not a replacement of the 1924 function [12].

Tabulated values for the CIE 1931 color-matching functions are printed in [1], in most textbooks on colorimetry, and in ASTM standard E 308-90 [13–14]. In accordance with an agreement between the CIE and the International Standards Organization (ISO) in 1991, most CIE standards will be published as joint double-logo standards by ISO and the CIE. CIE Publication S002 was republished with some modifications and error corrections as “International Standard ISO/CIE 10527: CIE Standard Colorimetric Observers.” This document presents the definitions and specifications of the color matching functions, describes the derivation of the data of the “standard colorimetric observers” and discusses practical application of the color matching functions. The standard colorimetric observers are defined by their tabulated values between 360 and 830 nm in 1-nm intervals to seven significant digits. Standard ISO/CIE 10527 is available from the CIE National Committees or from the CIE Central Bureau in Vienna. Computer disks D001 and

D002 containing data for both 1931 and 1964 standard colorimetric observers as well as spectral power distribution data described in Section 11.8, are also available.

The color matching functions, $\bar{x}(\lambda)$, $\bar{y}(\lambda)$, and $\bar{z}(\lambda)$, are used as weighting functions for a given spectral radiant flux distribution. A given spectral flux distribution (which may be spectral radiant flux, spectral irradiance, or spectral radiance) is also called a color stimulus, designated here by $Q(\lambda)$. By integrating $Q(\lambda)$ over the visible spectrum three times, with each of the spectral weighting functions $\bar{x}(\lambda)$, $\bar{y}(\lambda)$, and $\bar{z}(\lambda)$, one obtains three color coordinates X , Y , and Z that uniquely characterize the color of the stimulus spectrum $Q(\lambda)$:

$$\begin{aligned} X &= k \int Q(\lambda) \bar{x}(\lambda) d\lambda \\ Y &= k \int Q(\lambda) \bar{y}(\lambda) d\lambda \\ Z &= k \int Q(\lambda) \bar{z}(\lambda) d\lambda \end{aligned} \quad (11.3)$$

The three quantities X , Y , and Z are known as *tristimulus values* for the source spectrum $Q(\lambda)$. They uniquely characterize the color of this spectral distribution of light. As is done with many practical measurement systems, for real colorimetric measurements, the integrals in (11.3) are replaced with sums over equal wavelength intervals. (See the Appendix for more information about this process.)

To simplify analysis of colors with this system, a simple transformation is applied to normalize the above quantities. X , Y , and Z , are replaced by x , y , and z , respectively, as follows:

$$\begin{aligned} x &= \frac{X}{X + Y + Z} \\ y &= \frac{Y}{X + Y + Z} \\ z &= \frac{Z}{X + Y + Z} \end{aligned} \quad (11.4)$$

The resulting transformed values, x , y , and z , are called *chromaticity coordinates* since they are used to locate points in a three-dimensional color space. As a result of this normalization $x + y + z = 1$ and what was a three-dimensional color space has been transformed into a two-dimensional one. Since $z = 1 - x - y$, only x and y have to be specified to designate any color in this system uniquely. Note also that the constant k is eliminated in (11.4). See Figure 11.6 for a depiction of the CIE 1931 two-dimensional tristimulus color space. Sometimes a point in this color space is called a *color*. It can also be called a *chromaticity*.

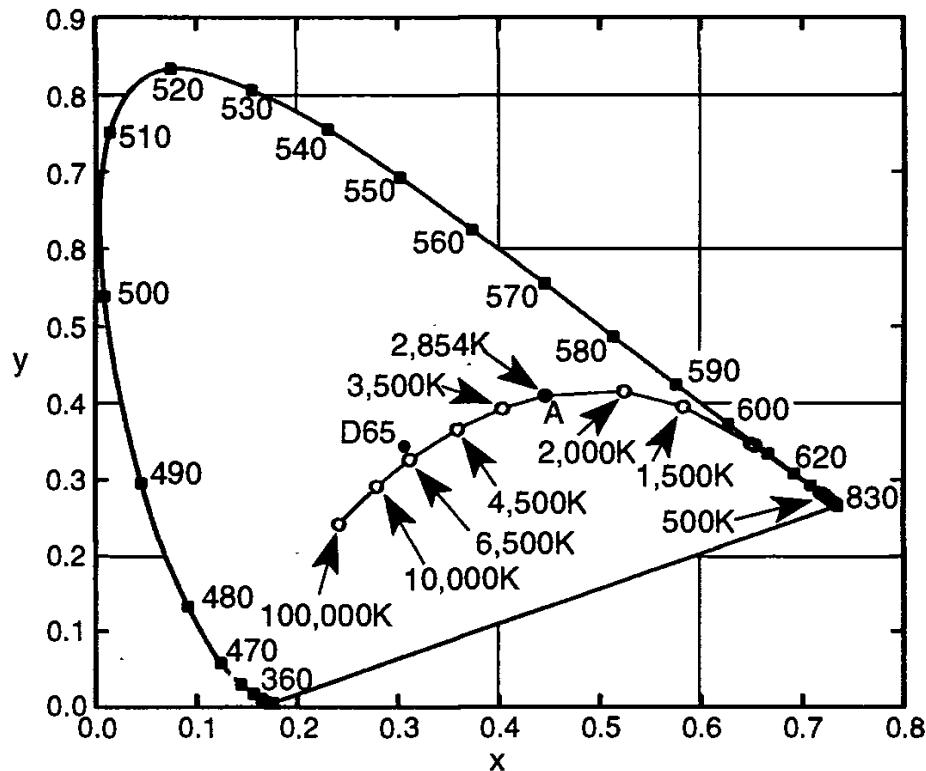


Figure 11.6. Plot of the CIE 1931 color space, with the colors of blackbody radiators at temperatures ranging from 500 to 100,000K. Monochromatic wavelengths along the spectrum locus are given in nanometers. A and D65 refer to CIE standard illuminants.

With this system of color specification, the monochromatic colors lie on a locus surrounding and bounding an area of an x - y plot of color points. A monochromatic color can be represented by letting the stimulus $Q(\lambda)$ in (11.3) be 1.0 for some wavelength λ and 0.0 for all others. The boundary of this area can be determined by replacing $Q(\lambda)$ in the summation versions of (11.3) with the number 1.0 for some wavelength λ (and zero for all others) and using (11.3) and (11.4) to determine the coordinates $(x'(\lambda), y'(\lambda))$ of the corresponding monochromatic color on the x - y plot. The result is

$$\begin{aligned} x'(\lambda) &= \frac{\bar{x}(\lambda)}{\bar{x}(\lambda) + \bar{y}(\lambda) + \bar{z}(\lambda)} \\ y'(\lambda) &= \frac{\bar{y}(\lambda)}{\bar{x}(\lambda) + \bar{y}(\lambda) + \bar{z}(\lambda)} \end{aligned} \quad (11.5)$$

If this calculation is repeated for each wavelength from 360 to 830 nm, the plot shown in Figure 11.6 for the boundary of the CIE 1931 tristimulus color space,

called the *spectrum locus*, is obtained. Monochromatic wavelengths from 360 to 830 nm are indicated on the diagram. The color white (neutral or achromatic color) lies in the interior of this space around the point designated by the coordinates $x = 0.333$ and $y = 0.333$. The colors (hues) become more saturated (closer to monochromaticity) as one moves away from this color toward the locus of monochromatic colors having maximum saturation. The colors indicated by "A" and "D65" are of standard illuminants, explained in Section 11.8.

Chapter 3 contains a discussion of blackbody radiation. If the Planckian (blackbody) spectral distribution represented by (3.4) is used in (11.3) and (11.4) for each different temperature from 500K to 100,000K, the points indicated by open circles and connected by the solid curve shown in Figure 11.6 are obtained.

If light having the color indicated as "A" on the diagram is mixed with light having another color, such as blackbody radiation at 4,500K, the color of the mixture will lie on a straight line joining these two points. The ends of the spectrum locus at 360 and 830 nm are joined by a straight line because mixtures of these two colors will have chromaticities lying on this line, and colors below this line are not possible [15].

Both the hue and saturation of a color are specified by the x and y chromaticity coordinates of a color. The lightness (value, brightness, or luminosity) is specified by capital Y , given in (11.3). Because of this, the system is sometimes referred to as the CIE Yxy (or at other times as the xyY) color system [16].

Designing a radiance meter to have filter/detector combinations producing an instrumental spectral sensitivity that matches each of the three CIE spectral tristimulus color matching functions (\bar{x} , \bar{y} , and \bar{z}) results in a device called a *colorimeter*, since it can directly measure the tristimulus values X , Y , and Z of the radiant flux passing through its aperture to its detector(s). Such colorimeters are commercially available. Many of the photometer manufacturers listed in Table 9.1 also market colorimeters.

Since the $\bar{y}(\lambda)$ color matching function is identical to the V-lambda function, most colorimeters are capable of measuring luminance and/or illuminance. In addition, according to Miller and Schneider [2], "It is not practical to build a single filter which can match the entire $\bar{x}(\lambda)$ function from 380 to 760 nm, so the 380 to 500 nm portion is disregarded and compensated in the calibration. However, on some high-performance photometer/colorimeters, a fourth filter is added which matches the 380 to 500 nm portion of the $\bar{x}(\lambda)$ function. This fourth filter, which is designated the X_b or X' filter, is added because it results in a slight improvement in measuring accuracy, especially when measuring sources with much blue content. . . . Obviously, the more closely the colorimeter spectral 'trim' matches the CIE color matching function, the better the colorimeter's accuracy when measuring sources with different spectral energy distributions. When proper calibrating sources are available, photometers with all four tristimulus filters have proven to be accurate to within ± 0.003 to ± 0.005 chromaticity units." The "chromaticity

units" quoted here are values of the chromaticity coordinates in the CIE 1931 color system.

11.5 CIE 1964 SUPPLEMENTARY OBSERVER COLOR SYSTEM

The CIE 1931 standard colorimetric observer was based on work performed by Guild [17] and Wright [18]. The colorimetric system based on this observer is still in use for general colorimetry. Subsequent to the 1931 adoption of this standard, however, further work over a period of years, detailed in [1], indicated that the data used may not adequately represent the color matching properties of the average observer with normal color vision and that the color matching functions were too low in the region from 380 to 460 nm. The 1931 system was based on observers looking at small 2° angular fields of view of relatively low luminance. When the field of view increases, it was later discovered, the results are slightly different. These problems were reviewed and studied for a number of years, resulting in a decision in 1964 to retain the CIE 1931 standard observer data for general colorimetry and to *supplement* it with new color matching functions for use whenever a more accurate correlation with visual color matching of fields of large angular subtense (more than 4 deg) is desired.

New color matching functions were specified, named $\bar{x}_{10}(\lambda)$, $\bar{y}_{10}(\lambda)$, and $\bar{z}_{10}(\lambda)$, with the subscript added to indicate that these functions are applicable for a 10-deg field rather than the 2-deg field of the 1931 system. The new functions were obtained by a direct method that did not involve an appeal to the CIE spectral luminous efficiency function, $V(\lambda)$. The latter had been identified previously as part of the problem with the 1931 system. The new system depended on actual measurements of the relative flux distribution of the spectrum studied.

Another aspect of the new system is that it is intended to be used at luminance levels high enough to exclude participation by the rod receptors in the eye from the determination of color. This observation condition is important because "rod intrusion" can upset the predictions of the standard observer system when light levels are too low.

The new 1964 10-deg color matching functions, similar in appearance to the 1931 ones, are shown in Figure 11.7, and the corresponding CIE 1964 color space is plotted in Figure 11.8 [15].

11.6 CIE 1976 UNIFORM COLOR SPACE

In spite of the improvements provided by the 1964 color space for certain applications, there have been persistent problems with both the 1931 and the 1964 color spaces. They produce a nonuniform distribution of colors, and anomalies whenever certain color difference calculations are performed. This has been a vexing problem

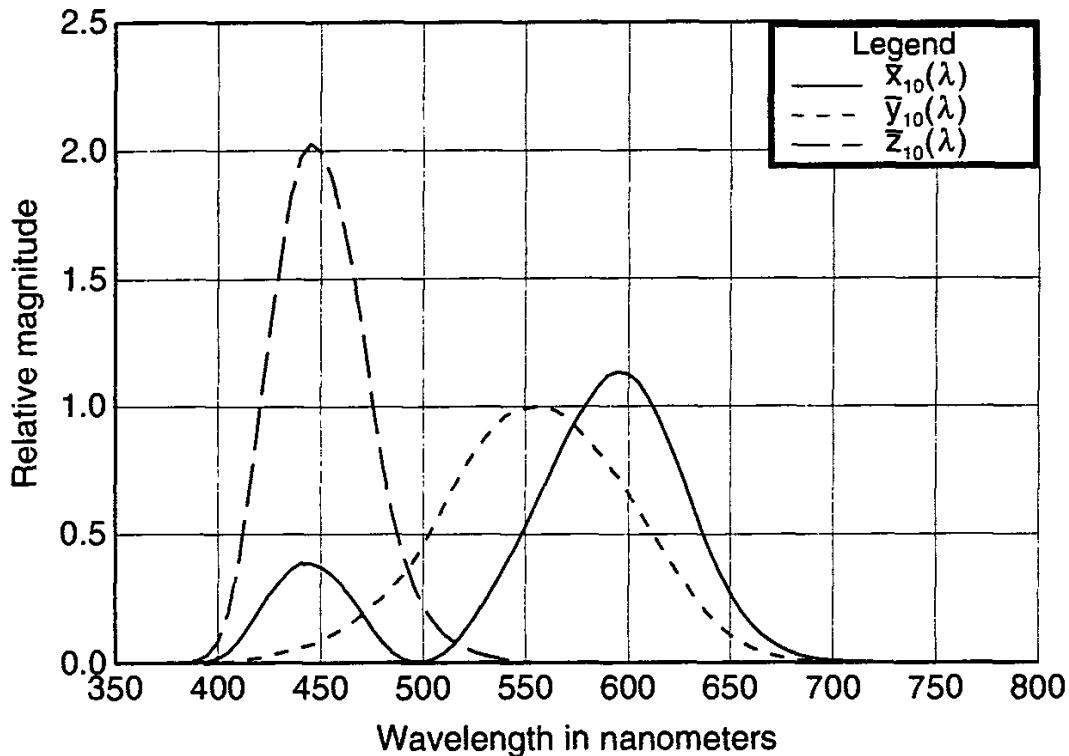


Figure 11.7 CIE 1964 tristimulus color matching function spectra, for a 10-deg field and high luminance levels where rod vision contributes little to color perceived predominantly by cone vision.

for the CIE Colorimetry Committee for years [1]. In an attempt to at least partially overcome the problems, a uniform color spacing (UCS) scheme was developed and adopted in 1976. It is based on the following straightforward transformation of either the 1931 or 1964 chromaticity coordinates (x, y, z) to the UCS coordinates (u', v', w'):

$$\begin{aligned} u' &= \frac{4x}{-2x + 12y + 3} \\ v' &= \frac{9y}{-2x + 12y + 3} \\ w' &= 1 - u' - v' \end{aligned} \tag{11.6}$$

Whether the 1931 or 1964 chromaticity coordinates are used depends upon the angular subtense of the color specimens being compared. If this angle is from 1 to 4 degrees, then the 1931 values of x and y should be used to calculate u' and v' . Otherwise, the 1964 values should be used and the results should be called u'_{10} and

