Aberration Theory Made Simple

Second Edition

Tutorial Texts Series

- *Aberration Theory Made Simple, Second Edition,* Virendra N. Mahajan, Vol. TT93
- *Modeling the Imaging Chain of Digital Cameras,* Robert D. Fiete, Vol. TT92
- *Cells Illuminated:* In Vivo *Optical Imaging*, Lubov Brovko, Vol. TT91
- *Polarization of Light with Applications in Optical Fibers*, Arun Kumar, Ajoy Ghatak, Vol. TT90
- *Digital Fourier Optics: A MATLAB Tutorial*, David G. Voeltz, Vol. TT89
- *Optical Design of Microscopes*, George Seward, Vol. TT88
- *Analysis and Evaluation of Sampled Imaging Systems*, Richard H. Vollmerhausen, Donald A. Reago, Ronald Driggers, Vol. TT87
- *Nanotechnology: A Crash Course*, Raúl J. Martin-Palma and Akhlesh Lakhtakia, Vol. TT86
- *Direct Detection LADAR Systems*, Richard Richmond, Stephen Cain, Vol. TT85
- *Optical Design: Applying the Fundamentals*, Max J. Riedl, Vol. TT84
- *Infrared Optics and Zoom Lenses, Second Edition*, Allen Mann, Vol. TT83
- *Optical Engineering Fundamentals, Second Edition*, Bruce H. Walker, Vol. TT82
- *Fundamentals of Polarimetric Remote Sensing*, John Schott, Vol. TT81
- *The Design of Plastic Optical Systems*, Michael P. Schaub, Vol. TT80
- *Fundamentals of Photonics*, Chandra Roychoudhuri, Vol. TT79
- *Radiation Thermometry: Fundamentals and Applications in the Petrochemical Industry*, Peter Saunders, Vol. TT78
- *Matrix Methods for Optical Layout*, Gerhard Kloos, Vol. TT77
- *Fundamentals of Infrared Detector Materials*, Michael A. Kinch, Vol. TT76
- *Practical Applications of Infrared Thermal Sensing and Imaging Equipment, Third Edition*, Herbert Kaplan, Vol. TT75
- *Bioluminescence for Food and Environmental Microbiological Safety*, Lubov Brovko, Vol. TT74
- *Introduction to Image Stabilization*, Scott W. Teare, Sergio R. Restaino, Vol. TT73
- *Logic-based Nonlinear Image Processing*, Stephen Marshall, Vol. TT72
- *The Physics and Engineering of Solid State Lasers*, Yehoshua Kalisky, Vol. TT71
- *Thermal Infrared Characterization of Ground Targets and Backgrounds, Second Edition*, Pieter A. Jacobs, Vol. TT70
- *Introduction to Confocal Fluorescence Microscopy*, Michiel Müller, Vol. TT69
- *Artificial Neural Networks: An Introduction*, Kevin L. Priddy and Paul E. Keller, Vol. TT68
- *Basics of Code Division Multiple Access (CDMA),* Raghuveer Rao and Sohail Dianat, Vol. TT67
- *Optical Imaging in Projection Microlithography*, Alfred Kwok-Kit Wong, Vol. TT66
- *Metrics for High-Quality Specular Surfaces*, Lionel R. Baker, Vol. TT65
- *Field Mathematics for Electromagnetics, Photonics, and Materials Science*, Bernard Maxum, Vol. TT64
- *High-Fidelity Medical Imaging Displays*, Aldo Badano, Michael J. Flynn, and Jerzy Kanicki, Vol. TT63
- *Diffractive Optics–Design, Fabrication, and Test*, Donald C. O'Shea, Thomas J. Suleski, Alan D. Kathman, and Dennis W. Prather, Vol. TT62
- *Fourier-Transform Spectroscopy Instrumentation Engineering*, Vidi Saptari, Vol. TT61
- *The Power- and Energy-Handling Capability of Optical Materials, Components, and Systems*, Roger M. Wood, Vol. TT60
- *Hands-on Morphological Image Processing*, Edward R. Dougherty, Roberto A. Lotufo, Vol. TT59
- *Integrated Optomechanical Analysis*, Keith B. Doyle, Victor L. Genberg, Gregory J. Michels, Vol. TT58
- *Thin-Film Design: Modulated Thickness and Other Stopband Design Methods*, Bruce Perilloux, Vol. TT57
- *Optische Grundlagen für Infrarotsysteme*, Max J. Riedl, Vol. TT56
- *An Engineering Introduction to Biotechnology*, J. Patrick Fitch, Vol. TT55
- *Image Performance in CRT Displays*, Kenneth Compton, Vol. TT54
- *Introduction to Laser Diode-Pumped Solid State Lasers*, Richard Scheps, Vol. TT53

(For a complete list of Tutorial Texts, see http://spie.org/x651.xml.)

Aberration Theory Made Simple

Second Edition

Virendra N. Mahajan

Tutorial Texts in Optical Engineering
Volume TT93

SPIE
PRESS
Bellingham, Washington USA

Library of Congress Cataloging-in-Publication Data

Mahajan, Virendra N.
Aberration theory made simple / Virendra N. Mahajan. -- 2nd ed.
p. cm. -- (Tutorial texts series ; v. TT93)
Includes bibliographical references and index.
ISBN 978-0-8194-8825-1
1. Aberration. 2. Imaging systems. 3. Geometrical optics. I. Title.
QC671.M34 2011
621.36--dc23

2011025097

Published by

SPIE
P.O. Box 10
Bellingham, Washington 98227-0010 USA
Phone: +1 360.676.3290
Fax: +1 360.647.1445
Email: books@spie.org
Web: http://spie.org

First printing

Printed in the United States of America.

Introduction to the Series

Since its inception in 1989, the Tutorial Texts (TT) series has grown to cover many diverse fields of science and engineering. The initial idea for the series was to make material presented in SPIE short courses available to those who could not attend and to provide a reference text for those who could. Thus, many of the texts in this series are generated by augmenting course notes with descriptive text that further illuminates the subject. In this way, the TT becomes an excellent stand-alone reference that finds a much wider audience than only short course attendees.

Tutorial Texts have grown in popularity and in the scope of material covered since 1989. They no longer necessarily stem from short courses; rather, they are often generated independently by experts in the field. They are popular because they provide a ready reference to those wishing to learn about emerging technologies or the latest information within their field. The topics within the series have grown from the initial areas of geometrical optics, optical detectors, and image processing to include the emerging fields of nanotechnology, biomedical optics, fiber optics, and laser technologies. Authors contributing to the TT series are instructed to provide introductory material so that those new to the field may use the book as a starting point to get a basic grasp of the material. It is hoped that some readers may develop sufficient interest to take a short course by the author or pursue further research in more advanced books to delve deeper into the subject.

The books in this series are distinguished from other technical monographs and textbooks in the way in which the material is presented. In keeping with the tutorial nature of the series, there is an emphasis on the use of graphical and illustrative material to better elucidate basic and advanced concepts. There is also heavy use of tabular reference data and numerous examples to further explain the concepts presented. The publishing time for the books is kept to a minimum so that the books will be as timely and up-to-date as possible. Furthermore, these introductory books are competitively priced compared to more traditional books on the same subject.