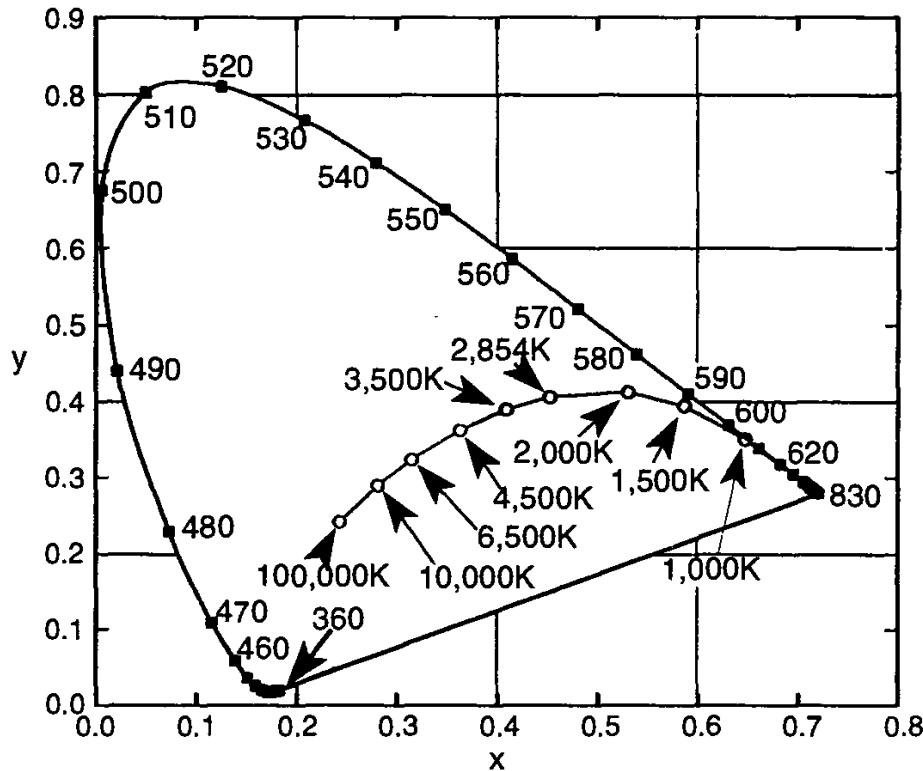


Figure 11.8 CIE 1964 color space, with plots of the colors of blackbody radiators at temperatures ranging from 500 to 100,000K. Monochromatc wavelengths are shown in nanometers.

v'_{10} . (The primes ('') are used on u , v , and w in the formulas above to distinguish them from the now obsolete unprimed CIE 1960 (u , v , w) system.)

The color spacing produced by the above transformations is known to be perceptually more nearly uniform than that of the CIE (x , y) diagrams illustrated in Figures 11.6 and 11.8. The resulting color spaces are plotted in Figures 11.9 and 11.10, one using the CIE 1931 colorimetric observer and the second using the CIE 1964 one. The new UCS color spaces are intended to apply to comparisons of differences between object colors of the same size and shape, viewed in identical white to middle-grey surroundings by an observer photopically adapted to a field of chromaticity not too different from that of average daylight [1].

As an outgrowth of the CIE UCS work, two *three-dimensional* spaces have been defined, primarily to standardize color difference measurement practices in the field. They are called the CIE 1976 ($L^*u^*v^*$) color space and the CIE 1976 ($L^*a^*b^*$) color space. They were introduced, nonmathematically, in Sections 11.3 and 11.3.2. Specific mathematical relationships have been developed to connect the *hue*, *chroma*, and *value* coordinates of the Munsell color system described in Section 11.3.2 with the (x , y , z) system described in Sections 11.4 and 11.5.

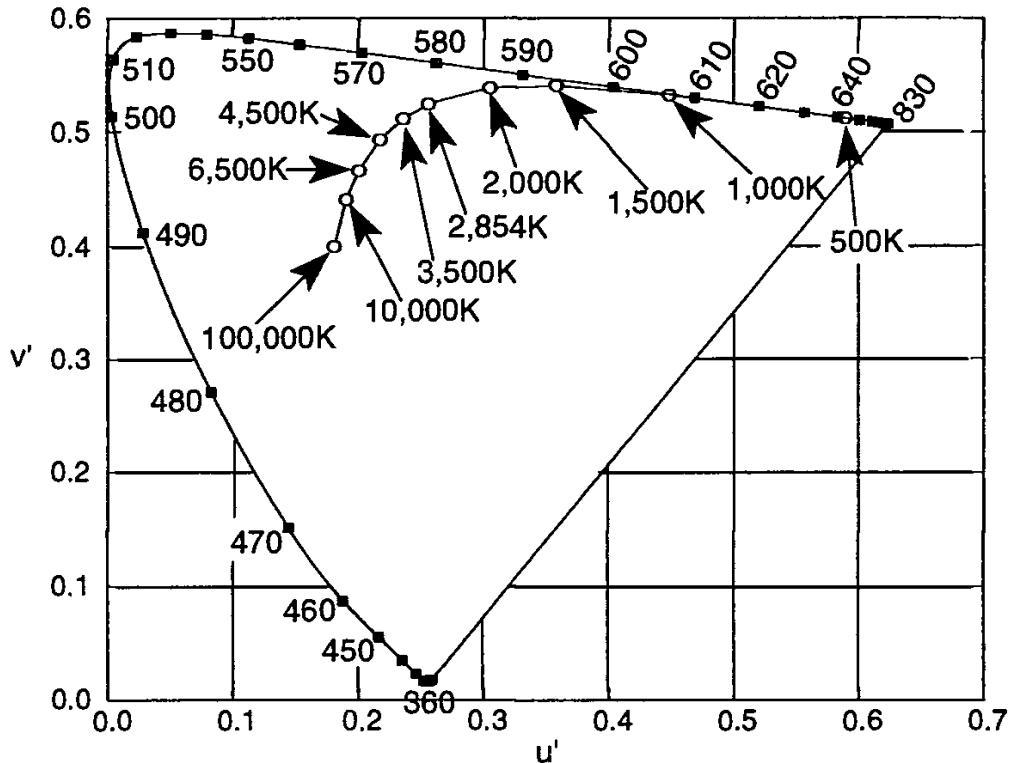


Figure 11.9 CIE 1976 uniform color space for CIE 1931 observer with 2-deg field.

To determine the new coordinates, one measures or otherwise determines the CIE 1931 or CIE 1964 tristimulus values (X , Y , Z) for the radiation whose color is to be determined. See Eq. 11.3. Then one uses equations (11.6) to calculate u' and v' for this stimulus. The coordinates Y_n , u'_n , and v'_n of a specified white color stimulus must also be known. With this information the new uniform color space coordinates can be determined from

$$\begin{aligned} L^* &= 116 \left(\frac{Y}{Y_n} \right)^{1/3} - 16, & \frac{Y}{Y_n} > 0.008856 \\ u^* &= 13 L^* (u' - u'_n) \\ v^* &= 13 L^* (v' - v'_n) \end{aligned} \quad (11.7)$$

The difference ΔE_{uv}^* between two color stimuli is calculated as the Euclidean distance between the points representing them in the CIE 1976 ($L^*u^*v^*$) color space (known as CIELUV 1976):

$$\Delta E_{uv}^* = \sqrt{(\Delta L^*)^2 + (\Delta u^*)^2 + (\Delta v^*)^2} \quad (11.8)$$

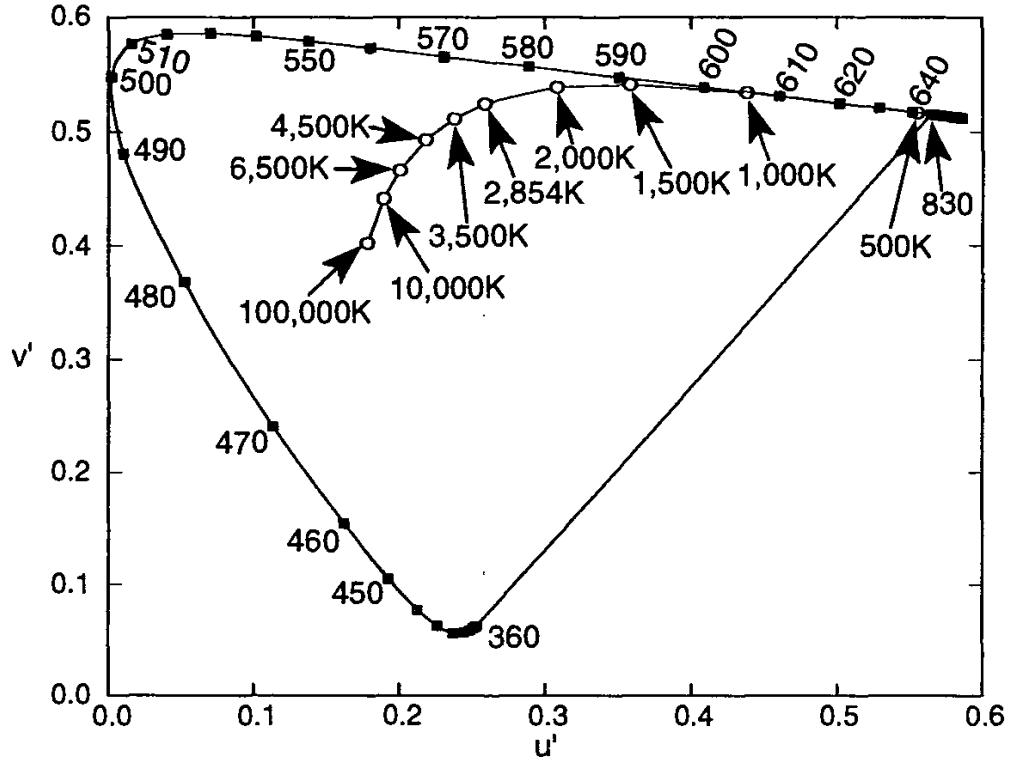


Figure 11.10 CIE 1976 uniform color space for CIE 1964 observer with 10-deg field.

Formulas for the CIE 1976 ($L^*a^*b^*$) color space (CIELAB 1976) are similar to the above, but these are written in terms of the (X, Y, Z) coordinates for the color stimulus directly and (X_n, Y_n, Z_n) for white light:

$$L^* = 116 \left(\frac{Y}{Y_n} \right)^{1/3} - 16 \quad \frac{Y}{Y_n} > 0.008856$$

$$a^* = 500 \left[\left(\frac{X}{X_n} \right)^{1/3} - \left(\frac{Y}{Y_n} \right)^{1/3} - \left(\frac{Z}{Z_n} \right)^{1/3} \right] \quad \frac{X}{X_n} > 0.008856 \quad (11.9)$$

$$\frac{Y}{Y_n} > 0.00856$$

$$b^* = 200 \left[\left(\frac{Y}{Y_n} \right)^{1/3} - \left(\frac{Z}{Z_n} \right)^{1/3} \right] \quad \frac{Z}{Z_n} > 0.00856$$

$$\Delta E_{ab}^* = \sqrt{(\Delta L^*)^2 + (\Delta a^*)^2 + (\Delta b^*)^2} \quad (11.10)$$

The concept of a color difference, the distance between two points on a color space diagram, has been discussed previously. Equations 11.8 and 11.10 provide mathematical quantifications for this concept. For more information on CIELUV and CIELAB color spaces and how they are used, the reader is referred to the technical literature covering the subject. Uniform color spaces such as CIELUV and CIELAB are important for the proper determination of color temperature, as described in the next section.

11.7 COLOR TEMPERATURE

The concept of color temperature was introduced in Section 3.3 on color and distribution temperatures of blackbody radiation. As illustrated in Figures 11.6, 11.8, 11.9, and 11.10, the color of the radiation emitted by a blackbody in the visible portion of the spectrum varies as the temperature is changed (over the range of temperatures where there is significant emission in the visible). *Color temperature* applied to an illuminant is the temperature of a blackbody whose color is the *same* as that of the illuminant. If the color of the illuminant does not quite match that of any blackbody, then a modification of the color temperature, called the *correlated color temperature*, is available. That temperature which produces light from a blackbody having a color *closest* to the color of the source of interest having the same brightness and under the same viewing conditions is called the correlated color temperature of the source [1, 5].

For illuminants with colors quite close to the blackbody radiator colors plotted in Figures 11.6, 11.8, 11.9, and 11.10, it is a simple matter to look at the relevant plot and determine by eye which blackbody color they come closest to. As the colors of illuminants move away from the blackbody locus, however, choosing the right color temperature is not as clear-cut. A strategy is needed for extending the concept of color temperature to these colors. The question is, how do we determine which of all possible colors on a CIE chromaticity diagram are closest to the color of a blackbody at some temperature as we move away from the blackbody curve on the diagram? This is an example for which the problem of determining color differences really becomes important.

Equations (11.7) and (11.9) provide a method for determining the “closeness” of two colors on the relevant UCS chromaticity diagrams. These equations can form the basis for determining the closeness of colors to blackbody colors on the chromaticity diagrams. The problem is that if we use different chromaticity diagrams and different “isotemperature lines” to do this analysis, we are likely to come up with different correlated color temperatures for the same source, a problem to be avoided if the concept of correlated color temperature is to have value.

Fortunately, the CIE has standardized the procedure for determining correlated color temperature [1]: “The recommended method . . . is to determine on

a chromaticity diagram the temperature corresponding to the point on the Planckian locus that is intersected by the agreed isotherm line containing the point representing the stimulus. The isotherm lines presently recommended are those normal to the Planckian locus in a chromaticity diagram in which $2v'/3$ is plotted against u' , where v' and u' are the coordinates of the CIE 1976 uniform chromaticity scale diagram . . .”

The transformation provided in this definition converts the (u', v') coordinates of the 1976 UCS standard to those (u, v) of the obsolete 1960 scale, which was designed so that the isotherm lines would be straight lines perpendicular to the Planckian locus. Even though the 1960 UCS color space is obsolete, it is still used in determining correlated color temperature. A widely used plot of lines of constant-correlated color temperature on a portion of the CIE 1931 chromaticity diagram can be found in [2] and [17]. It is based on Kelly's 1963 work [19]. The interested reader can find similar diagrams in the *Handbook of Optics* [20] and in a 1968 paper by Robertson [21]. An approximate version of this diagram, using the isotherm lines as computed by Kelly is shown in Figure 11.11.