When a proposal for a text is received, each proposal is evaluated to determine the relevance of the proposed topic. This initial reviewing process has been very helpful to authors in identifying, early in the writing process, the need for additional material or other changes in approach that would serve to strengthen the text. Once a manuscript is completed, it is peer reviewed to ensure that chapters communicate accurately the essential ingredients of the science and technologies under discussion.

It is my goal to maintain the style and quality of books in the series and to further expand the topic areas to include new emerging fields as they become of interest to our reading audience.

James A. Harrington
Rutgers University

To My

Wife **Shashi Prabha**

Son **Vinit Bharati**

Daughter **Sangita Bharati**

FOREWORD TO THE FIRST EDITION

It is a distinct pleasure for me to write this short foreword to Dr. Virendra Mahajan's tutorial text, *Aberration Theory Made Simple.* I write it not because I am particularly knowledgeable about aberration theory—in fact, it may be because I am not particularly knowledgeable that I was invited! This is a Tutorial Text, and as a lifelong educator I am also a lifelong learner and I should be able to learn from this text; and I did.

This text is prepared in the ideal way for a tutorial. It comes as a direct result of teaching this material to a wide range of audiences in a wide range of locations; so it has been tried and tested. The "student guinea pigs" have performed their invaluable service so that those of us who come along later have the benefit of their and the author's labors.

Dr. Mahajan has lived up to his title and made aberration theory simple. Of course, I should caution the reader that simple is relative. Some topics do not yield easily to simple yet accurate descriptions. Those readers who insist that "rays" are the most important components of any analysis of optical systems, whether aberrant or not, will be very satisfied with the first half of the book, but may wish to ignore the second half. They should not. Those who are enamored with the wave approach (like me) will immediately read the second half of this book and applaud, but not go back and read the first half. They should! I did!

I am pleased that Dr. Mahajan has provided a significant list of references in addition to the bibliography at the end of the book. This will be of considerable value to the reader. Not incidentally, SPIE Optical Engineering Press will also publish a Milestone volume on *Effects of Aberrations in Imaging Systems* with Dr. Mahajan as the editor. Thus, each of us will be able to have an authoritative companion volume that contains reprints from the world's literature that will no doubt verify that this current Tutorial Text is indeed *Aberration Theory Made Simple.*

Brian J. Thompson June 1991
Rochester, New York

TABLE OF CONTENTS

ABERRATION THEORY MADE SIMPLE

PREFACE TO THE SECOND EDITION

I wrote *Aberration Theory Made Simple* some 20 years ago to provide a clear, concise, and consistent exposition of what aberrations are, how they arise in optical imaging systems, and how they affect the quality of optical images formed by them, both in terms of geometrical and diffraction optics. Later, I expanded this Tutorial Text into a textbook under the title *Optical Imaging and Aberrations* in two parts, one on Ray Geometrical Optics and the other on Wave Diffraction Optics. Detailed mathematical derivations missing in the Tutorial Text are given in this textbook, along with problems at the end of each chapter.

In this second edition of *Aberration Theory Made Simple*, I have updated the sign convention for Gaussian optics to the Cartesian sign convention, as used in advanced books on geometrical optics and in the optical design software programs. The quantities such as object and image distances that are numerically negative are indicated in figures with a parenthetical negative sign (–). Thus a reader will find a change in the sign of some parameters in equations in the part on geometrical optics when compared with those in the first edition. In this new edition, I have deleted certain advanced details that are available in the long textbook. Deletions include the plots of the optical transfer function for primary aberrations. I have added some new material as well, such as the centroid and standard deviation of ray aberrations, spot diagrams for primary aberrations, golden rule of optical design about relying on such diagrams, update of 2D PSFs for primary aberrations, aberration-free optical transfer function of systems with annular and Gaussian pupils, Zernike polynomials for circular pupils and the corresponding polynomials for annular and Gaussian pupils, effect of longitudinal image motion on an image, lucky imaging in ground-based astronomy, and adaptive optics. I have also added a brief summary at the end of each chapter, highlighting the essence of its content. It is hoped that these additions will be helpful to the readers of this edition of *Aberration Theory Made Simple*.

The second edition of *Aberration Theory Made Simple* has been translated into Russian by Professor Irina Livshits of National Research University of Information Technologies, Mechanics and Optics, Saint Petersburg, Russia. This Russian edition is available from the university by contacting her at <ecenter-optica10@yandex.ru>.

Virendra N. Mahajan
El Segundo, California

June 2011

PREFACE TO THE FIRST EDITION

Aberration theory is a subject that is as old and fascinating as the field of optics. It is, however, a cumbersome subject that many students of optics do not appreciate fully. The purpose of this tutorial book is to provide a clear, concise, and consistent exposition of what aberrations are, how they arise in optical imaging systems, and how they affect the quality of images formed by them. Its emphasis is on physical insight, problem solving, and numerical results. It is intended for engineers and scientists who have a need and/or a desire for a deeper and better understanding of aberrations and their role in optical imaging and wave propagation. Although some knowledge of Gaussian optics and an appreciation for aberrations would be useful, they are not prerequisites. What is needed is dedication and perseverance. A novice trying to learn this subject without investing much time will probably be disappointed in spite of the title of the book. The book is not intended for teaching lens design or optical testing. However, it is hoped that those working in these fields will benefit from it. It should be useful to students who may want to learn aberration theory without having to go through any lengthy derivations.

These derivations are omitted out of necessity for brevity and in keeping with the spirit of these tutorials. These tutorials have been adapted from my lectures for a graduate course entitled “Advanced Geometrical Optics," which I have been teaching in the Electrical Engineering-Electrophysics Department of the University of Southern California since 1984. They were originally developed for a short course on optical imaging and aberrations, which I taught at The Aerospace Corporation to Aerospace and Air Force personnel. They were then expanded for a short course I have been teaching at the Optical Society of America and SPIE meetings. Generally speaking, only the primary aberrations of optical systems are discussed here; they provide the first and a significant step beyond Gaussian imaging. Although a knowledge of these aberrations is very useful, they may not sufficiently describe the imaging properties of a high-quality optical system. Higher-order aberrations in such systems are often determined by ray tracing them.

This book is organized in two parts: Part I is on ray geometrical optics and Part II is on wave diffraction optics. The first chapter introduces the concepts of aperture stop and entrance and exit pupils of an optical imaging system. The wave and ray aberrations are defined and wavefront defocus and tilt aberrations are discussed. Various forms of the primary aberration function of a rotationally symmetric system are given, and how this function changes as the aperture stop of the system is moved from one position to another is discussed. The aberration function for the simplest imaging system, namely, a single spherical refracting surface, is given. Finally, a procedure by which the aberration function of a multielement system may be calculated is described. This chapter provides a foundation for the next six chapters.

Chapters 2-6 give the primary aberrations of simple systems, such as a thin lens, plane-parallel plate, spherical mirror, Schmidt camera, and a conic mirror. Numerical problems are discussed here and there to illustrate how to apply the formulas given in

these chapters. Part I of the book ends with chapter 7, where the aberrated images of a point object based on geometrical optics are discussed. Thus the ray spot diagrams and, in particular, the spot sizes for primary aberrations are discussed. The concept of aberration balancing, based on geometrical optics to reduce the size of an image spot, is introduced.