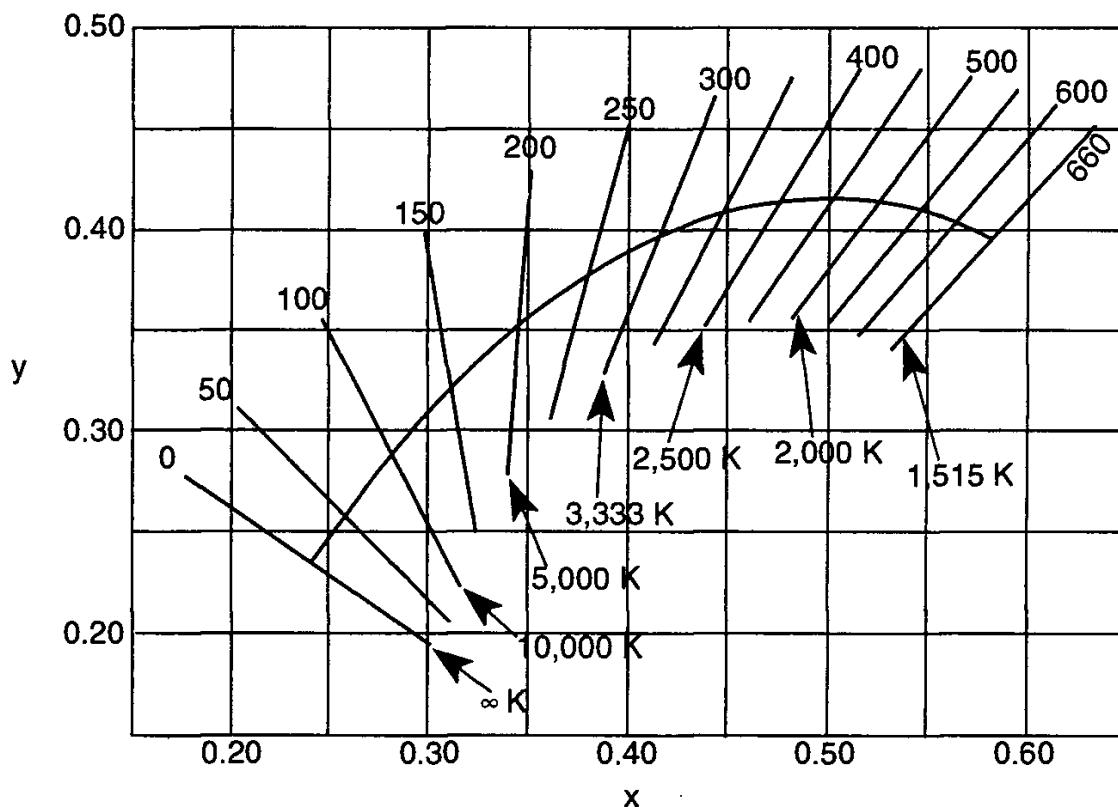


Figure 11.11 Portion of CIE 1931 chromaticity diagram showing isotherm lines as computed by Kelly [19]. Absolute temperatures are indicated below the lines and reciprocal megakelvin ($MK^{-1} = 10^6 \cdot K^{-1}$) temperatures are shown above them. Taking the reciprocal of the temperature in reciprocal megakelvin and multiplying by 10^6 yields the temperature in degrees K.

11.8 STANDARD ILLUMINANTS AND REFLECTION COLORIMETRY

As was mentioned at the beginning of this chapter, the color of an object depends upon the color of the light illuminating it, the source spectrum. The term *illuminant* is used to mean the theoretical spectral distribution of light illuminating an object. The word "source" is restricted to refer to real, practical sources of light [22]. The explanatory comments on p. 39 of [1] further clarify this idea:

A distinction is made between "illuminant" and "source". The term "source" refers to a physical emitter of light, such as a lamp or the sun and sky. The term "illuminant" refers to a specific spectral power distribution, not necessarily provided directly by a source, and not necessarily realizable by a source. The present recommendation first defines "illuminants" by relative spectral power distributions and then "sources". The definition of the sources is considered secondary as it is conceivable that new developments in lamps and filters will bring about improved sources that represent the illuminants more accurately and are more suitable for laboratory use.

Specifying standard illuminants is important for the proper assessment of the inherent color of objects. Of course, the spectral reflectance of an object surface can be measured over the whole visible portion of the spectrum. Since this reflectance spectrum is independent of illuminant, it can be thought of as representing the *inherent color* of the surface and one could use it directly as if it were a color stimulus and determine its chromaticity coordinates this way. The result would be independent of the illuminant, but would represent the actual color of the object *as perceived by a human* only if the object is illuminated by a source with a constant spectral flux distribution over the visible. The only way to produce a reflected flux spectral distribution that matches the spectral reflectance of the object is to illuminate the object with a spectrally flat source. A spectrum such as this is difficult to produce and does not represent any source likely to be encountered in everyday life. It is more reasonable and practical for standard illuminants to be defined that are more characteristic of the light sources found in the real world and that are relatively easy to design, build, and maintain. Thus, colors of reflecting objects are defined with reference to the standard spectral distribution of the illuminant.

Electric lamps are ubiquitously present in nearly every building intended for human habitation in the modern era. Thus electric lamp spectra are important candidates for standardization as illuminants. On the other hand, the bulk of human eye evolution probably took place before so-called "artificial sources" were invented, under conditions of more naturalistic daylight illumination [23]. Daylight is therefore an important source of light, one desired by clothing and fabric buyers, and by painters, sculptors, and other artists for which proper rendering of the colors of objects is essential.

Solar radiation is the operating “fuel” on which all the natural ecosystems and fluid circulatory systems of the planet depend. As nonrenewable energy sources become more difficult and expensive to locate and extract, and as pollution levels produced by them increase, human civilization will be turning toward increased use of solar radiation to power its energy consuming systems. Daylight is particularly well qualified as a replacement for nonrenewable energy sources since it is ready for use, without much modification, for daytime illumination of building interiors. Increased knowledge of the availability of this important light resource will be important for progress in the transition to renewable energy sources. This makes the daylight carried by solar radiation an important candidate for a standard illuminant.

There are a number of other possible sources of light to be considered as candidates. Choosing amongst them depends primarily upon the intended application and secondarily upon their ease of use and maintenance.

The concept of measuring the color rendering properties of illuminants is discussed in the next section. In this section, attention is focused on describing illuminants that have been standardized for colorimetry and for use in determining the colorimetric properties of materials.

The *IES Lighting Handbook* discusses the need for standard illuminants and sources and offers the following survey of standards writing organizations dealing with light sources [24]:

Standard specifications for the color and spectral quality of light sources for special applications have been adopted in several fields. Some have been adopted formally, as in the graphic arts field, by the IES; others, as for photographic prints and transparencies, by the American National Standards Institute; and still others by the American Society for Testing and Materials. Other less formal standard specifications are recommended, as those for textiles by the American Association of Textile Chemists and Colorists, and for diamond grading by the Gemological Institute of America. In many of these, the target standard is the spectral quality of daylight at around 7,500K, sometimes used with an additional lamp at lower color temperature, sometimes with the addition of ultraviolet.

11.8.1 Blackbody Illuminants

Because of the importance of blackbodies in radiometry and photometry for color rendering purposes, and because the Planckian spectral distribution can be calculated precisely, it is natural to expect blackbody spectra to play a role in colorimetry as standard illuminants. Accordingly, the CIE defines standard illuminant A as light from a *full radiator* (meaning a blackbody) at absolute temperature

2,856K according to "The International Practical Temperature Scale, 1968 (ITS-68)." Thus the relative spectral flux distribution can be calculated using Planck's radiation formula, (3.4), for any set of wavelengths over the visible. Alternatively, the data can be taken from tables of values for this illuminant at 1-nm intervals.

According to Wyszecki and Stiles [5], the relative spectral radiant power distribution of illuminant A has not changed since it was first recommended by the CIE in 1931. However, the value for the radiation constant c_2 has changed. The CIE changed the temperature from the original value of 2,848 in order to preserve the spectral shape of illuminant A. Making adjustments of this sort have become a policy of the CIE so that future changes in the international temperature scale do not change the spectral distribution of illuminant A.

The ITS was changed again in 1990. According to Mielenz [25], "The color temperature assigned to standard illuminant A is a descriptive parameter which depends on the value of c_2 assigned in the international temperature scale. Its value is approximately 2,856K on the IPTS-68 and remains the same on the ITS-90 because the value of c_2 has not been changed."

With standard illuminant A defined by its spectral radiant power distribution, the next question is one of how to approximate it with a real source of light. It is possible to approximate illuminant A with a tungsten filament lamp. Since tungsten is not a true blackbody radiator, a temperature of 3,650K is needed to achieve an approximation of the spectral shape of a blackbody at 2,856K. If a tungsten filament is operated close to its melting point temperature of 3,650K, tungsten can evaporate from the filament and condense on the inside of the glass envelope. Nevertheless, tungsten filament lamps were used for many years in many national standardizing laboratories to provide an approximation of illuminant A [26].

Zubler and Mosley developed the quartz halogen lamp [27] around 1959. The quartz envelope permits operation at higher temperatures and the halogen participates in a cyclic chemical reaction that deposits evaporated tungsten back on the hot filament. This results in longer lasting lamps operated at elevated temperatures [28]. The system works best if the inside surface of the quartz envelope is maintained at a temperature near to 600°C. The spectral radiant flux from these lamps is similar to that of a Planckian radiator and includes a relatively correct quantity of UV radiation. UV radiation is normally lost with ordinary tungsten filament lamps. This is largely the result of the low UV transmittance of the glass envelope. Quartz has a higher UV transmittance. Because of the nonuniform spectral transmittance of quartz and the spectral variation in tungsten emittance, a quartz halogen lamp is operated at a somewhat elevated temperature (typically near 3,000K) to provide a good spectral match to a blackbody radiator at 2,856K. Operating a gas-filled tungsten filament lamp with a fused quartz envelope in this manner is the CIE-recommended standard source for illuminant A.

11.8.2 Daylight Illuminants

Although the color rendering quality of daylight is superb, it has the unfortunate characteristic of changing from minute to minute and hour to hour throughout the day. Nevertheless, it is *the* common reference and presumed best spectral model for illuminating objects to produce the best reflected color.

Consequently, there have been a number of measurements over the years of the spectral distributions of irradiance over the visible spectrum from sun and sky on clear, cloudy, and overcast days, for varying solar altitude angles (angle from the horizon to the center of the solar disk). Judd, MacAdam, and Wyszecki [29] reported on a number of measurements of these spectral distributions and this work formed the basis for the definitions of some standard daylight illuminants. Chromaticity coordinates for a large number of daylight measurements are plotted along with the Planckian locus (from 3,500 K to ∞) and a set of correlated color temperature isotemperature lines in the *IES Lighting Handbook* (1984 reference volume, Figure 5-27). The locus of chromaticities taken to be typical of daylight conforms to the relation

$$y = 2.870x - 3.000x^2 - 0.275 \quad (11.11)$$

on a portion of the CIE 1960 UCS chromaticity diagram, (taken from the caption to Figure 2 in [29]).

The need for daylight-like illuminants led the CIE to propose standard illuminants B and C. Illuminant B was intended to represent direct sunlight with a correlated color temperature of approximately 4,874K and illuminant C was intended to represent average daylight with a correlated color temperature of 6,774K. Relative spectral power distributions for these sources are tabulated and plotted in [17]. Both of these illuminants are currently considered to be inadequate in representing the intended variations of daylight, as both are seriously deficient in the ultraviolet region. UV radiation plays an important role in fluorescent materials. Illuminants B and C have been completely dropped from ISO/CIE International Standard 10526 [30].

In order to realize illuminants B and C, the CIE previously specified that standard sources B and C be realized by interposing specified spectrally selective filters between a source intended to match illuminant A and the object to be illuminated. The filters consisted of two water solutions, B_1 and B_2 for illuminant B and C_1 and C_2 for illuminant C, in colorless glass cells in series. The specific chemicals dissolved in the water for these four filters, called Davis-Gibson filters, are provided in Table 1 (3.3.5) on p. 148 of [17].

Standard daylight illuminants B and C have given way to a new illuminant, designated D₆₅, recommended by the CIE in 1964 as the main standard daylight illuminant. It is intended to represent a phase of natural daylight with a correlated

color temperature of approximately 6,500 K. It is currently defined by a specific spectral power distribution, independent of any blackbody temperature. The exact value for the temperature associated with the D₆₅ distribution depends upon the value adopted for the constant c_2 in Planck's radiation formula and on the convention used for assigning correlated color temperature to stimuli whose chromaticities do not lie exactly on the Planckian locus. According to ISO/CIE standard 10526: 1991, "Using the convention that lines of constant correlated colour temperature are normal to the Planckian locus in a chromaticity diagram in which $2v'/3$ is plotted against u' , the correlated colour temperature was equal to 6,500 K when $c_2 = 1.4380 \times 10^{-2} \text{ m}\cdot\text{K}$; with $c_2 = 1.4388 \times 10^{-2} \text{ m}\cdot\text{K}$, it is equal to $6,500(14,388/14,350)\text{K}$." u' and v' are coordinates in the CIE 1976 uniform chromaticity scale diagram depicted in Figure 11.9, and the color temperature for illuminant D₆₅ currently is 6,517K.

The spectral distribution of flux in standard illuminant D₆₅ is tabulated at 1-nm intervals in ISO/CIE Standard 10526 from 300 nm to 830 nm and is available on CIE computer disk CIE D001: 1988. This standard practice document makes the statement: "At present, no source is recommended for realizing CIE standard illuminant D₆₅. Practical sources intended for this purpose can be assessed by a method described in CIE Publication 51."

The spectral distribution of illuminants A and D₆₅ are plotted in Figure 11.12. CIE Publication 51 specifies [31] that the spectral power distribution of a source being considered to represent a standard illuminant is determined by spectroradiometry over the near ultraviolet and visible wavelength ranges. This measured spectral distribution is normalized by using the CIE 1964 Supplementary Standard Observer color matching function $\bar{y}_{10}(\lambda)$. The resulting normalized spectral distribution is designated $S_n(\lambda)$. The chromaticity coordinates (u'_{10}, v'_{10}) of the test source are determined and the resulting point on the CIE 1976 UCS diagram for the CIE 1964 Supplementary Standard Observer (Figure 11.10) must lie within a circle of radius 0.015 centered on the point on the diagram representing the color of standard illuminant D₆₅. Transforming this circle on the CIE 1976 UCS diagram to the CIE 1964 color space results in an ellipse centered at the D₆₅ point having approximate (x_{10}, y_{10}) color coordinates of (0.32, 0.33). Due to the effect of metamerism, CIE Publication 51 provides additional limitations on the nature of the source standard spectral distribution. These limitations are based on the computation of a set of *metamerism indices* computed according to both CIELAB and CIELUV color difference formulas.

If the description of this section sounds complicated, it is. The complications result both from the necessarily abbreviated description provided and the real complexity of the subject matter. The interested reader should consult the references cited for more information.

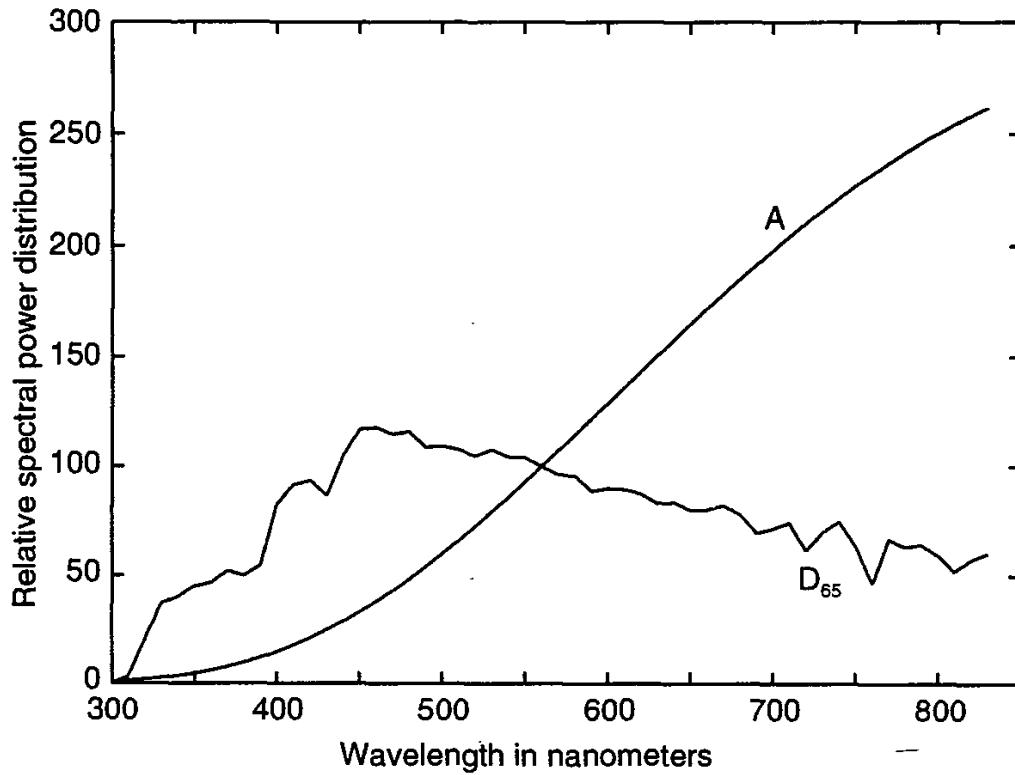


Figure 11.12. Relative spectral power distributions of CIE standard illuminants A and D₆₅.