In Part II, chapters 8-11 discuss the effects of aberrations on the image of a point object based on wave diffraction optics. Chapter 8 considers systems with circular exit pupils. The aberration-free characteristics of such systems are described in terms of the point-spread and optical transfer function. How the aberrations affect these functions is discussed, and aberration tolerances are obtained for a given Strehl or a Hopkins ratio. The concept of aberration balancing, based on wave diffraction optics, to maximize Strehl or Hopkins ratios is discussed. Systems with annular and Gaussian pupils are considered in chapter 9. The effect of obscuration on the point-spread function and on aberration tolerance is discussed. Similarly, the effect of Gaussian amplitude at the exit pupil is discussed. The content of this chapter provides a basis for assessing the effects of aberrations on the optical performance of reflecting telescopes, such as Cassegrain and Ritchey-Chrétien, and on the propagation of laser beams.

The line of sight of an aberrated system is discussed in chapter 10 in terms of the centroid of its point-spread function. It is pointed out that only coma type aberrations change the centroid. Random aberrations are considered in chapter 11, where the time averaged point-spread and optical transfer functions for random image motion and aberrations introduced by atmospheric turbulence are discussed. Part II of the book ends with chapter 12, where a brief discussion is given on how the aberrations of a system may be observed and recognized interferometrically.

Each chapter is written to be as independent of the others as possible, although some are more so than others. For example, chapter 7 may be followed by chapter 1. Except for the first few sections of chapter 1, it is not necessary to understand Part I in order to understand Part II. However, reading Part II without Part I would be like knowing half of a story. Chapter 12 may be read at any time; however, the reason for using certain specific values of defocus, for example, in the case of spherical aberration, may not be understood unless the concepts of aberration balancing discussed in chapters 7 and 8 are understood. On the matter of references to the literature on aberration theory, I have listed under the bibliography those books that treat this subject to some or a large extent. These are the ones I have had the opportunity to read and benefit from. On the wave diffraction optics, I have given references in the text either for historical reasons (such as the papers by Airy and Lord Rayleigh) or because the work is relatively recent and has not appeared in books. Additional references are given after the bibliography for further study on part of the reader.

Finally, I would like to thank those who have helped me with the preparation of this book. I have had many discussions with Dr. Bill Swantner on geometrical optics and Dr. Richard Boucher on diffraction optics. Dr. Boucher also did computer simulations of the

point-spread functions and interferograms and prepared the photographs for this book. Prof. Don O'Shea provided critical and valuable comments when he reviewed this book. Helpful comments were also provided by Prof. R. Shannon. The Sanskrit verse and its translation on p. xxiii were provided by Dr. S. Sutherland, University of California at Berkeley. The manuscript and its many revisions were typed by Iva Moore. The final version was produced by Betty Wenker and Candy Worshum. I thank The Aerospace Corporation for providing help and facilities to prepare this book. I also thank Dr. Roy Potter and Eric Pepper of the SPIE staff for suggesting and facilitating the preparation of this book, which was carefully edited by Rick Hermann. I cannot thank my wife and children enough for their patience during the course of this work and so I dedicate this book to them.

Virendra N. Mahajan June 1991
El Segundo, California

SYMBOLS AND NOTATION

a	radius of exit pupil	q	shape factor
A_i	aberration coefficient	r_c	radius of circle
AS	aperture stop	r_0	atmospheric coherence diameter
CR	chief ray	R	radius of reference
B_d	peak defocus value	$R_n^m(\rho)$	Zernike circle radial polynomial
B_t	peak tilt value	$R_n^m(\rho;\epsilon)$	Zernike annular radial polynomial
D	pupil diameter	$R_n^m(\rho;\gamma)$	Zernike-Gauss radial polynomial
e	eccentricity	S	Strehl ratio
f	focal length	S_p	area of exit pupil
F	focal ratio	W	wave aberration
GR	general ray	$x,\ y$	rectangular coordinates of a point
h	object height	z	optical axis, axial distance
h'	image height	$Z_n^m(\rho,\theta)$	Zernike circle polynomial
H	Hopkins ratio	$Z_n^m(\rho,\theta;\epsilon)$	Zernike annular polynomial
I	irradiance	$Z_n^m(\rho,\theta;\gamma)$	Zernike-Gauss circle polynomial
J_n	nth order Bessel function	$\vec{v}_i$	image spatial frequency vector
l	distance	$\vec{v}_o$	object spatial frequency vector
L	image distance from exit pupil	v	normalized spatial frequency
LSF	line spread function	τ	optical transfer function
m	pupil magnification	Ψ	phase transfer function
M	image magnification	$\rho = r/a$	normalized radial coordinate
MCF	mutual coherence function	ω	Gaussian beam radius
MR	marginal ray	β	field angle
MTF	modulation transfer function	ϕ	polar angle of frequency vector
n	refractive index, integer	ϵ	obscuration ratio
N	Fresnel number	γ	truncation ratio
OTF	optical transfer function	Δ	longitudinal defocus
p	position factor	Φ	phase aberration
P	object point, image power	r,θ	polar coordinates of a point
P'	Gaussian image point	λ	optical wavelength
P_{ex}	power in the exit pupil	ξ,η	normalized rectangular coordinates
$P(\cdot)$	pupil function	σ_W	standard deviation of wave aberration
PSF	point-spread function	σ_Φ	standard deviation of phase aberration
PTF	phase transfer function	σ_F	standard deviation of figure errors

अनन्तरत्नप्रभवस्य यस्य हिमं न सौभाग्यविलोपि जातम् ।
एको हि दोषो गुणसन्निपाते निमज्जतीन्दोः किरणेष्विवाङ्कः ॥

Anantaratnaprabhavasya yasya himaṃ na saubhāgyavilopi jātam |
Eko hi doṣo guṇasannipāte nimajjatīndoḥ kiraṇesvivāṅkaḥ ||

The snow does not diminish the beauty of the Himālayan mountains which are the source of countless gems. Indeed, one flaw is lost among a host of virtues, as the moon's dark spot is lost among its rays.

Kālidāsa *Kumārasambhava* 1.3

CHAPTER 1
Optical Aberrations

1.1 INTRODUCTION

This chapter starts with the concepts of *aperture stop* and *entrance* and *exit pupils* of an optical imaging system. Certain special rays, such as the *chief* and the *marginal*, are defined. The *wave aberration* associated with a ray is defined and its relationship to the corresponding transverse *ray aberration* is given. Representations of wavefront *defocus* and *tilt* aberrations are given. We introduce different forms of the *primary aberration function* of a *rotationally symmetric system.* How this function changes as the aperture stop of the system is moved from one position to another is discussed. The primary aberration function for the simplest imaging system, namely, a single spherical refracting surface, is given for an arbitrary position of the aperture stop. Finally, we outline a procedure by which the aberration function of a multielement system may be calculated. This procedure is utilized in later chapters, for example, to calculate the aberration of a *thin lens* (Chapter 2) and *a plane-parallel plate* (Chapter 3). This chapter forms the basis of Part I on *geometrical optics.*

1.2 OPTICAL IMAGING

An optical imaging system consists of a series of refracting and/or reflecting surfaces. The surfaces refract or reflect light rays from an object to form its image. The image obtained according to geometrical optics in the *Gaussian approximation*, i.e., according to Snell's law in which the sines of the angles are replaced by the angles, is called the *Gaussian image.* The Gaussian approximation and the Gaussian image are often referred to as the *paraxial approximation* and the *paraxial image*, respectively. We assume that the surfaces are rotationally symmetric about a common axis called the *optical axis* (*OA*). Figure 1-1 illustrates the imaging of an on-axis point object P_0 and an off-axis point object P, respectively, by an optical system consisting of two thin lenses. (For definition of a thin lens, see Section 2.2.) P' and P_0' are the corresponding Gaussian image points. An object and its image are called *conjugates* of each other, i.e., if one of the two conjugates is an object, the other is its image.