11.8.3 Reflection Colorimetry

ASTM Standard E 308-90 offers [13] a defined procedure for obtaining the tristimulus values X , Y , and Z as well as X_{10} , Y_{10} , and Z_{10} from spectral reflectance or transmittance data for specimens of objects. Spectral power distributions are tabulated from 380 to 760 nm in 5-nm intervals for standard illuminants A, C, D₅₀, D₅₅, D₆₅, and D₇₅. (Although the CIE recommends only the illuminant D₆₅ for daylight currently, the ASTM standard still provides data for “daylight” illuminants at the color temperatures of 5,000, 5,500, and 7,500 K anyway.) If high accuracy is required, calculation at 1-nm intervals is suggested. Letting $S(\lambda)$ stand for the spectral distribution of the selected standard illuminant, $\bar{x}(\lambda)$, $\bar{y}(\lambda)$, and $\bar{z}(\lambda)$ stand for either the CIE 1931 or 1964 color matching functions, and $R(\lambda)$ stand for either the spectral reflectance factor or spectral transmittance factor of the material sample, the tristimulus values for the sample are given by

$$\begin{aligned} X &= k \sum_{380}^{780} R(\lambda) S(\lambda) \bar{x}(\lambda) \\ Y &= k \sum_{380}^{780} R(\lambda) S(\lambda) \bar{y}(\lambda) \\ Z &= k \sum_{380}^{780} R(\lambda) S(\lambda) \bar{z}(\lambda) \end{aligned} \quad (11.12)$$

where k is a normalization constant given by

$$k = \frac{100}{\sum_{380}^{780} S(\lambda) \bar{y}(\lambda)} \quad (11.13)$$

(Reflectance factor is defined in Section 6.7.2. Transmittance factor can be given a similar definition.) The chromaticity coordinates are then obtained in the usual way using (11.4). When a color space more nearly uniform than the above is desired, CIELAB or CIELUV are recommended and the equations given previously for these spaces are provided in the standard. By this means, the colors of objects are determined from the color of light transmitted or reflected by them when they are theoretically illuminated by a standard illuminant or actually illuminated by a standard source intended to emulate a standard illuminant.

11.9 COLOR RENDERING INDEX

One way to be sure that objects look their intended color is to always illuminate them with light of the same spectral distribution as the standard illuminant used to determine their color by the method outlined in Section 11.8.3. This procedure is not practical in general, however, since it is not possible to control the spectral distributions of all light sources illuminating objects after they leave the manufacturer's premises. The question then becomes one of how different objects look when illuminated by other sources. This is the core question of a subset of illuminating engineering dealing with what is called *color rendering*.

Illuminating objects that are supposed to look green with light having plenty of red and blue (but little green) will change the apparent color of the object and lead to poor color rendering. A more subtle problem is that two sources of the same color can cause the same colored object to look different when illuminated by the two sources. The reason for this is that the two sources of the same color can have drastically different spectral distributions, and the spectral distribution of reflectance (or transmittance) of the object can be strongly varying and can there-

fore respond differently to the differently weighted regions of the source spectrum. See Section 11.2.

The term color rendering refers to the ability of a source of light to properly render the colors of a variety of differently colored objects. In general, the broader and more uniform the spectral distribution of a light source over the visible, the better will be its color rendering properties. The best color rendering is achieved by a source whose spectral distribution matches that of the standard illuminant used to specify the colors of objects as described in the previous section.

Until the last several decades, the most common sources of light for illuminating objects were daylight and incandescent lamps. Both of these sources have a generally broad distribution of wavelengths. Although these sources are different in color, humans have become accustomed to the generally slight color changes exhibited by objects when going from one source to the other. With the advent of gas discharge lamps, tolerance of color changes is no longer true. It is the nature of light emitted from electrically ionized gases such as those found in neon signs and low-pressure gas discharge tubes to be concentrated in a number of narrow, almost monochromatic, spectral lines spread irregularly across the visible portion of the spectrum. The particular spectral lines produced depend upon the gas or mixture of gases used in the discharge lamp. Such lamps generally have poor color rendering properties. Picking up one or two of the colors in an object and rejecting the remaining ones, these sources produce a shift in the color appearance of the object.

In an effort to produce spectral distributions from such lamps that provide better color rendition, phosphor blends were added to the inside of the glass envelopes of these lamps to convert by fluorescence some of the near-UV radiation produced to a broader spectrum of visible light. Spikes or concentrations of relatively monochromatic radiation are still emitted by these lamps, however, and color rendering problems can still persist.

As a result of these developments in electric lamp technology, the CIE developed a system for rating the color rendering properties of lamps [32]. The CIE vocabulary defines color rendering as the "effect of an illuminant on the colour appearance of objects by conscious or subconscious comparison with their colour appearance under a reference illuminant."

Note that the color rendering property of a lamp is not an absolute characteristic—it must necessarily depend upon the spectral properties of the reference illuminant with which it is compared. The CIE has specified a method of comparing lamps with reference illuminants that results in the calculation of what is called a color rendering index (CRI).

Of course, the first question that arises is what illuminant should be used as the standard one? It was decided to have a different reference illuminant for each color temperature produced by a lamp. Some lamps intentionally have a rather warm appearance, resulting from a preponderance of red wavelengths and a min-

imal amount of blue. Such lamps impart this same “warm” appearance to objects illuminated by them. Others are intended for high brightness illumination of objects with a full range of wavelengths present, producing color renderings more like that produced by natural daylight.

The CIE took advantage of this natural division of lamp types and specified two different types of reference illuminants. One is for lamps producing correlated color temperatures below 5,000K and another for correlated color temperatures above 5,000K. In the first case, the reference illuminant is a blackbody radiator having a color temperature near to the correlated color temperature of the lamp being evaluated. In the second case, the reference illuminant is one of a series of theoretical spectral power distributions emulating different phases of natural daylight.

Once a reference illuminant appropriate to the lamp being tested has been selected, one must next choose a set of standard-colored test objects. The CIE has specified fourteen test color samples for calculating the CRI. Test colors 1 through 8 are of medium saturation and they are distributed rather uniformly around the hue circle in color space. Test colors 9 through 12 are more saturated. Colors 13 and 14 approximate Caucasian complexion and leaf green, respectively.

Light from the reference illuminant and from the test lamp is theoretically made to fall on each of these standard test objects and the color coordinates of the reflected illumination are determined for each object and for each of the two sources using the methods outlined in Section 11.8.3. To determine the color rendering of the test lamp, the differences ΔE_i in reflected color between the two lamps is calculated for the i th one of each of the fourteen test objects and the resulting differences are averaged over the first eight test objects to obtain the general color rendering index R_a :

$$R_a = \frac{1}{8} \sum_{i=1}^8 (100 - 4.6 \Delta E_i) \quad (11.14)$$

with the ΔE_i being calculated using the CIE 1964 Color-Difference Formula, similar to (11.8), but with ΔL^* replaced with Δw^* . This formula is scaled (by the 4.6 factor) so that a warm white fluorescent lamp has a general color rendering index (GCRI) of 50 [6]. Standard test colors 9 through 14 are used for more specialized color rendering applications. This scheme places the color rendering index for warm white fluorescent lamps about midway along a scale from 0 to 100. Values less than 0 are possible.

For more detailed information about the determination of color rendering indices, the reader is referred to the CIE standard [32] and the *IES Lighting Handbook* [24]. Sample CRI values for several common sources of light are pro-

vided in Table 11.1 along with their CIE 1931 chromaticity coordinates and approximate correlated color temperatures.

Table 11.1
Color Rendering Indices of Common Lamp Types

<i>Name of Light Source</i>	<i>CIE 1931 Chromaticity Coordinates</i>		<i>Correlated Color Temperature</i>	<i>CIE General Color Rendering Index</i>
	<i>x</i>	<i>y</i>		
Warm white fluorescent	0.436	0.406	3020	52
Cool white fluorescent	0.373	0.385	4250	62
Simulated D ₆₅ standard illuminant	0.313	0.325	76520	91
Clear mercury vapor	0.326	0.390	5710	15
Metal halide	0.396	0.390	3720	60
High-pressure Xenon	0.324	0.324	5920	94
High pressure sodium	0.519	0.418	2100	21
Low-pressure sodium	0.569	0.421	1740	-44
DXW tungsten halogen	0.424	0.399	3190	100

REFERENCES

- [1] CIE. "Colorimetry," Publ. CIE No. 15.2, Central Bureau of the CIE, Kegelgasse 27, A-1030 Vienna, Austria, 1986. Also available from the CIE National Committees of member countries. In the U.S., it is available from TLA-Lighting Consultants, Inc., 7 Pond St., Salem, MA 01970-4893.
- [2] Miller, K. A., and W. E. Schneider. *Colorimetry: Methods and Tools*, Photonics Design & Applications Handbook, 1990 ed., Laurin Publishing Co., Inc.: Berkshire Common, P. O. Box 1146, Pittsfield, MA.
- [3] Robertson, A. R. "Color Perception," *Physics Today*, Dec. 1992, pp. 24-29.
- [4] Grum, F. and C. J. Bartelson. *Optical Radiation Measurements*, Vol. 2, Color Measurement, Academic Press: New York, NY, 1980.
- [5] Wyszecki, G., and W. S. Stiles. *Color Science: Concepts and Methods*, Quantitative Data and Formulae, 2nd ed., John Wiley & Sons: New York, NY, 1982.
- [6] IES Color Committee. "Color and Illumination," IES-DG-1-1990, Ill. Eng. Soc. No. Amer., 345 E. 47 St., New York, NY 10017, 1990.
- [7] CIE. *International Lighting Vocabulary*, Publ. No. 17.4, Central Bureau of the CIE, A-1033 Vienna, P. O. Box 169, Austria, 1987. Jointly published as "International Electrotechnical Vocabulary," Publication 50(845), Bureau Central de la Commission Electrotechnique Internationale, 3 rue de Varembe, Geneve, Suisse, 1987.
- [8] Bartelson, C. J. "Colorimetry," Ch. 3 in *Optical Radiation Measurements*, Vol. 2, ed. by F. Grum and C. J. Bartelson, Academic Press: New York, NY, 1980.
- [9] Kelley, K. L., and D. B. Judd. *Color: Universal Language and Dictionary of Names*, NIST Special Publication SP440, Washington DC: U. S. Gov. Printing Office, 1978.

-
- [10] Munsell, A. H. *A Color Notation*, Munsell Color Company, Macbeth Div. of Kollmorgen Corp., Box 230, Newburgh, NY 12550, undated.
 - [11] *Munsell Book of Color*, Munsell Color Company, Macbeth Div. of Kollmorgen Corp., Box 230, Newburgh, NY 12550, undated.
 - [12] CIE. "CIE 1988 2° Spectral Luminous Efficiency Function for Photopic Vision," CIE Technical Report Publ. No. CIE 86, 1st ed., CIE Central Bureau, Kegelgasse 27, A-1030 Vienna, Austria, 1990.
 - [13] ASTM. "Test Method for Computing the Colors of Objects by Using the CIE System," E 308-90, Amer. Soc. Test. Materials, 1916 Race St., Philadelphia, PA 19103, 1991.
 - [14] ASTM. *ASTM Standards on Color and Appearance Measurement*, 3rd ed., Amer. Soc. Test. Materials: Philadelphia, PA, 1991.
 - [15] It is noted that the CIE color matching functions are provided to 830 nm at the edge of the infrared portion of the spectrum. The spectrum locus for the 1964 standard observer reverses direction at around 770 nm, so that a line drawn from 360 to 830 nm on the corresponding chromaticity diagram fails to include a few visible colors that would fall below this line.
 - [16] Minolta Corporation. "Precise Color Communication—Color Control from Feeling to Instrumentation," Minolta Camera Company, Ltd., 3-13, 2-Chome, Azuchi-Mochi, Chuo-Ku, Osaka 541, Japan, 1991.
 - [17] Guild, J. "The Colorimetric Properties of the Spectrum," *Phil. Trans. Roy. Soc. A*, Vol. 230, 1931, pp. 149–187.
 - [18] Wright, W. D. "A Re-determination of the Trichromatic Coefficients of the Spectral Colors," *Trans. Opt. Soc.*, Vol. 30, 1928–1929, pp. 141–164.
 - [19] Kelly, K. L. "Lines of Constant Correlated Color Temperature Based on MacAdam's (u,v) Uniform Chromaticity Transformation of the CIE Diagram," *J. Opt. Soc. Am.*, Vol. 53, 1963, p. 999.
 - [20] Driscoll, W. G., and W. Vaughn. *Handbook of Optics*, McGraw-Hill Book Co.: New York, NY, 1978.
 - [21] Robertson, A. R. "Computation of Correlated Color Temperature and Distribution Temperature," *J. Opt. Soc. Am.*, Vol. 58, 1968, p. 1528.
 - [22] Simon, F. T. "Modern Illuminants," in *Optical Radiation Measurement*, Vol. 2, ed. by F. Grum and C. J. Bartelson, Academic Press: New York, NY, 1980.
 - [23] Corth, R., "What is 'Natural' Light?" *Lighting Design and Application*, April 1983, pp. 34–40.
 - [24] Kaufman, J. E., and J. F. Christensen, eds. *IES Lighting Handbook 1984 Reference Volume*, Illuminating Engineering Society of North America: New York, NY, 1984.
 - [25] Mielenz, K. D. "Effects of the International Temperature Scale of 1990 (ITS-90) on CIE Documentary Standards for Radiometry, Photometry, and Colorimetry," *Optical Radiation News*, No. 55, Council on Optical Radiation Measurements, Spring, 1991.
 - [26] Tarrant. A. Private communication, 1993.
 - [27] Zubler, E. G., and F. A. Mosby. *Illum. Eng.*, Vol. 54, 1959, p. 734.
 - [28] Simon, F. T. "Modern Illuminants," *Optical Radiation Measurements*, Vol. 2, ed. by F. Grum and C. J. Bartelson, Academic Press: New York, NY, 1980, p. 152.
 - [29] Judd, D. B., D. L. MacAdam, and G. Wyszecki. "Spectral Distribution of Typical Daylight as a Function of Correlated Color Temperature," *J. Opt. Soc. Am.*, Vol. 54, 1964, p. 1031. Summarized in: *Illum. Eng.*, Vol. LX, April, 1965, p. 271.
 - [30] ISO/CIE. "CIE Standard Colorimetric Illuminants," ISO/CIE 10526: 1991 (E), International Organization for Standardization, Case postale 56, CH-1211, Geneve, Switzerland.
 - [31] CIE. "A Method for Assessing the Quality of Daylight Simulators for Colorimetry," Publ. No. 51 (TC-1.3) 1981, CIE Central Bureau, Vienna, Austria. 31.
 - [32] CIE. *Method of Measuring and Specifying Colour Rendering Properties of Light Sources*, 2nd ed., CIE Publ. No. 13.2, CIE Central Bureau, Vienna, 1974.