An aperture in the system that physically limits the solid angle of the rays from a point object the most is called the *aperture stop* (*AS*). For an extended (i.e., a nonpoint) object, it is customary to consider the aperture stop as the limiting aperture for the axial point object, and to determine vignetting, or blocking of some rays, by this stop for off-axis object points. The object is assumed to be placed to the left of the system so that initially light travels from left to right. The image of the stop by surfaces that precede it in the sense of light propagation, i.e., by surfaces that lie between it and the object, is called the *entrance pupil* (*EnP*). When observed from the object side, the entrance pupil appears to limit the rays entering the system to form the image of the object. Similarly, the image of the aperture stop by surfaces that follow it, i.e., by surfaces that lie between it and the

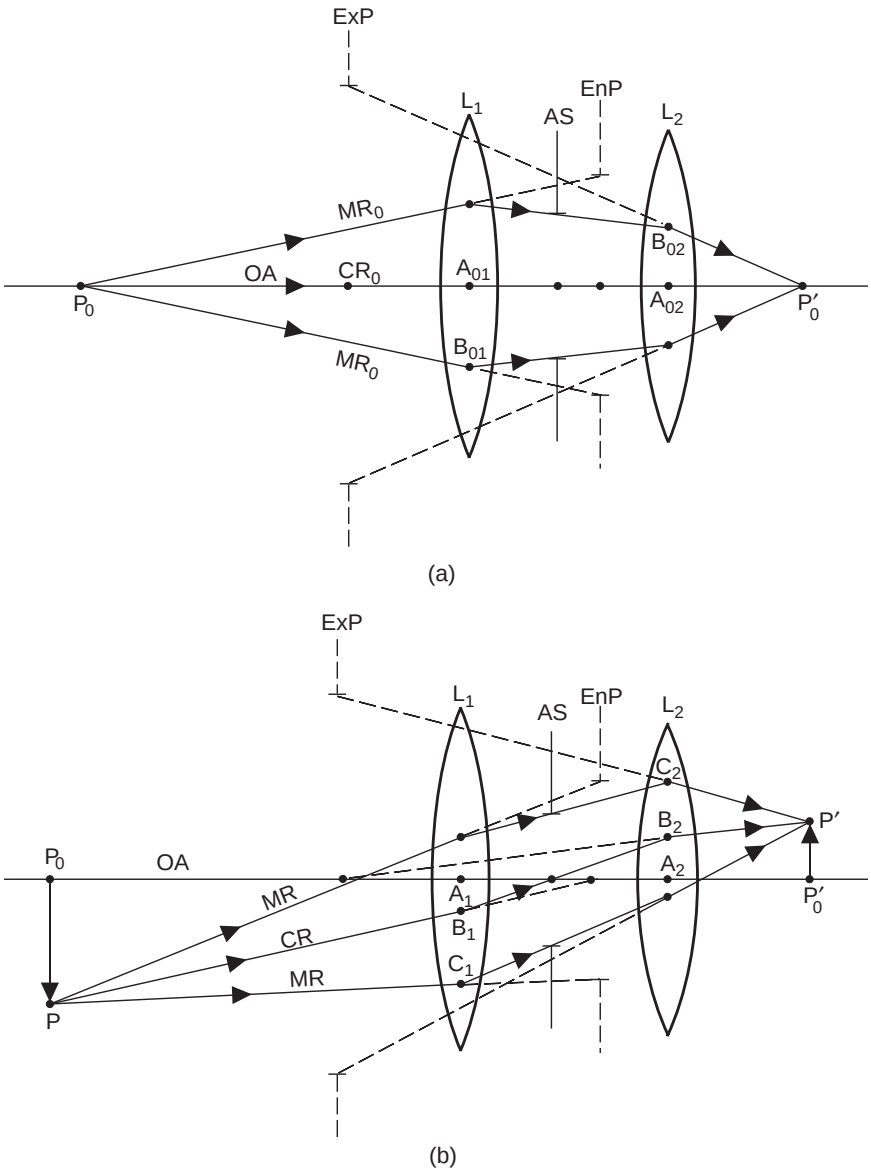


Figure 1-1. (a) Imaging of an on-axis point object P_0 by an optical imaging system consisting of two lenses L_1 and L_2. *OA* is the optical axis. The Gaussian image is at P_0'. *AS* is the aperture stop; its image by L_1 is the entrance pupil *EnP*, and its image by L_2 is the exit pupil *ExP*. CR_0 is the axial chief ray, and MR_0 is the axial marginal ray. (b) Imaging of an off-axis point object *P*. The Gaussian image is at *P′*. *CR* is the off-axis chief ray, *MR* is the off-axis marginal ray.

image, is called the *exit pupil* (*ExP*). The object rays reaching its image appear to be limited by the exit pupil. Since the entrance and exit pupils are images of the stop by the surfaces that precede and follow it, respectively, the two pupils are conjugates of each other for the whole system; i.e., if one pupil is considered as the object, the other is its image formed by the system.

An object ray passing through the center of the aperture stop and appearing to pass through the centers of the entrance and exit pupils is called the *chief* (or the *principal) ray* (*CR*). An object ray passing through the edge of the aperture stop is called a *marginal ray* (*MR*). The rays lying between the center and the edge of the aperture, and, therefore, appearing to lie between the center and edge of the entrance and exit pupils, are called *zonal rays.*

It is possible that the stop of a system may also be its entrance and/or exit pupil. For example, a stop placed to the left of a lens is also its entrance pupil. Similarly, a stop placed to the right of a lens is also its exit pupil. Finally, a stop placed at a single thin lens is both its entrance and exit pupils.

1.3 WAVE AND RAY ABERRATIONS

In this section, we define the wave aberration associated with a ray and relate it to its transverse ray aberration in an image plane. The *optical path length* of a ray in a medium of refractive index *n* is equal to *n* times its geometrical path length. If rays from a point object are traced through the system and up to the exit pupil such that each one travels an optical path length equal to that of the chief ray, the surface passing through their end points is called the system *wavefront* for the point object under consideration. If the wavefront is spherical with its center of curvature at the Gaussian image point, we say that the Gaussian image is perfect. If, however, the wavefront deviates from this *Gaussian spherical wavefront*, we say that the Gaussian image is aberrated. The optical deviation (i.e., geometrical deviations times the refractive index *n* of the image space) of the wavefront along a certain ray from the Gaussian spherical wavefront is called the *wave aberration* of that ray. It represents the difference between the optical path lengths of the ray under consideration and the chief ray in traveling from the point object to the reference sphere. Accordingly, the wave aberration associated with the chief ray is zero. The wave aberration associated with a ray is positive if it has to travel an extra optical path length, compared to the chief ray, in order to reach the Gaussian spherical wavefront. The Gaussian spherical wavefront is also called the *Gaussian reference sphere.*

Figures 1-2a and 1-2b illustrate the reference sphere *S* and the aberrated wavefront *W* for on- and off-axis point objects whose Gaussian images lie at P_0' and P', respectively. The coordinate system is also illustrated in these figures. We choose a right-hand coordinate system such that the optical axis lies along the *z* axis. The object, entrance pupil, exit pupil, and the Gaussian image lie in mutually parallel planes that are perpendicular to this axis, with their origins lying along the axis. We assume that a point object such as *P* lies along the x axis. The *zx* plane containing the point object and the optical axis is called the *tangential* or the *meridional plane.* The Gaussian image P' lying in the Gaussian image plane along its x axis also lies in the tangential plane. This may be seen by a consideration of a tangential object ray and Snell's law according to which the incident and refracted or reflected rays at a surface lie in the same plane. The chief ray

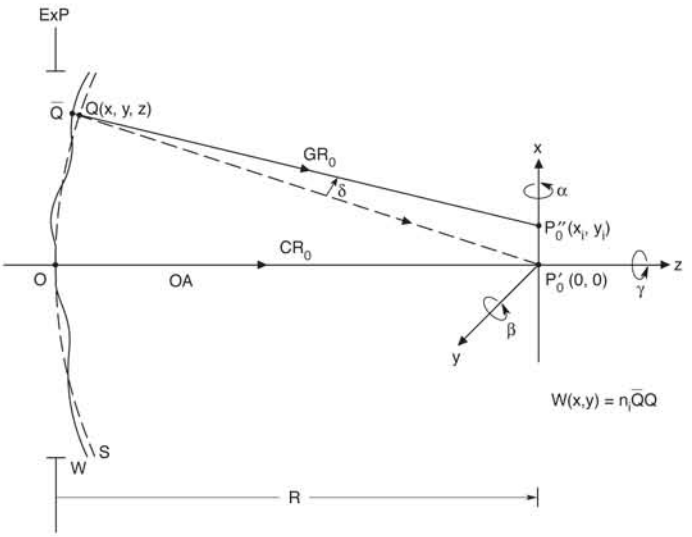


Figure 1-2a. Aberrated wavefront for an on-axis point object P_0. The reference sphere S of radius of curvature R is centered at the Gaussian image point P_0'. The wavefront W and reference sphere pass through the center O of the exit pupil ExP. A right-hand Cartesian coordinate system showing x, y, and z axes is illustrated, where the z axis is along the optical axis of the imaging system. Angular rotations α, β, and γ about the three axes are also indicated. CR_0 is the chief ray, and a general ray GR_0 is shown intersecting the Gaussian image plane at P_0''.

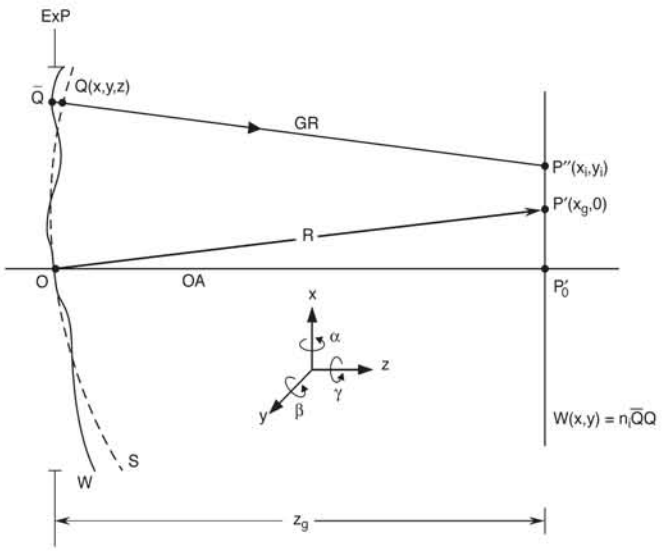


Figure 1-2b. Aberrated wavefront for an off-axis point object P. The reference sphere S of radius of curvature R is centered at the Gaussian image point P'. The value of R in this figure is slightly larger than its value in Figure 1-2a. GR is a general ray intersecting the Gaussian image plane at the point P''. By definition, the chief ray (not shown) passes through O, but it may or may not pass through P'.

always lies in the tangential plane. The plane normal to the tangential plane but containing the chief ray is called the *sagittal plane.* As the chief ray bends when it is refracted or reflected by an optical surface, so does the sagittal plane.

Consider an image ray such as GR in Figure 1-2b passing through a point Q with coordinates (x, y, z) on the reference sphere of radius of curvature R centered at the image point. We let $W(x, y)$ represent its wave aberration $n\overline{Q}Q$, since z is related to x and y by virtue of Q being on the reference sphere. It can be shown that the ray intersects the Gaussian image plane at a point P'' whose coordinates with respect to the Gaussian image point P' are approximately given by

$$(x_i, y_i) = \frac{R}{n}\left(\frac{\partial W}{\partial x}, \frac{\partial W}{\partial y}\right) \quad . \tag{1-1}$$

[Equation (1-1) has been derived by Mahajan, Born and Wolf, and Welford. However, Welford uses a sign convention for the wave aberration that is opposite to ours.]

The displacement $P_0'P_0''$ in Figure 1-2a (or $P'P''$ in Figure 1-2b) of a ray from the Gaussian image point is called its *geometrical* or *transverse ray aberration*, and its coordinates (x_i, y_i) in the Gaussian image plane relative to the Gaussian image point are called its ray aberration components. Since a ray is normal to a wavefront, the ray aberration depends on the shape of the wavefront and, therefore, on its geometrical path difference from the reference sphere. The division of W by n in Eq. (1-1) converts the optical path length difference into geometrical path length difference. When an image is formed in free space, as is often the case in practice, then $n = 1$. The angle $\delta \simeq P_0'P_0''/R$ between the ideal ray QP_0' and the actual ray QP_0'' is called the *angular ray aberration.*

The distribution of rays from a point object in an image plane is called the *ray spot.* (Such diagrams are discussed in Chapter 7.) When the wavefront is spherical with its center of curvature at the Gaussian image point, then the wave and ray aberrations are zero. In that case, all of the object rays transmitted by the system pass through the Gaussian image point, and the image is *perfect.* We shall refer to $W(x, y)$ as the wave at a projected point (x, y) in the plane of the exit pupil. If (r, θ) represent the corresponding polar coordinates, they are related to the rectangular coordinates according to

$$(x, y) = r(\cos\theta, \sin\theta) \quad . \tag{1-2}$$

1.4 DEFOCUS ABERRATION

We now discuss defocus wave aberration of a system and relate it to its longitudinal defocus. Consider an imaging system for which the Gaussian image of a point object is located at P_1. As indicated in Figure 1-3, let the wavefront for this point object be spherical with a center of curvature at P_2 (due to field curvature discussed in Section 1.6) such that P_2 lies on the line OP, joining the center 0 of the exit pupil and the Gaussian image point P_1'. The aberration of the wavefront with respect to the Gaussian reference