Appendix

Correspondence Between Finite Elements and the Calculus

A.1 INTRODUCTION

Modern calculus was developed by Sir Isaac Newton in England around 1666, in order to help him solve problems in physics, and by Baron von Leibniz in Germany. Leibniz published his version of the differential calculus in 1684, three years before Newton's treatise on the subject. It is the Leibniz notation that we use today, a notational system well adapted to the concepts of radiation transfer quantified in radiometry and photometry. Two forms of the calculus are used throughout this book, as in the following examples

$$\frac{d\Phi}{d\lambda} \quad \frac{d\Phi}{ds_o} \quad \int L \cos \theta d\omega \quad (\text{A.1})$$

The first two operations above are called derivatives and the last is called an integral. Definitions will now be given for the mathematical operations symbolized by these notations, using finite differences in the quantities involved.

Once the finite difference approach has been used to *define* differentiation and integration, the approach will be turned back around and used to show how finite differences can provide approximate formulas to replace derivatives and integrals with simple quotients and sums.

This discussion has two purposes. First, differential and integral calculus may seem very mysterious to people not trained in its intricacies. The finite differences that are used to replace the derivatives and integrals should seem less imposing. Secondly, when one makes actual measurements of quantities in radiometry or photometry, the data is obtained over finite intervals of the variables involved. Differential and integral calculus are seldom needed. If this is the case, one might ask, why the bother with differentiation and integration at all?

One answer is that the notation and concepts of differentiation and integration are very precise and exact. There are no approximations associated with solving most problems using this system of mathematics. Another answer is found in Chapter 4 where it is shown that real problems of flux transfer from a source surface to a receiving surface can become very difficult mathematically. Expressing these problems in terms of integral calculus provides powerful tools for solving these difficult problems exactly, at least in theory. Tables of the integrals of a very large number of commonly found functions have been published [1–2]. If a function that needs to be integrated can be found in one of these tables, then an enormous amount of work can be avoided trying to perform the integrations step by step.

In summary, the notation and manipulative rules of differential and integral calculus are very useful in setting up and solving problems of radiative transfer. In reducing the resulting solutions to practice or in dealing with measurement results, however, we leave the realm of calculus and convert our equations to their finite difference forms. The latter part of this appendix is devoted to showing how this conversion is done and to providing examples to illustrate the processes involved.

A.2 DEFINITION OF THE DERIVATIVE

The meaning of the differential notation dQ/dx is this: A finite, measurable difference $\Delta Q = Q(x + \Delta x) - Q(x)$ in some quantity Q is to be divided by the corresponding difference Δx in some other quantity on which the first quantity depends (is functionally dependent). After the differences in numerator and denominator have been shrunk down continuously until they reach a limit where they are immeasurably, infinitesimally small, the derivative of the function $Q(x)$ with respect to x is defined:

$$\frac{dQ(x)}{dx} \equiv \lim_{\Delta x \rightarrow 0} \frac{\Delta Q}{\Delta x} \quad (\text{A.2})$$

Even though the denominator on the right in the above equation is made to shrink to zero, the fraction itself does not blow up (become arbitrarily large) because the numerator shrinks along with the denominator and the quotient remains finite. If one looks at a plot of the function $Q(x)$ versus x over the range of Q , like the one shown in Figure A.1, it is clear that the quotient on the right side of (A.2) is the average slope of the curve over the range from $x = x_1$ to $x + \Delta x = x_2$.

The reason for shrinking Δx down toward zero in the definition of the derivative is so that the quotient in (A.2) is averaged over a smaller and smaller range Δx of x . In the limit as Δx goes to zero, the *average* slope becomes the *actual* slope of the curve at the value of x where the derivative is desired.

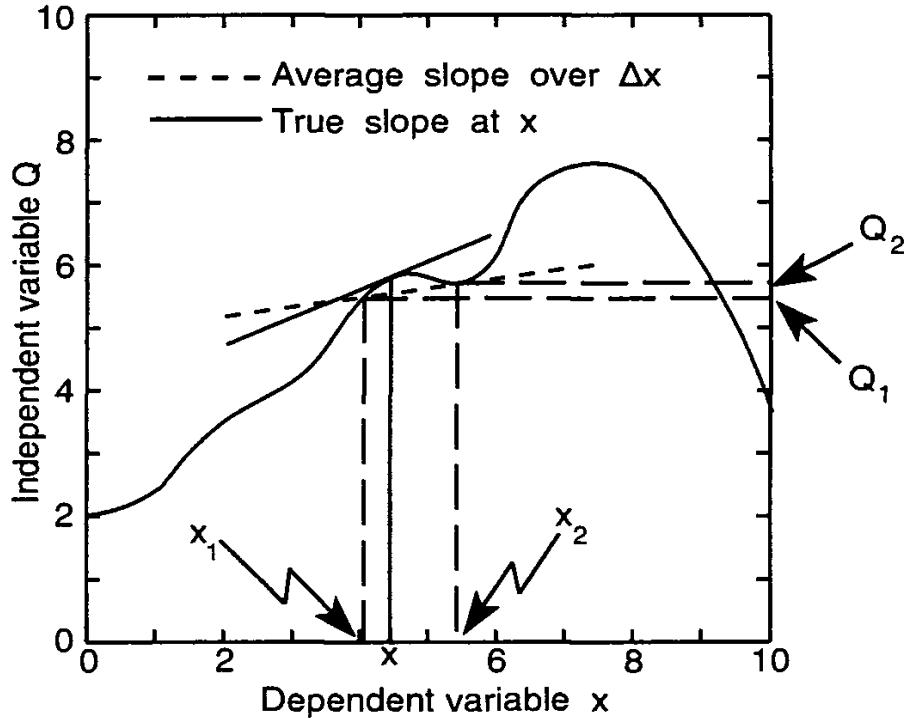


Figure A.1 Illustration of the equivalence between the derivative of a function and the slope of the plot of that function.

Now let's turn the process around and go the other way. Derivatives are fine for mathematical functions $Q(x)$, but infinitesimally small physical quantities, or differences in those quantities, are unmeasurable. In practice, one measures a finite amount of or change in some radiometric quantity ΔQ and divides it by a finite amount of or change in the other quantity Δx on which it depends. The smaller Δx can be made to be the better the quality of the measurement, but in the end what one actually measures is an average of the radiometric quantity over the range Δx of the independent variable.

Obviously, for the purposes of determining the value of the derivative at a specific value of the independent variable x , the differences used in the ratio are desired to be as small as possible, but they cannot be zero or they will not be measurable. Thus, one can easily approximate derivatives wherever they appear with ratios of finite differences.

An example of the use of the derivative, is in the definition of irradiance, (1.9) in Chapter 1:

$$E = \frac{d\Phi}{ds_o} \quad (\text{A.3})$$

Writing this using the limit notation yields

$$E = \frac{d\Phi}{ds_o} = \lim_{\Delta s_o \rightarrow 0} \frac{\Delta\Phi}{\Delta s_o} \quad (\text{A.4})$$

where $\Delta\Phi$ is the quantity of flux that impinges on or passes through small area Δs , containing the point at which the irradiance is to be evaluated. In practice, one measures a small but finite quantity of flux $\Delta\Phi$ distributed over a small but finite area Δs_o and determines the average irradiance over the area by taking the ratio $\Delta\Phi/\Delta s_o$.

Another example is in the definition of spectral irradiance, (1.10) in Chapter 1:

$$E_\lambda = \frac{dE}{d\lambda} \quad (\text{A.5})$$

An equivalent way of writing this quantity is

$$\frac{dE(\lambda)}{d\lambda} = \lim_{\Delta\lambda \rightarrow 0} \frac{\Delta E}{\Delta\lambda} \quad (\text{A.6})$$

where ΔE is the quantity of irradiance within the wavelength interval $\Delta\lambda$ centered at wavelength λ .

A.3 DEFINITION OF THE INTEGRAL

Consider some quantity, say B , defined to be the derivative of some other quantity, say Q , with respect to an independent variable such as x , i.e.

$$B = \frac{dQ}{dx} \quad (\text{A.7})$$

Since B might be different for different values of x , B is written as a function $B(x)$ of x . Knowing $B(x)$, it is desired to find how much of Q there is over some range of x .

The range from x_1 to x_2 is divided into N equal intervals of width Δx . Then $B(x)$ is evaluated in the middle of each of these intervals, as is illustrated in Figure A.2. The value of $B(x)$ in the i th interval is designated B_i . Next the products of all the B_i values with the width Δx of each interval are summed up and the limit is taken as Δx goes to zero (and, correspondingly, N goes to infinity). This process leads to a definition of the integral Q_{12} of the function $B(x)$ over the interval from x_1 to x_2 :

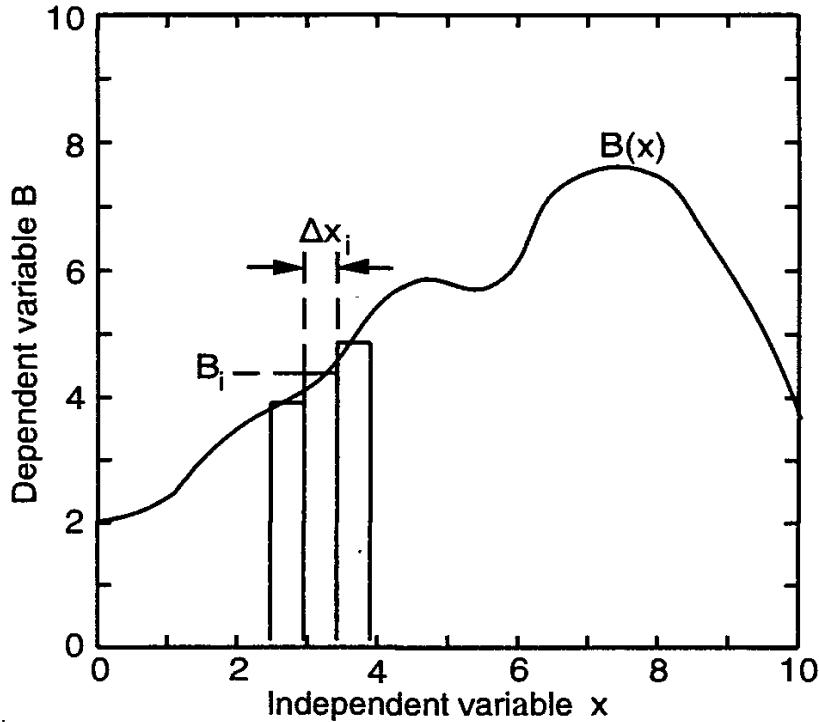


Figure A.2 The continuous function $B(x)$ is approximated by a finite set of discrete values, B_i at each of a set of finite values x_i .

$$Q_{12} \equiv \int_{x_1}^{x_2} B(x)dx = \lim_{\substack{\Delta x \rightarrow 0 \\ N \rightarrow \infty}} (B_1 \Delta x + B_2 \Delta x + B_3 \Delta x + \cdots + B_N \Delta x) \quad (\text{A.8})$$

or

$$Q_{12} \equiv \int_{x_1}^{x_2} B(x)dx = \lim_{\substack{\Delta x \rightarrow 0 \\ N \rightarrow \infty}} \sum_{i=1}^N B_i \Delta x = \lim_{\substack{\Delta x \rightarrow 0 \\ N \rightarrow \infty}} \Delta x \sum_{i=1}^N B_i \quad (\text{A.9})$$

Another way of expressing the process of integration is to solve (A.7) for dQ to obtain

$$dQ = B(x)dx \quad (\text{A.10})$$

Then, to get the total amount of Q over some range of x , say from x_1 to x_2 , both sides of (A.10) are integrated over that range as follows:

$$Q = \int dQ = \int_{x_1}^{x_2} B(x)dx \quad (\text{A.11})$$

Notice that the integral of a function $B(x)$ over a certain range $[x_1, x_2]$ is equivalent to the “area” under a plot of $B(x)$ versus x over that range. It is approximated by the sum shown in (A.9) of the areas of all the rectangles of width Δx and height B_i shown in Figure A.2 over the range $[x_1, x_2]$, with the “limit” part of (A.9) removed. The smaller Δx and the larger N are, the better will be the approximation.

A.4 INTEGRALS AS SUMS

Now that the integral is defined in terms of finite elements, the process can be reversed. As was just mentioned, sums can be used to approximate integrals. In measurement situations, for example, one might have some finite amount of a radiometric quantity B_i measured at each of a finite number of values of the dependent variable x . The results are multiplied by the finite differences Δx_i in the dependent variable for each value of B_i that was measured. The smaller the Δx_i ’s can be made to be, the closer the measurement represents the true situation.

The integrals of spectrally and/or angularly varying flux distributions found in this book cannot generally be represented by simple algebraic or transcendental functions, since they are usually the result of measurements taken at discrete intervals of the variables involved. Their integrals generally cannot be found in published tables of integrals of mathematical functions, unless the discrete measurements can be well fitted to some analytical curve whose functional description *can* be integrated analytically. The integrals of discrete values can however be approximated by measuring finite elements of the quantities involved and calculating the finite sums as indicated above. The fundamental correspondence is shown as

$$\int_{x_1}^{x_2} B(x)dx \approx \sum_{i=1}^N B_i \Delta x_i \quad (\text{A.12})$$

If all the Δx_i ’s are of the same width, Δx , then

$$\int_{x_1}^{x_2} B(x)dx \approx \Delta x \sum_{i=1}^N B_i \quad (\text{A.13})$$

with $\Delta x = (x_2 - x_1)/N$. All the integrals found in this book can be approximated with sums.

An example of this process is shown in the next section. It is an evaluation of the first integral equation found in this book, (1.17) in Chapter 1, an integral of the radiance over solid angle.

There are various techniques for improving the accuracy of the approximations shown above. For example, one can guess at a functional form for B_i in each interval Δx_i , then integrate this analytically, separately for each interval, and then sum up the resulting integrals. The same improvement in accuracy can be achieved by making the intervals Δx_i much smaller.

In many cases double or triple integrals are involved, such as

$$E = \int_0^\infty \int_0^{2\pi} \int_0^{\pi/2} L_\lambda(\theta, \phi) \cos \theta \sin \theta d\theta d\phi d\lambda \quad (\text{A.14})$$

and the expression of these in terms of multiple summations can be challenging. The integration over solid angle, in particular, can be tricky and is addressed in the next section.

Readers desiring more information about calculus are directed to any of a number of either new or old textbooks on this subject. An interesting one was written originally by Silvanus P. Thompson, first President of the Illuminating Engineering Society of London. It was first published in 1908, and was reprinted not too many years ago [3]. A more recent one by Johnson and Kiokemeister was published in 1960 [4]. A still more recent one is the 1987 text by Stein [5].

A.5 SUMS OVER SOLID ANGLES

The need to integrate over a solid angle is encountered frequently in this book. Usually spherical coordinates are used, the element $d\omega$ of solid angle is written as $d\omega = \sin \theta d\theta d\phi$, and the integration becomes one over the ranges of θ and ϕ that cover the solid angle of interest. For example, when a hemispherical solid angle is the range of integration, the limits of the two integrations are from 0 to 2π for ϕ and 0 to $\pi/2$ for θ . If one attempts to represent these integrations with sums over equally spaced intervals in θ and ϕ , errors in the results can be produced.

Here is how to do it correctly. Although in what follows integrations are performed over a hemispherical solid angle, the results can be transformed to integrals over smaller or larger ones. The treatment begins with (1.17) and (1.18) in Chapter 1, reproduced here as

$$E = \int_{2\pi} L(\theta, \phi) \cos \theta d\omega \quad (\text{A.15})$$

$$E = \int_0^{2\pi} \int_0^{\pi/2} L(\theta, \phi) \cos \theta \sin \theta d\theta d\phi \quad (\text{A.16})$$

Suppose that values for $L(\theta, \phi)$ in (A.15) are provided only at N directions over the hemisphere. First we associate with each of these directions a small but finite solid angle $\Delta\omega_i$ in such a manner that the N solid angles all add up to the 2π sr of the hemisphere; that is,

$$\sum_{i=1}^N \Delta\omega_i = 2\pi \quad (\text{A.17})$$

With the discrete but finite values for $\Delta\omega_i$ in hand, the finite element version of (A.15) can be written

$$E = \sum_{i=1}^N L_i \cos \theta_i \Delta\omega_i \quad (\text{A.18})$$

If the directions for the L_i values are uniformly distributed over the hemisphere, then all the $\Delta\omega_i$ can be the same, equal to $2\pi/N$, and (A.18) becomes

$$E = \frac{2\pi}{N} \sum_{i=1}^N L_i \cos \theta_i \quad (\text{A.19})$$

Now let's examine the case where the directions in which L is measured are *not* uniformly distributed over the hemisphere, but are based on equal intervals in θ and ϕ , so that values L_{ij} for the radiance/luminance are known at the angles specified by

$$\begin{aligned} \theta_i &= \Delta\theta \times (i - 1) & i = 1, 2, 3, \dots, N \\ \phi_j &= \Delta\phi \times (j - 1) & j = 1, 2, 3, \dots, M \end{aligned} \quad (\text{A.20})$$

If the locations of these directions are plotted on a sphere, one will find them very much more closely spaced near $\theta = 0$ than at $\theta = 90$ deg. We see the same sort of spacing pattern in the intersections of latitude and longitude lines on a global map of the earth. In consequence, (A.15) cannot just be replaced with a double sum, over i and over j . Such an equation would assume equal solid angles for every combination of i and j , whereas the spacing between the directions of these solid angles varies with i , resulting in a de facto weighting of the data that gives more emphasis to directions with small θ . The result is a serious distortion of the true angular distribution of L over the hemispherical solid angle of integration.

One way to get around this problem is to write the finite element version of (A.15) as follows:

$$E = \sum_{j=1}^M \sum_{i=1}^N L_{ij} \cos \theta_i \Delta\omega_{ij} \quad (\text{A.21})$$

and replace the $\Delta\omega_{ij}$ with $\sin \theta_i \Delta\theta \Delta\phi$ to obtain

$$E = \Delta\theta \Delta\phi \sum_{j=1}^M \sum_{i=1}^N L_{ij} \cos \theta_i \sin \theta_i \quad (\text{A.22})$$

where it is seen that the new $\sin \theta_i$ factor weights the sum in such a manner as to compensate for the fact that the directions are more closely spaced near $\theta = 0$. Now the solid angles near $\theta = 0$ are effectively smaller than they are near $\theta = 90$ deg.

In the case that the data for L are not equally spaced over θ and ϕ , the more general version of (A.18) is used:

$$E = \sum_{j=1}^M \sum_{i=1}^N L_{ij} \cos \theta_i \sin \theta_i \Delta\theta_i \Delta\phi_j \quad (\text{A.23})$$

The alert reader will notice that (A.23) resembles (A.16). Equation (A.23) is, in fact, the sums-of-finite-elements version of (A.16).

In the case that N and M are large numbers, on the order of hundreds or thousands, the above sums should produce fairly accurate results without the need for interpolation between adjacent directions or other accuracy enhancing procedures. However, if N and M are limited to fairly small values, errors can be introduced due to the discretization process, portions of which can be overcome with a modification of (A.21).

Let $\Delta\omega_i$ be defined with the following exact equation:

$$\Delta\omega_{ij} = \int_{\phi_j}^{\phi_{j+1}} \int_{\theta_i}^{\theta_{i+1}} \sin \theta d\theta d\phi \quad (\text{A.24})$$

whose integrated value is given by

$$\Delta\omega_{ij} = \Delta\phi (\cos \theta_i - \cos \theta_{i+1}) \quad (\text{A.25})$$

assuming equal interval data in ϕ with $\Delta\phi = 2\pi/M$. Again the size of $\Delta\omega_{ij}$ properly shrinks with decreasing θ . With this change, (A.21) becomes

$$E = \Delta\phi \sum_{j=1}^M \sum_{i=1}^N L_{ij} \cos \theta_i (\cos \theta_i - \cos \theta_{i+1}) \quad (\text{A.26})$$

The principal results for this section are (A.23) and (A.26). When computer programs are written to evaluate these formulas with real data for L_{ij} , a test for bugs in the program can be made by assuming that L is constant in all directions with the value of 1.0. The result of the summations in these equations should therefore be π , as given in (1.19) and (1.20) in Chapter 1.

A program to evaluate the above equations was written and subsequently run for values of N ranging from 50 to 5,000 with the results shown in Table A.1, where $E1$ is obtained using (A.23) and $E2$ comes from using (A.26).

It can be seen from these results that (A.26) ($E2$ in Table A.1) converges slightly faster to the true value than does (A.23) ($E1$ in Table A.1). The convergence is plotted in Figure A.3, and it can be seen that an error of less than 0.5% is achieved with 200 points using (A.26).

Table A.1
Evaluation of (A.23) and (A.26) With $L = 1$

N	50	100	150	200	500	1000	2000	5000
E1	3.2191	3.1803	3.1674	3.1610	3.1493	3.1457	3.1435	3.14237
E2	3.2034	3.1727	3.1624	3.1572	3.1479	3.1447	3.14316	3.14222

Throughout this book extensive use is made of the integral calculus. It is a convenient notation for expressing mathematical definitions and relationships in radiometry and photometry. It is a convenient language to use in solving problems of radiation transfer, as is seen in Chapter 5. However, every integral can be replaced with a finite sum, as shown above. The smaller the intervals, the better the approximation.

With modern high-speed digital computers and a variety of programming languages and easy-to-use programs for performing repeated calculations, many problems can be solved to an acceptable degree of accuracy without ever having to use an integral sign or integrate a function mathematically.

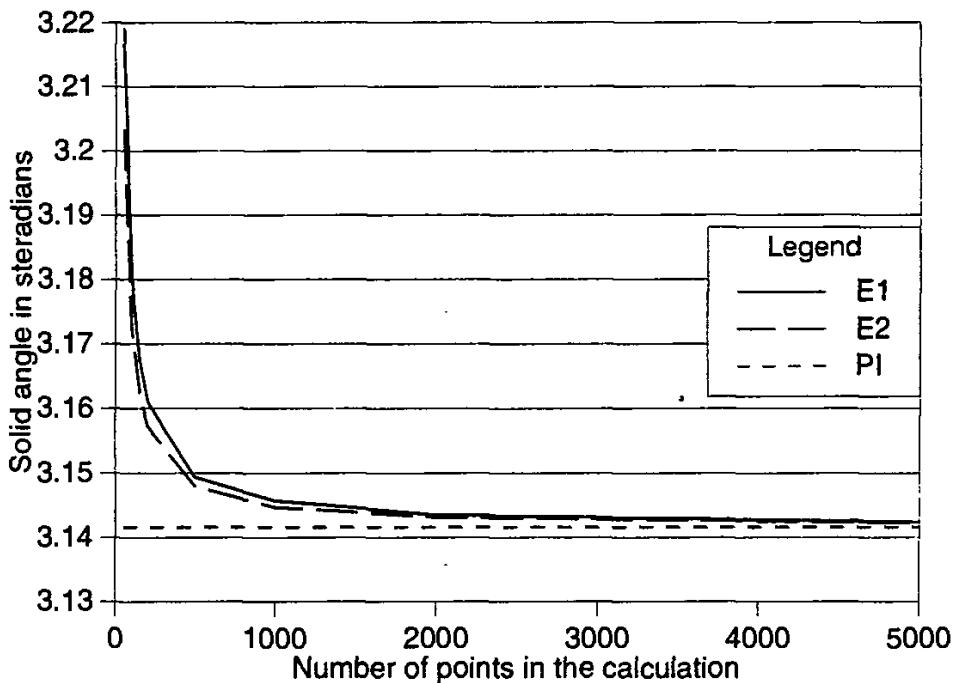


Figure A.3 Plot of (A.18) and (A.22) calculations with increasing values of $N = M$.

REFERENCES

- [1] Weast, R. C., *CRC Handbook of Chemistry and Physics*, 66th ed., Section A, CRC Press, Inc.: Boca Raton, FL, 1985.
- [2] Peirce, B. O., and R. M. Foster. *A Short Table of Integrals*, 4th ed., Ginn and Company: New York, NY, 1957.
- [3] Thompson, S. *Calculus Made Easy*, 3rd ed., Macmillan: London, UK, 1908. St. Martin's Press (subsidiary of Macmillan), New York, 1970.
- [4] Johnson, R. E., and F. L. Kiokemeister. *Calculus with Analytic Geometry*, 2nd ed., Allyn and Bacon: Boston, MA, 1960.
- [5] Stein, S. K. *Calculus with Analytic Geometry*, 4th ed., McGraw-Hill Book Co.: New York, NY, 1987.

About the Author

Dr. Ross McCluney is a principal research scientist at the Florida Solar Energy Center, where his duties include scientific research and program management. His primary research interest is in the energy and illumination performance of windows and other daylighting systems, but he also pursues work in the optical aspects of solar energy collection and analysis. He has a strong interest in environmental ethics and serves as vice president and member of the board of directors of the Earth Ethics Research Group, Inc. of Seminole, Florida. His previous book, on the environmental problems of South Florida, was published by the University of Miami Press in 1971.

Dr. McCluney is past chairman of the Technical Committee on fenestrations of the American Society of Heating, Refrigerating, and Air Conditioning Engineers; a member of the Daylighting Committee of the Illuminating Engineering Society; a member of the U.S. National Committee of the International Lighting Commission (CIE); and a founding member of the CIE's technical committee on international daylight and solar radiation measurements.

Dr. McCluney obtained a bachelor's degree in physics from Rhodes College in Memphis in 1963 and a master's degree in physics from the University of Tennessee in 1967. His research at the University of Tennessee involved the analysis of subharmonic generation in water through observations of changes in laser diffraction patterns produced by the subharmonics. Subsequently, he worked as a development engineer for Eastman Kodak Company in Rochester, New York, where he developed holographic interferometers for change testing of large-aperture optical systems. He used this technique at the University of Miami to develop a ten-pass holographic interferometer for diagnostic testing of electromagnetic plasmas.

Dr. McCluney received his doctorate in physics from the University of Miami in Coral Gables, Florida, in 1973. His dissertation research involved the scattering of light by marine phytoplankton. He worked as an optical oceanographer for NASA's Goddard Space Flight Center in Greenbelt, Maryland, from 1973 through 1975, where his work focused on the remote measurement of ocean color, including

the design and development of instrumentation for measuring the optical properties of seawater and the development of radiative transfer models for the spectral distribution of radiance emerging from natural waters. He has taught an undergraduate course on college algebra and a graduate course on optical oceanography. He began work at the Florida Solar Energy Center in January 1976, soon after it was established by the Florida legislature.

McCluney has served as a technical consultant to numerous U.S. corporations and government agencies in the United States and Canada. He served as the technical consultant on the design of the world's largest sundial which is located in an administrative office building of Disney World at Lake Buena Vista, Florida.

Index

- Aberrations, 243, 254–63
astigmatism, 259–60
chromatic, 257–58
coma, 258–59
correcting, 261
defined, 254
diffraction limit and, 261–63
distortion, 258
field curvature, 261
image quality and, 264
spherical, 254–57
- Absorptivity, spectral, 129
- Absolute radiometry, 322–26
passive/active modes, 324
self-calibration, 325
uses, 323–24
- Absorptance
defined, 128
equality of average, 151
spectral, 148
- Acceptance angle, 117
decreasing, 123
- Acceptor energy, 202
- Achromats, 258
- Active cavity radiometers, 322
illustrated, 323
- Additive filtering, 226
- Albedo, 160
bond, 160
normal, 160
- American Society for the Testing of Materials (ASTM), 133
- Angles of incidence, 112
- Angular dependence, 156–60
- Antireflection coating, 294
- Aperture stop, 256
- Array devices. *See* CCD array detector
- Asphere, 257
- Astigmatism, 259
illustrated, 260
- Attenuation coefficient, 143
- Band gap, 199
- Bands, 197
conduction, 198, 199
valence, 197–98, 199
- Basic radiance, 114
- Biconical, 157
transmittance/reflectance, 158
- Bidirectional reflectance distribution function (BRDF), 161
- Bidirectional transmittance distribution function (BTDF), 161
- Bimetallic strip, 190
- Blackbodies
6000K, spectrum, 67
defined, 64, 133
emission from, 72–73
emittance of, 133
exitance spectra for, 66, 67, 68
experimental approximation of, 73–74
illustrated, 74
illuminants, 367–68
radiation, 63–64
color temperature, 71–72
distribution temperature, 71–72
emission, 73
exchanging, 103
luminous efficacy of, 69–71
- See also* Greybodies
- Blazing, 291
- Bougie, 340

- Brewster's angle, 140
- Brightness, 349
- luminance and, 45–47
- Bulk devices, 187
- Burning glass, 251
- Calculus, 377–78
- definition of the derivative, 378–80
 - definition of the integral, 380–82
 - integrals as sums, 382–83
 - sums over solid angles, 383–87
- Calibration
- detector, 227–29, 276
 - methods for, 228–29
 - photometer/radiometer, 314–29
 - approaches, 314–15
 - broadband irradiance standard
 - sources, 318–19 - drift and, 315
 - methods, 306
 - self-calibration, 325
 - standard lamps and, 315–16
 - transfer standards, 316–18
 - using standard sources, 327–28 - spectroradiometer, 328–29
- Carbon black, 189
- CCD array detector, 207
- hybrid, 209
- Charge carriers, 202
- Charge-coupled device (CCD), 207
- See also* CCD array detector
- Chopping wheel, 213, 278
- Chroma, 349
- illustrated, 350
- Chromatic aberrations, 257–58
- defined, 257
 - illustrated, 257
 - longitudinal, 258
- See also* Aberrations
- Chromatic adaptation, 348
- Chromatic difference of magnification, 258
- Chromaticity, 356
- Chromaticity coordinates, 356
- CIE, 4
- 1924 standard photometric observer, 34
 - 1931 color system, 354–59
 - 1964 Color-Difference Formula, 374
 - 1964 supplementary observer color system, 359
 - 1976 ($L^*a^*b^*$) color space, 352
 - 1976 uniform color space, 359–64
- Definitions 129
- International Lighting Vocabulary*, 128, 129, 133, 336, 345
- spectral luminous efficiency, 36
- terminology standards, 1, 2
- Circularly variable interference filter, 287, 314
- Circular polarization, 23
- illustrated, 24
- Coherence, 80
- Collimated beam, 16–17
- Colorimeter, 358
- Colorimetry, reflection, 371–72
- Color matching functions, 354
- plots of, 355
 - tristimulus, 360
 - as weighting functions, 356
- Color rendering, 372–73
- defined, 372
- Color rendering index (CRI), 372–75
- defined, 373
- Colors, 344–49
- concepts of, 345
 - inherent, 366
 - of monochromatic beams, 343
- Color science, 343–75
- Color space, 357
- 1964, with plots, 361
 - 1976, uniform, 362, 363
- Color specification, 349–54
- attributes, 349
- Color spectral distribution, 347
- Color stimulus, 347, 348
- achromatic stimuli, 348
 - metameric, 348
 - primary, 354
- Color stimulus function, 347
- tristimulus values, 348
- Color temperature, 71, 364–65
- applied to illuminant, 364
 - correlated, 71, 364
 - of the source, 71
 - standardized, 364–65
- Coma, 258–59
- illustrated, 260
- Compound parabolic concentrator, 268
- illustrated, 269
- Concentration ratio
- defined, 122
 - determination, 121
- Concentrators
- compound parabolic, 268–69
 - nonimaging, 268