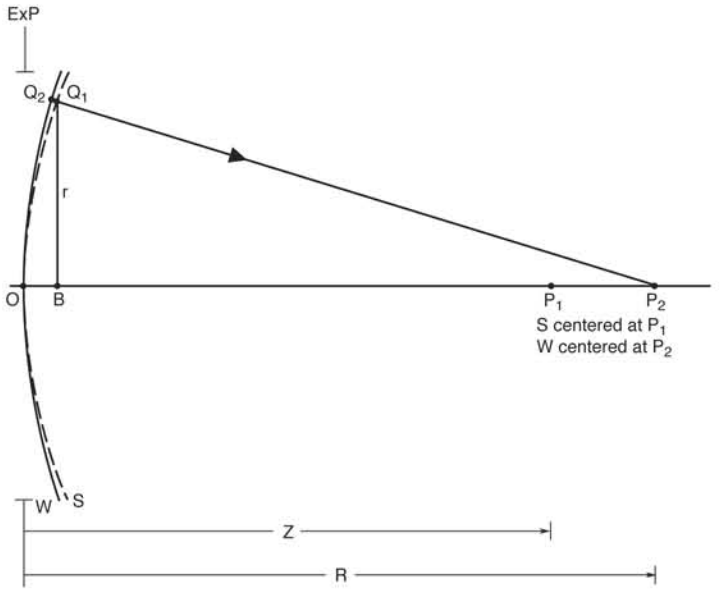


Figure 1-3. Wavefront defocus. Defocused wavefront *W* is spherical with a radius of curvature *R* centered at P_2. The reference sphere *S* with a radius of curvature *z* is centered at P_1. Both *W* and *S* pass through the center *O* of the exit pupil *ExP*. The ray Q_2P_2 is normal to the wavefront at Q_2. *OB* represents the sag of Q_1.

sphere is its optical deviation from it. This deviation is given by nQ_2Q_1, where *n* is the refractive index of the image space and Q_2Q_1, as indicated in the figure, is approximately equal to the difference in the sags of the reference sphere and the wavefront at a height *r*. (The *sag* of a surface at a certain point on it represents its deviation at that point along its axis of symmetry from a plane surface that is tangent to it at its vertex). Thus, the *defocus wave aberration* at a point Q_1 at a distance *r* from the optical axis is given by

$$W(r) = \frac{n}{2}\left(\frac{1}{z} - \frac{1}{R}\right)r^2 \quad , \tag{1-3a}$$

z and R are the radii of curvature of the reference sphere *S* and the spherical wavefront *W* centered at P_1 and P_2, respectively, passing through the center *O* of the exit pupil, and *r* is the distance of Q_1 from the optical axis. We note that the defocus wave aberration is proportional to r^2. If $z \simeq R$, then Eq. (1-3a) may be written

$$W(r) \simeq -\frac{n}{2}\frac{\Delta}{R^2}r^2 \quad , \tag{1-3b}$$

where $\Delta = z - R$ is called the *longitudinal defocus*. We note that the defocus wave aberration and the longitudinal defocus have numerically opposite signs. The ray aberrations corresponding to a defocus wave aberration are discussed in Chapter 7.

A defocus aberration is also introduced if the image is observed in a plane other than the Gaussian image plane. Consider, for example, an imaging system forming an

aberration-free image at the Gaussian image point P_2 (note that the Gaussian image is now located at P_2). Thus, the wavefront at the exit pupil is spherical passing through its center Q with its center of curvature at P_2. Let the image be observed in a defocused plane passing through a point P_1, which lies on the line joining Q and P_2. For the observed image at P_1 to be aberration free, the wavefront at the exit pupil must be spherical with its center of curvature at P_1. Such a wavefront forms the reference sphere with respect to which the aberration of the actual wavefront must be defined. The aberration of the wavefront at a point Q_1 on the reference sphere is given by Eqs. (1-3a) and (1-3b).

If the exit pupil is circular with a radius a, then Eq. (1-3b) may be written

$$W(\rho) = B_d \rho^2 \quad , \tag{1-3c}$$

where $\rho = r/a$ is the normalized distance of a point in the plane of the pupil from its center and

$$B_d \simeq -n\Delta/8F^2 \tag{1-3d}$$

represents the peak value of the defocus aberration with $F = R/2a$ as the *focal ratio* or the *f-number* of the image-forming light cone. Note that a positive value of B_d implies a negative value of Δ. Thus, an imaging system having a positive value of defocus aberration Δ can be made defocus free if the image is observed in a plane lying farther from the plane of the exit pupil, compared to the defocused image plane, by a distance $8B_dF^2/n$. Similarly, a positive defocus aberration of $B_d \simeq -n\Delta/8F^2$ is introduced into the system if the image is observed in a plane lying closer to the plane of the exit pupil, compared to the defocus-free image plane, by a distance Δ.

1.5 WAVEFRONT TILT

Now we describe the relationship between a wavefront tilt and the corresponding tilt aberration. As indicated in Figure 1-4, consider a spherical wavefront centered at P_2 in the Gaussian image plane passing through the Gaussian image point P_1. The wave aberration of the wavefront at Q_1 is its optical deviation nQ_2Q_1 from a reference sphere centered at P_1. It is evident that, for small values of the ray aberration P_1P_2, the wavefront and the reference sphere are tilted with respect to each other by an angle β. The wavefront tilt may be due to distortion discussed in Section 1.6 and/or due to an inadvertently tilted element of the imaging system. The ray and the wave aberrations can be written

$$x_i = R\beta \tag{1-4}$$

and

$$W(r,\theta) = n\beta r\cos\theta \quad , \tag{1-5a}$$

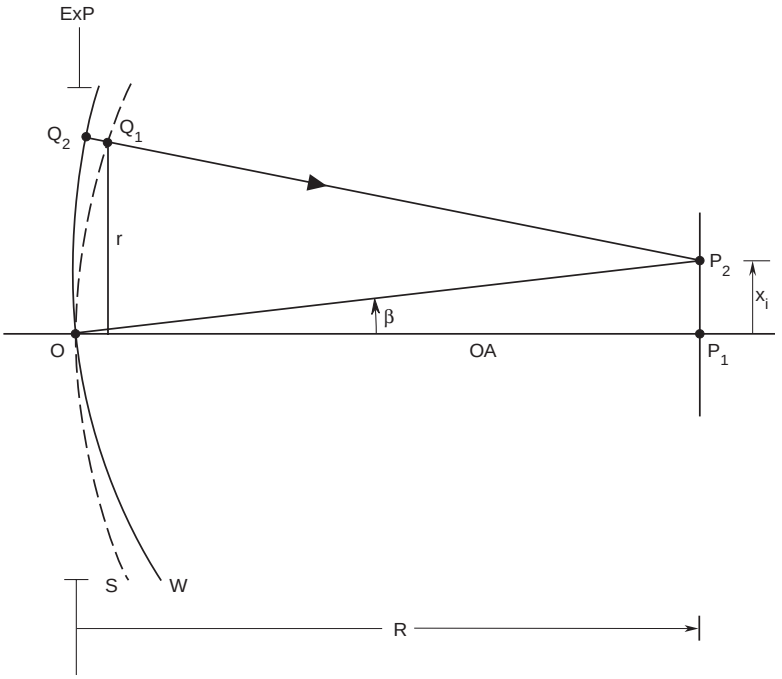


Figure 1-4. Wavefront tilt. The spherical wavefront *W* is centered at P_2 while the reference sphere *S* is centered at P_1. Thus, for small values of P_1P_2, the two spherical surfaces are tilted with respect to each other by a small angle $\beta = P_1P_2/R$, where *R* is their radius of curvature. The ray Q_2P_2 is normal to the wavefront at Q_2.

respectively, where $P_1P_2 = x_i$ and (r, θ) are the polar coordinates of the point Q_1. Both the wave and ray aberrations are numerically positive in Figure 1-4.