-
- Conduction band, 198
 illustrated, 199
- Conductors, 130–31, 201
 relative energy band position, 200
- Cone receptors, 345–46
- Configuration factor, 92, 100–102
 alternate names for, 100
 with alternative notation, 101–2
 classes, 101
 differential, 101
 expression, 100
 geometry, 105
- Conical angle, 157
 geometry, 157
- Contrast
 detection, 48
 threshold, 49
- Conversion factors
 for depreciated units, 339
 SI to IP, 338
- Cosine response correction, 221–24, 274
 with integrating sphere, 223
 method for, 221
- Critical angle, 139
- Dark current, 180
- Dark output, 180
- Daylight illuminants, 369–71
- Depletion zone, 204
- Depth of field, 256
- Detectivity, 218
 normalized, 218
 illustrated, 230
- See also* Quantum efficiency; Responsivity
- Detectors, 8
 calibration of, 227–29, 276
 methods for, 228–29
- CCD array, 207
 classification of, 187–209
 dark output, 180
 defined, 179
 detectivity, 218
 illustrated, 230
 determining signals from, 183
 example, 229–334
 flux conditioning prior to, 220–27
 flux falling on, 186
 manufacturers of, 240–41
 multi-element, 205–9
 NEP, 217–18
 noise, 209–11
 categories, 210
- electrical, 210
 impact, 211
 radiation, 210
 random, 209
 white, 211
- performance, 215–20
 photon noise-limited, 219–20
- photoconductive, 203, 212
- photoemissive detectors, 191–96
 with pre-amplifiers, 233
- pyroelectric, 190–91
 sensitivity of, 235
- quadrant, 205–6
- quantum efficiency, 216–17
- radiance meter, 119–20
- response function, 181
- responsivity of, 180, 216
- semiconductor, 197–205
 sensitivity of, 180
 photocathode spectral radiant, 231
 pyroelectric, 235
 range of, 229
 of semiconductor types, 232
- signal conditioning after, 227
- signals of, 180
- spectral response, 179, 182
 constant, 185
 flat, 183
 limited, 184
 narrow-band, 186
 non-flat, 185
 spectral sensitivity for, 229
- target area, 31
- temperature and, 186
- thermal, 187–91
- Dielectric, 136
 optical properties, 136–41
- Diffraction
 gratings, 289
 order of, 289
- Diffraction grating monochromator, 290
- Diffraction limit, 261–63
 defined, 262
- Diffusion, 163
 isotropic, 163
 nonselective, 163
- Directional angle, 157
 geometry for, 157
- Directional dependencies, 128, 148, 156
- Directional-hemispherical, 157
 reflectance, 158

-
- Disability glare, 50–52
 avoiding, 55
 defined, 51
 illustrated, 52
- Discomfort glare, 52–53
 defined, 52
 illustrated, 53
- Dispersion, 141, 257
- Distortion, 258
 barrel, 258
 illustrated, 259
 pin cushion, 258
- Distribution temperature, 71–72
- Document on Colorimetry*, 353
- Doping, 202
- Dual-port comparison method, 278–79
 illustrated, 278
- Dynodes, 193
- Electromagnetic radiation, 20–22
 defined, 21
 polarization of, 22–25
 sources, 294–95
 See also Radiation
- Elementary beam, 109–10
 exitant, 111
 illustrated, 109
 incident, 111
 refracted, 112
- Elliptical polarization, 23
- Emissivity, 9
 defined, 132–33
 term usage, 133
- Emittance
 for blackbody, 133
 defined, 132
 directional, 133
 equality of average, 151
 hemispherical, 133
 definition of, 134–35
 material values, 134
 normal, 133
 material values, 135
 spectral
 directional, 136
 tungsten, 135
 total, 135
- Encyclopedia of Chemical Technology*, 197
- Endings
 -ance, 128–29
 usage, 130
 -ivity, 128
- Eppley Laboratory, 309
- Equivalency, 142
- Etendue, 123
- Excitation levels, 198
- Exitance, 9
 of greybodies, 133
 luminous, 38
- Extinction coefficient, 143
- Extraterrestrial solar spectrum, 67
- FEL lamps, 321
- Fermi level, 200, 201
- Field curvature, 261
- Field darkening, 253
- Field emission, 193
- Flux. *See* Photon flux; Radiant flux
- Flux conditioning, 220–27
 cosine response correction, 221–24
 photopic correction, 224
 spectral filtering, 224–27
- Flux density, 20
 area increase and, 91
- Flux distribution, 265–66
- Flux transfer calculations, 87–105
 configuration factor, 100–102
 geometry and definitions, 87–100
 Monte Carlo method, 99–100
 net exchange of radiation, 102–3
- F-number, 119, 246, 299
- Focal length, 245
- Focal plane, 245
- “Gershun tube,” 116, 267
 illustrated, 117
- Glare
 disability, 50–52, 55
 discomfort, 52–53
 illuminance and, 54–55
- Glass, 294
 See also Lens; Windows
- Golay cell, 190
- Goniometers, 295–96
 schematic illustration of, 295
- Gonioradiometer, 295
- Grating monochromator, 289–92
 diffraction, 291
 wavelength scanning
 mechanism, 290
- Gratings
 blazing of, 290
 diffraction, 289–90
 operating equation for, 289

-
- Greybodies, 63
defined, 133
exitance of, 133
See also Blackbodies
- Halogen lamps
FEL, 321
quartz, 76, 320
tungsten, 76
- Handbook of Chemistry and Physics*, 134, 135
- Handbook of Fundamentals*, 133
- Handbook of Optics*, 365
- Haze meters, 279, 297
- Hemispherical angle, 157
geometry, 157
- Hemispherical emittance, 133
definition of, 134–35
material values, 134
- Holes, 202
- “Hot mirrors,” 172
illustrated, 173
- Hue, 349
- IES Lighting Handbook*, 34, 48, 49, 55, 295
color rendering indices, 374
illuminance categories and values, 56
illumination and, 54
Reference Volume, 36, 81
standard illuminants and, 367
- IES Lighting Ready Reference*, 55
- Illuminance, 3, 46
categories/values, 56
defined, 38
glare and, 54–55
luminance and, 54
meters, 311–12
selection, 55–56
spherical, 336
standard sources, 326–27
- Illuminants, 344
blackbody, 367–68
daylight, 369–71
defined, 366
standard, 344, 366–71
- Illumination, 54–56
quality of, 54
quantities of, 54
- Image plane, 248
- Images
quality of, 263–64
aberrations and, 264
virtual, 249–50
- spectral irradiance distributions for, 79
- Insulators, 200
relative energy band position, 200
- Integrating spheres, 223–24
applications, 279
constant radiance/luminance and, 328
efficiency vs. wall reflectance, 273
radiation detector and, 275, 276
for reflectance measurement, 277
- Intensity, 11
defined, 20
luminous, 3, 37–38
photon, 27
radian, 10–11
spectral radiant, 11
- Interfaces, 129
corrugated, 150
multiple reflections between, 152
optical properties, 130–41
patterned, 150
reflectivity at, 137, 139, 140
transmissivity at, 139, 140
- Interference, 81
- Interference filters, 237
angle of incidence and, 285
circularly variable, 287, 314
defined, 314
linearly variable, 287
long pass, 286
multilayer, 284
short pass, 286
spectral passband shape for, 284
technology, 287
See also Spectral filters
- International Commission on Illumination. *See CIE*
- International Committee for Weights and Measures, 336
- International Daylight Measurement Programme (IDMP), 309
- International Energy Agency (IEA), 309
- International Lighting Vocabulary*, 128, 129, 133, 336, 345
- International standard candle, 340
- Inverse square law, 19, 91
- Ionization energy, 199
- IP system of units, 333, 336
SI conversion factors, 338
- Irradiance, 3
angle of incidence vs., 16
collimated beam and, 16
defined, 9, 20

-
- Irradiance (*cont.*)
- importance of, 10
 - leaving surface, 9
 - photon, 27
 - radiance vs., 15–16
 - radiometers, 306–10
 - restricted spectral band, 310–11
 - spectral. *See* Spectral irradiance
 - spherical, 336
- Isotropic diffuser, 163
- perfect, 164
- Kirchhoff's law, 149
- finite wavelengths and, 162
- Lambertian radiators, 13–15
- Lambertian surface, 18
- defined, 164
- Lambert's Cosine Law, 13–15, 91
- geometry, 15
- Lamps, 11
- halogen, 76, 320, 321
 - standard, 315–16
- Lasers
- radiometer power, 310
 - spectral irradiances, 80
- Lateral color, 258
- Lens
- design, 246
 - f-number, 119
 - focal plane, 245
 - glasses, 294
 - idealized (thin), theory, 245–50
 - negative, 249
 - perfect, 248
 - magnification/demagnification with, 251
 - positive, 249
 - speed, 119
 - spherical aberration in, 255
- "Lens formula in Gaussian form," 248
- LI-COR, Inc., 336–37
- Light, 33–36
- defined, 33
 - levels of, adaptation to, 48
 - primary sources of, 346
 - See also* Wavelengths
- Light-emitting diode (LED), 79–80
- spectral exitance of, 80
- Lighting system luminous efficacy, 43–44
- electric, 44
- Linearly variable interference filter, 287
- Linear polarization, 22
- Lorentz-Mie scattering theory, 147
- Luminaires, 11
- Luminance, 3
- brightness and, 45–47
 - defined, 38
 - meters, 311–12
 - spacial distribution of, 46
 - standard sources, 326–27
 - variations, 47
 - vision and, 47–50
- Luminosity, 340
- Luminous efficacy, 41–44
- of blackbody radiation, 69–71
 - on linear scale, 71
 - on logarithmic scale, 70
 - values, 69
 - defined, 43
 - lighting, 43–44
 - from light sources, 43
 - radiation, 41–43
 - from radiation source, 44
 - spectral, function, 42
- Luminous exitance, 38
- Luminous fluence rate, 336
- Luminous flux, 3
- defined, 37
 - useful, 44
- Luminous intensity, 3
- defined, 37–38
- Luxon, 339
- Mesopic vision, 36
- Metal-insulator semiconductor (MIS) capacitors
- Metal junction devices, 187
- Metamerism, 36
- Metamerism indices, 370
- Metamers, 348
- Metric, 333
- principles, 324–36
 - units for radiometry/photometry, 336
- Microchannels, 196
- Microphonics, 211
- Mirrors, 246
- disadvantage of, 246
 - off-axis operation, 247
 - perfect, 247
 - spherical aberration in, 255
- Modulation, 213
- methods, 213–14
 - need for, 214
- Modulation transfer function (MTF), 264
- Monochromatic radiation, 1

-
- Monochromators, 280–93
 grating, 289–92
 prism, 287–89
 scanning, 287–93
 spectral filters, 280–87
 spectrometers, 292–93
- Monte Carlo calculations, 99–100
 setting up, 100
- Multilayer interference filter, 284
- Multiplier phototube (MPT), 194
 schematic, 194
 variations, 196
- Munsell Book of Color*, 351
- Munsell Color System, 349–51
 chroma, 349
 color solid, 351
 illustrated, 350
 schematic, 351
- National Institute of Standards and Technology (NIST), 317–18, 320
- National Standards Laboratories, 329
- NEP. *See* Noise equivalent power
- Nephelometer, 296–97
- Noise, detector, 209–11
 categories, 210
 defined, 209
 electrical, 210
 impact, 211
 radiation, 210
 random, 209
 white, 211
- Noise equivalent power (NEP), 217–18
 wavelength dependencies, 234
- Nonblackbodies, 63
- Normal emittance, 133
 material values, 135
- N-type extrinsic semiconductor, 202
- Object plane, 248
- Optical axis, 244–45
 centered ray and, 244
- Optical concentration
 limits, 120–23
 ratio determination, 121
- Optical materials, refractive indices of, 141
- Optical properties
 broadband, 160–64, 165–67
 bulk medium, 131, 142–48
 conductor, 130–31
 definitions, 168–69
 dielectric, 136–41
- directional dependence of, 127
 incident/emergent beam directions and, 132
- interface, 130–41
- of materials, 127–75
- nonconductor, 131–32
- photopic, 167
- of plane parallel plates, 148–56
- of rough surfaces, 141
- specular, 162–64
- surface, 130–41
- surface emission, 132–36
- Optical systems, 120–23, 243–300
 centered, 244
 diffraction limited, 262
 focal length, 245
 illustrated, 268
 imaging, 243
 properties, 246
 nonimaging, 243, 266–69
 objectives of, 267
 radiance incidence on, 120
 throughput, 269–71
See also Optical axis
- Optical transfer function (OTF), 264
- Oriel Corporation, 310
- Parsons Black, 189
- Path radiance, 115
- Peltier effect, 187
- Pencil of rays, 45, 108–9
 envelope, 108
 exitent, 109
 focusing, onto small point, 160
 illustrated, 109
 incident, 109
- Perfect vacuum
 emission into, 72–73
 refractive index of, 72
- Photocathode, 192
- Photoconductive detectors, 203
 noise spectrum for, 212
- PbS, 205
- Photodiodes, 193
 circuit connection schemes, 204
- PIN, 205
- silicon, 204
- Photoelectron, 193
- Photoemissive detectors, 191–96
 response time, 196
- Photoemissivity, 191
- Photometers, 303–30
 calibration 314 20