Once again, for a system with a circular exit pupil of radius *a*, Eq. (1-5) may be written

$$W(\rho, \theta) = na\beta\rho\cos\theta \quad , \tag{1-5b}$$

or

$$W(\rho, \theta) = B_t\rho\cos\theta \quad , \tag{1-5c}$$

where $B_t = na\beta$ is the peak value of the tilt aberration. Note that a positive value of B_t implies that the wavefront tilt angle β is also positive. Thus, if an aberration-free wavefront is centered at P_2, then an observation with respect to P_1 as the origin implies that we have introduced a tilt aberration of $B_t\rho\cos\theta$.

1.6 ABERRATION FUNCTION OF A ROTATIONALLY SYMMETRIC SYSTEM

The aberration function $W(r, \theta; h')$ of an optical imaging system with an axis of rotational symmetry depends on the object height *h* or the image height *h'* from the

optical axis, and pupil coordinates (r, θ) of a point in the plane of the exit pupil, through three *rotational invariants* h'^2, r^2, and $h'r\cos\theta$. The aberration terms of degree 4 in the rectangular coordinates of the object and pupil points are called *primary aberrations.* Thus, the *primary aberration function* consists of a sum of five terms, e.g.,

$$W(r, \theta; h') = {}_0a_{40}r^4 + {}_1a_{31}h'r^3\cos\theta + {}_2a_{22}h'^2r^2\cos^2\theta + {}_2a_{20}h'^2r^2 + {}_3a_{11}h'^3r\cos\theta \quad , \tag{1-6}$$

where the subscripts of the aberration coefficients ${}_ia_{jk}$ represent the powers of h', r, and $\cos\theta$, respectively. Note that there is no term in h'^4 since the aberration of the chief ray ($r = 0$) must be zero. Since the wave aberration W has dimensions of length, the dimensions of the coefficients ${}_ia_{jk}$ are inverse length cubed. The *order* of an aberration term is equal to the sum of the powers of h' and r, i.e., it is equal to its degree in the (x, y) coordinates of the object (or its image) and pupil points. Since the order of a primary aberration is 4, they are called the *fourth-order wave aberrations.* They are also called the *Seidel aberrations.* Since the ray aberrations are related to the wave aberrations by a spatial derivative [see Eq. (1-1)], their degree is lower by one. Accordingly, the primary aberrations are also called the *third-order ray aberrations.* The coefficients ${}_0a_{40}$, ${}_1a_{31}$, ${}_2a_{22}$, ${}_2a_{20}$, and ${}_3a_{11}$ represent the coefficients of *spherical aberration*, *coma*, *astigmatism*, *field curvature*, and *distortion*, respectively.

From Eq. (1-6), we note that only spherical aberration is independent of the object or image height. The field curvature, in its dependence on the pupil coordinates (r, θ), is like the defocus aberration discussed in Section 1.4. However, the field curvature represents a defocus aberration that depends on the field h', thus requiring a curved image surface for its elimination. On the other hand, pure defocus aberration, such as that produced by observing the image in a plane other than the Gaussian image plane, is independent of the field h'. Similarly, distortion depends on the pupil coordinates as a wavefront tilt. However, distortion depends on the field as h'^3, but the wavefront tilt produced by a tilted element in the system would be independent of h'.

For simplicity, we will use the notation a_s, a_c, a_a, a_d (d for defocus), and a_t (t for tilt) to represent the coefficients of spherical aberration, coma, astigmatism, field curvature, and distortion, respectively. For an optical system with a circular pupil of radius a, we can use the normalized radial variable $\rho = r/a$, suppress the explicit dependence on image height h', and write the primary aberration function in the form

$$W(\rho, \theta) = A_s\rho^4 + A_c\rho^3\cos\theta + A_a\rho^2\cos^2\theta + A_d\rho^2 + A_t\rho\cos\theta \quad , \tag{1-7}$$

where A_i are the *peak aberration coefficients* given by

$$A_s = a_s a^4, \; A_c = a_c h' a^3, A_a = a_a h'^2 a^2, A_d = a_d h'^2 a^2, \; A_t = a_t h'^3 a \quad . \tag{1-8}$$

It should be clear that, since $0 \le \rho \le 1$ and $0 \le \theta \le 2\pi$, a peak aberration coefficient A_i, as the name implies, represents the maximum value of the corresponding aberration. This

value occurs at a point on the edge of the pupil, i.e., for a marginal ray with $\theta = 0$. With the image height suppressed, the field curvature and distortion coefficients A_d and A_t are similar to the corresponding defocus and tilt coefficients B_d and B_t, respectively, considered in Sections 1.4 and 1.5.

1.7 EFFECT OF CHANGE IN APERTURE STOP POSITION ON THE ABERRATION FUNCTION

Now we consider how the primary aberration function of a system changes due to a change in the position of its aperture stop. We remind ourselves that the wave aberration associated with a ray represents the difference between its optical path length and that of the chief ray in traveling from a point object to the reference sphere. Moreover, the chief ray is that object ray that passes through the center of the aperture stop. Hence, since the chief ray changes as the position of the aperture stop is changed, the wave aberration of a ray also changes.

Consider, as indicated in Figure 1-5, an optical imaging system forming an image P' of an off-axis point object, at a height h' from the optical axis. Let the aperture stop of the system be located at a position such that its exit pupil is located at ExP_1. Let the primary aberration function of the system be given by $W_{Q1}(x_1, y_1; h')$ representing the aberration of an image-forming ray passing through a point (x, y) in the plane of the exit pupil with respect to the chief ray O_1P' passing through the center O_1 of the exit pupil.