-
- Photometers (*cont.*)**
- defined, 303
 - manufacturers, 313
 - spectral response, 304
 - working standard, 315
- See also Radiometers; Spectroradiometers*
- Photometry**, 1, 33–61
- defined, 33
 - definitions, 37–44
 - fundamental quantities, 3, 37
 - representative values, 40
 - metric units for, 336
 - radiative transfer and, 115
 - SI to IP conversion factors for, 338
- See also Radiometry*
- Photomultiplier tube (PMT)**, 194–95
- gain, 196
 - illustrated, 195
- Photon counter**, 26
- Photon flux**, 25–28
- defined, 26
 - units, 336–38
- Photon intensity**, 27
- Photon irradiance**, 27
- Photon number**, 26
- Photon radiance**, 27–28
- Photons**, 25
- Avogadro's number of, 336
 - fluctuation-limited detection, 219
 - incident, 217
 - noise-limited detection, 219
- Photopic correction**, 224
- Photopic optical properties**, 167
- Photopic spectral luminous efficiency**, 35
- function, 34
 - for photopic vision, 42
- Photosynthetic photon flux fluence rate (PPFFR)**, 337
- Phototopic vision**, 34
- photopic spectral luminous efficiency for, 42
- Photovoltaic detectors, silicon**, 232–33
- linearity for, 235
- Photovoltaic effect**, 203
- Pixels**, 174
- Planck's blackbody spectral radiation law**, 64–68
- modification of, 72, 136
- Planck's formula**, 71
- Plane angle**, 5
- illustrated, 5
- Plane parallel plates**, 148
- nonscattering media, 148–54
 - scattering geometry, 155
 - scattering media, 154–56
- Plane polarization**, 21
- term usage, 23
- Point spread function**, 263
- illustrated, 264
- Polarization**, 22–25
- degree of, 25
 - elliptical, 23
 - left circular, 23
 - linear, 22
 - illustrated, 23
 - plane, 21
 - term usage, 23
 - random, 24
 - right circular, 23
- Polarizing angle**, 140
- Polychromator**, 312
- Power**, 8
- equal, 24
- Poynting vector**, 21
- Principle planes**, 246
- Prism monochromator**, 287–89
- illustrated, 288
- P-type extrinsic semiconductor**, 202
- Pyranometers**, 189
- Pyrheliometers**, 189
- Pyroelectric detectors**, 190
- advantages, 191
 - disadvantages, 191
 - materials, 190
 - sensitivity of, 235
- Quadrant detector**, 205–6
- illustrated, 206
 - PbS, 206
- Quantum efficiency**, 216–17
- external, 217
- See also Detectivity; Responsivity*
- Radial projections**, 5
- Radian**, 6
- Radianc**, 3
- in absorbing media, 115
 - along a ray geometry, 108
 - basic, 114
 - defined, 11
 - emitted, 134
 - factor, 164
 - as field quantity, 107–8
 - function of, 12
 - geometry for, 12

- irradiance vs., 15–16
- path, 115
- photon, 27–28
- radiometers, 306–10
- in scattering media, 115
- spectral, 13
 - transmitted-to-incident, 148–49
 - at two parallel planes, 143
- through a lens, 114
- transmittance, 114
- Radiance invariance, 111
 - defined, 111
 - at interface, 112–14
- Radiance meter
 - concentrating, 116–20
 - detector, 119–20
 - diagram, 118
 - Gershun tube, 116
 - illustrated, 117
 - lensless, 116
 - performance analysis, 118
- Radiant efficiency, 44
- Radiant energy, 7
 - spectral, 8
- Radiant fluence rate, 336
- Radiant flux, 3
 - defined, 8
 - invisible, 43
 - spectral, 8
 - spectral distribution, 9
- Radiant intensity, 3, 10–11
 - defined, 10
 - spectral, 11
- Radiant power. *See* Radiant flux
- Radiation
 - blackbody, 63–64
 - color temperature, 71–72
 - distribution temperature, 71–72
 - emission, 73
 - exchanging, 103
 - luminous efficacy of, 69–71
 - chopping, 211–15
 - defined, 1
 - detection, 179–237
 - electromagnetic, 20–22
 - defined, 21
 - polarization of, 22–25
 - sources, 294–95
 - elementary beam of, 109–10
 - illustrated, 109
 - exchange
- for nonparallel planar source/receiver, 97
- luminous efficacy, 41–43
- monochromatic, 1
 - in visible spectrum, 36
- noise, 210
- reflected, 99
- spectral distribution of, 3
- wavelengths, 33
- Radiative transfer, 115
 - defined, 115
 - Monte Carlo and, 100
 - problem solutions, 103–4
- Radiometers, 303–30
 - absolute, 322–26
 - active cavity, 322–23
 - broadband, 303, 306–10
 - calibration, 314–29
 - methods, 306
 - using standard sources, 327–28
 - defined, 303
 - design factors, 305–6
 - illuminance, 311–12
 - irradiance
 - broadband, 306–10
 - restricted spectral band, 310–11
 - laser power, 310
 - luminance, 311–12
 - Mims, 80
 - radiance, broadband, 306–10
 - solid angle range in, 305
 - working standard, 315
 - See also* Photometers; Spectroradiometers
- Radiometry
 - absolute, 322–26
 - defined, 1
 - fundamental quantities, 3, 37
 - representative values for, 13
 - metric units for, 336
 - radiative transfer and, 115
 - SI to IP conversion factors for, 338
 - SI units for, 337
 - See also* Photometry
- Radiosity, 340
- Rayleigh scattering, 75
 - theory, 147
- Rays
 - in elementary beam, 111
 - increasing radiance of, 115
 - parallel, 245
 - in pencil of rays, 111
- Real image, 249

- Reflectance**
- absolute, 279
 - average, 272
 - biconical, 158
 - contrast, 49
 - defined, 128
 - diffuse, 163, 164
 - directional-hemispherical, 158
 - equations, 160–62
 - factor, 164
 - high wall, 274
 - integrating sphere, measurement of, 277
 - measurement, 156
 - for plate of glass, 154
 - regular, 263
 - solid angles and, 156–57
 - spectral, 131
 - defined, 148
 - specular, 162–63, 164
 - term usage, 129
- See also* Transmittance
- Reflection colorimetry**, 371–72
- Reflection/transmission**
- diffuse, 163
 - isotropic diffuse, 163–64
 - mixed, 163
 - perfect, 164
 - regular, 163
 - specular, 163
- See also* Reflectance; Transmittance
- Reflectivity**, 129
- defined, 129
 - at interface, 137, 139, 140
- Reflectometers**, 274–80, 296
- Refraction**, at smooth surface, 112
- Resistance**, 129
- Resistivity**, 129
- Responsivity**, 216
- See also* Detectivity; Quantum efficiency
- Retroreflection**, 164
- Reviews of Scientific Instruments*, 304
- Saturation**, 349
- Scanning monochromators**, 287–93
- grating, 289–92
 - prism, 287–89
 - spectrometers, 292–93
- Scattering**, 163
- Scattering indicatrix**, 145
- Scattering meters**, 296–97
- Scotopic vision**, 34
- Semiconductor devices**, 197–205
- conductors, 201
 - insulators, 200
- Semiconductors**, 201–5
- defined, 201
 - intrinsic, 202
 - N-type extrinsic, 202
 - P-type extrinsic, 202
 - relative energy band position, 200
- See also* Semiconductor devices
- Sensitivity**, 180
- photocathode spectral radiant, 231
 - of pyroelectric motion sensing detector, 235
 - ranges of, 229
 - of semiconductor detector types, 232
- Sensors**, 8
- defined, 179
- See also* Detectors
- Signal modulation**, 211–15
- Silicon photodiodes**, 204
- Silicon photovoltaic detectors**, 232–33
- linearity for, 235
 - spectral response of, 233
- Single-port substitution method**, 277
- SI (Système Internationale) system of units**, 333, 334–36
- coherent set of units, 335
 - IP conversion factors, 338
 - metric principles, 334–36
 - metric units for radiometry/photometry, 336
 - prefixes, 335
 - for radiometry/photometry, 337
- See also* IP system of units
- Snell's law**, 137, 162
- Society of Photo-Optical Instrumentation Engineers (SPIE)*, 304–5
- Solar constant**, 123
- Solid angle**, 6
- conical, 157
 - with constant flux, 91
 - defining element geometry, 17
 - directional, 157
 - direction and magnitude of, 132
 - hemispherical, 17, 157
 - illustrated, 6
 - infinitesimal element of, 132
 - pencil of rays, 45
 - projected, 271
 - sum over, 383–87
 - weighted, 271
- Source/receiver**

-
- flux transfer calculations, 87–105
 - geometry, 88
 - infinitesimal, areas, 101
 - nonparallel planar, 97
 - Speckles, 81
 - Spectral absorptance, 148
 - Spectral absorption coefficient, 144
 - Spectral absorptivity, 129
 - Spectral dependence, 164–65
 - Spectral distribution, 75
 - of blue sky radiation, 75
 - CIE vocabulary for, 2
 - color, 347
 - defined, 1
 - determining, 166
 - of flux falling on detector, 186
 - of flux in standard illuminant, 370
 - of incidence irradiance, 58
 - tungsten filament incandescent lamp, 173
 - Spectral emittance
 - directional, 136
 - tungsten, 135
 - Spectral filtering, 224–27
 - additive, 226
 - subtractive, 226
 - Spectral filters, 280–85
 - circularly variable, 284
 - holographic notch, 283–84
 - IR blocking glass, 284
 - linearly variable, 284
 - long-pass glass, 283
 - manufacturers of, 281
 - multilayer interference, 282
 - order-sorting, 291
 - spectral width at half maximum, 285
 - UV-VIS bandpass glass, 282
 - See also* Interference filters
 - Spectral irradiance, 10
 - of halogen lamps, 76
 - illustrated, 77
 - of infrared sources, 79
 - for laboratory radiation sources, 78
 - for lasers, 80
 - of terrestrial direct beam solar radiation, 75
 - transmitted-to-incident, 148
 - of xenon gaseous discharge lamp, 78
 - Spectral lines, 292
 - Spectral masking, 226–27
 - Spectral radiance, 13
 - transmitted-to-incident, 148–49
 - at two parallel planes, 143
 - Spectral radiant flux, 8
 - Spectral radiant intensity, 11
 - Spectral reflectance
 - defined, 148
 - of metallic mirror coatings, 131
 - Spectral scattering coefficient, 144
 - Spectral selectivity, 167–74
 - defined, 167
 - uses, 172, 174
 - Spectral transmissivity, 129
 - Spectral transmittance, 148
 - Spectral weighting functions, 150
 - Spectrograph, 289, 292, 312
 - Spectrometers, 292–93
 - defined, 292
 - Spectrophotometer, 293, 312
 - Spectroradiometers, 215, 292, 303, 312–14
 - attachments, 312
 - calibration of, 328–29
 - defined, 312
 - scanning, 312
 - See also* Photometers; Radiometers
 - Spectrum, 1
 - eye response, 169
 - See also* Spectral distribution
 - Spectrum locus, 358
 - Specular, 137
 - reflectance, 162–63, 164
 - transmittance, 162–63, 164
 - Spherical aberrations, 254–57
 - correcting, 257
 - defined, 254
 - in lens, 255
 - in mirror, 255
 - reducing, 256
 - See also* Aberrations
 - Spot diagram, 265
 - illustrated, 266
 - Standard lamps, 315
 - illustrated use, 316
 - Standard observer, 354
 - Stefan-Boltzmann law, 72
 - modification, 133
 - Steradian, 6
 - Subtractive filtering, 226
 - Surfaces
 - diffusely reflecting, 50
 - glossy, 50
 - interreflections between, 99
 - light reflection off, 51
 - matte, 50

-
- Surfaces (*cont.*)**
- net flux transferred between, 103
 - specularly reflecting, 50
- Target area, 31**
- Thermal detectors, 187–91**
- bulk, 190
 - Peltier effect and, 187–88
 - See also* Detectors
- Thermocouples, 188**
- Thermopile, 188**
- coatings, 89
 - compensated, 188
 - connection schematic, 189
 - pyranometer, 189
 - pyrheliometers, 189
 - thin film, 188–89
- Threshold contrast, 49**
- Throughput, 123, 269–71**
- defined, 270
 - geometric construction for, 270
- Total internal reflection (TIR), 139**
- Transfer standards, 316–17**
- primary, 316
 - secondary, 316–17
 - working standards, 317
- Transmissivity**
- at interface, 139, 140
 - spectral, 129
- Transmissometers, 274–80, 296**
- Transmittance**
- biconical, 158
 - defined, 128
 - diffuse, 163, 164
 - directional-hemispherical, 159
 - equations, 160–62
 - hot mirror, 173
 - measurement, 156
 - for plate of glass, 154
 - solid angles and, 156–57
 - spectral, 148
 - specular, 162–63, 164
 - window, spectra, 170
 - See also* Reflectance
- Trichromatic system, 347**
- primaries of, 347
- Tristimulus**
- 1964 color matching function, 360
 - colorimetry, 352–54
 - color space, 358
 - values, 348, 356
- Troland, 338–39**
- Turbidity, 297**
- Ulbricht sphere, 223**
- Ultraviolet, 33**
- photons, 76
- Unit circle, 5**
- USA Standard Nomenclature and Definitions for Illuminating Engineering, 3–4**
- UV-A, 79**
- UV-B, 79**
- UV-C, 79**
- Vacuum phototube, 192**
- schematic, 192
- Valance band, 197–98**
- illustrated, 199
- Vignetting, 253–54**
- illustrated, 253
- Virtual image, 249**
- illustrated, 250
- Visibility reference function, 50**
- Vision**
- adaptation, 48
 - luminance and, 47–50
- V-lambda correction, 328**
- V-lambda curve, 304, 355**
- V-lambda matching spectral response correcting filter, 224**
- Volume scattering, 129**
- Volume scattering function, 144–45**
- definition geometry, 146
- Wavelengths**
- notions, 4–7
 - radiation, 33
 - range of, 4
- Wien displacement law, 68–69**
- Windows, 293–94**
- for cold climates, 170
 - glasses used for, 294
 - for hot climates, 172
 - transmittance spectra, 170, 171
 - conical-hemispherical, 172
- World Meteorological Organization (WMO), 309**

The Artech House Optoelectronics Library

Brian Culshaw, Alan Rogers, and Henry Taylor, *Series Editors*

Acousto-Optic Signal Processing: Fundamentals and Applications, Pankaj Das

Amorphous and Microcrystalline Semiconductor Devices, Optoelectronic Devices,
Jerzy Kanicki, editor

Electro-Optical Systems Performance Modeling, Gary Waldman and John Wootton

The Fiber-Optic Gyroscope, Hervé Lefèvre

Field Theory of Acousto-Optic Signal Processing Devices, Craig Scott

High-Power Optically Activated Solid-State Switches, Arye Rosen and Fred Zutavern,
editors

Highly Coherent Semiconductor Lasers, Motoichi Ohtsu

Germanate Glasses: Structure, Spectroscopy, and Properties, Alfred Margaryan and
Michael A. Piliavin

Introduction to Electro-Optical Imaging and Tracking Systems, Khalil Seyrafi and
S. A. Hovanessian

Introduction to Glass Integrated Optics, S. Iraj Najafi

Introduction to Radiometry and Photometry, William Ross McCluney

Optical Control of Microwave Devices, Rainee N. Simons

Optical Document Security, Rudolf L. van Renesse

Optical Fiber Amplifiers: Design and System Applications, Anders Bjarklev

Optical Fiber Sensors, Volume I: Principles and Components, John Dakin and Brian
Culshaw, editors

Optical Fiber Sensors, Volume II: Systems and Applications, John Dakin and Brian
Culshaw, editors

Optical Interconnection: Foundations and Applications, Christopher Tocci and H. John
Caulfield

Optical Network Theory, Yitzhak Weissman

Optical Transmission for the Subscriber Loop, Norio Kashima

Principles of Modern Optical Systems, Volumes I and II, I. Andonovic and
D. Uttamchandani, editors

Reliability and Degradation of LEDs and Semiconductor Lasers, Mitsuo Fukuda

Semiconductors for Solar Cells, Hans Joachim Möller

Single-Mode Optical Fiber Measurements: Characterization and Sensing, Giovanni Cancellieri

For further information on these and other Artech House titles, contact:

Artech House
685 Canton Street
Norwood, MA 01602
617-769-9750
Fax: 617-769-6334
Telex: 951-659
email: artech@world.std.com

Artech House
Portland House, Stag Place
London SW1E 5XA England
+44 (0) 71-973-8077
Fax: +44 (0) 71-630-0166
Telex: 951-659