Now, suppose we move the aperture stop to a new position along the optical axis such that the corresponding new exit pupil is located at ExP_2 with its center at O_2. A

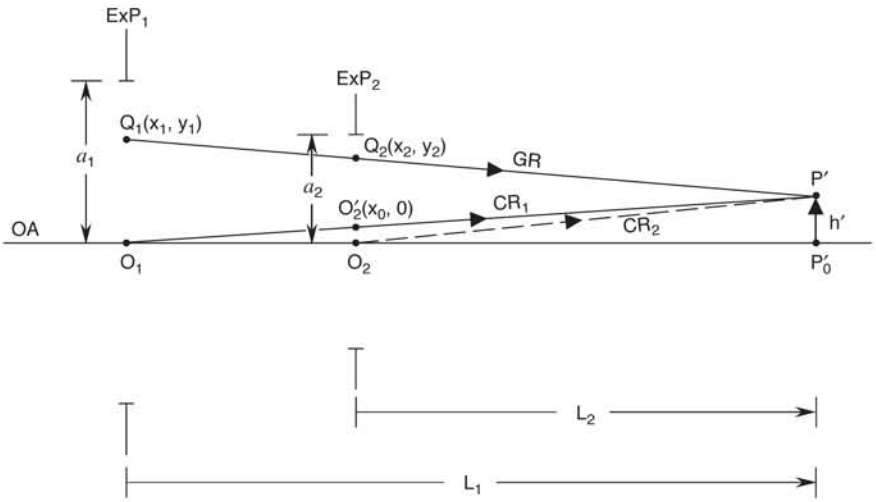


Figure 1-5. Exit pupils ExP_1 and ExP_2 corresponding to two positions of the aperture stop of an optical system forming a Gaussian image P' of an off-axis point object P (not shown). The chief rays CR_1 and CR_2 are for the pupils ExP_1 and ExP_2, respectively.

change in the stop position does not change the position of the image P'. Let L_1 and L_2 be the axial distances of the Gaussian image plane from the planes of the exit pupils ExP_1 and ExP_2, respectively. The chief ray O_1P' (or its extension) intersects the plane of exit pupil ExP_2 at O_2' with rectangular coordinates $(x_0, 0)$, where from similar triangles $O_1O_2O_2'$ and $O_1P_0'P'$ one finds that

$$x_0 = \frac{h'}{L_1}(L_1 - L_2) \quad . \tag{1-9}$$

Note that the *y* coordinate is zero because it lies on the chief ray, which in turn lies in the tangential plane *zx*.

The aberration of a ray Q_1P' with respect to the chief ray O_1P' represents the aberration at a point Q_1 with respect to the aberration at O_1 (which is zero by definition). It is also equal to the aberration of the ray Q_1P' at Q_2 with respect to the aberration at O_2', where Q_2 represents the point of intersection of the ray with the plane of the exit pupil ExP_2. It is evident from the geometry of Figure 1-5 that

$$(x_1, y_1) \simeq \frac{L_1}{L_2}(x_2 - x_0, y_2) \quad , \tag{1-10}$$

where (x_2, y_2) are the coordinates of Q_2 with respect to O_2 as the origin. Thus, the aberration at Q_2 with respect to its value at O_2' may be obtained by substituting Eq. (1-10) into the expression for $W_{Q_1}(x_1, y_1)$, i.e.,

$$W_{Q_2}(x_2, y_2) \simeq W_{Q_1}\left[\frac{L_1}{L_2}(x_2 - x_0, y_2)\right] \quad . \tag{1-11))}$$

Note that the aberration function referred to the new exit pupil is zero at $(x_0, 0)$. In order that the aberration at the center O_2 of the new exit pupil be zero, we define a new aberration function $W(x_2, y_2; h')$ with respect to the new chief ray O_2P' (not shown in Figure 1-5), i.e.,

$$\begin{aligned} W(x_2, y_2; h') &= W_{Q_2}(x_2, y_2) - W_{Q_2}(0, 0) \\ &= W_{Q_1}\left[\frac{L_1}{L_2}(x_2 - x_0, y_2)\right] - W_{Q_1}(-x_0 L_1/L_2, 0) \quad . \end{aligned} \tag{1-12}$$

Let the primary aberration function at ExP_1 be given by

$$\begin{aligned} W_{Q_1}(x_1, y_1; h') &= a_{s1}(x_1^2 + y_1^2)^2 + a_{c1}h'x_1(x_1^2 + y_1^2) + a_{a1}h'^2x_1^2 \\ &\quad + a_{d1}h'^2(x_1^2 + y_1^2) + a_{t1}h'^3x_1 \quad . \end{aligned} \tag{1-13}$$

Substituting Eq. (1-13) into Eq. (1-12) and noting from Figure 1-5 that the ratio of the radii of the two exit pupils is equal to the ratio of their distances from the Gaussian image plane, we can show that the old and the new peak aberration coefficients are related to each other according to

$$A_{s2} = A_{s1} \quad , \tag{1-14a}$$

$$A_{c2} = A_{c1} - 4bA_{s1} \quad , \tag{1-14b}$$

$$A_{a2} = A_{a1} - 2bA_{c1} + 4b^2A_{s1} \quad , \tag{1-14c}$$

$$A_{d2} = A_{d1} - bA_{c1} + 2b^2A_{s1} \quad , \tag{1-14d}$$

and

$$A_{t2} = A_{t1} - 2b\left(A_{a1} + A_{d1}\right) + 3b^2A_{c1} - 4b^3A_{s1} \quad , \tag{1-14e}$$

where

$$b = \left(L_1 - L_2\right)h'/a_1L_2 \quad . \tag{1-15}$$

In Eq. (1-15), a_1 is the radius of the exit pupil ExP_1. It is evident from Eqs. (1-14) that, because of a shift in the position of the aperture stop, an aberration of a certain order in pupil coordinates introduces aberrations of all lower orders as well. For example, a term in spherical aberration not only gives spherical aberration, but also introduces coma, astigmatism, field curvature, and distortion. From Eq. (1-14a), we note that the peak spherical aberration of a system is independent of the position of its aperture stop. Equation (1-14b) shows that if a system is free of spherical aberration, then the peak value of its coma is independent of the position of its aperture stop. It also shows that if spherical aberration is not zero, its coma can be made zero by selecting an aperture stop position corresponding to

$$b = \frac{A_{c1}}{4A_{s1}} \qquad \text{or} \qquad \frac{L_1}{L_2} = 1 + \frac{a_{c1}}{4a_{s1}} \quad . \tag{1-16}$$

Similarly, Eqs. (1-14c) and (1-14d) show that if a system is free of spherical aberration and coma, then the peak values of its astigmatism and field curvature are independent of the position of its aperture stop. Finally, Eq. (1-14e) shows that the peak value of distortion depends on the position of the aperture stop unless spherical aberration, coma, and the sum of astigmatism and field curvature are each zero.

It should be noted that the optical path length of a ray, or its optical path length difference with respect to another, does not change with a change in the position of the aperture stop. However, since the chief ray does change, the new aberration function merely describes the wave aberrations of rays with respect to the new chief ray. The position of the aperture stop also affects which and how many of the object rays are

transmitted by the system. Indeed, for high-quality imaging systems, a lens designer chooses the position of the aperture stop judiciously so that rays with large aberrations are blocked by it without a substantial loss in the amount of transmitted light.

An example of the utility of an appropriate position of the aperture stop is considered in Chapter 4, where it is shown that a spherical mirror with aperture stop located at its center of curvature suffers only from spherical aberration and field curvature. Coma, astigmatism, and distortion, which may be present for any other position of the aperture stop, are identically zero for this specific location. Indeed, such a location of the aperture stop forms the basis of the *Schmidt camera* discussed in Chapter 5.

1.8 ABERRATIONS OF A SPHERICAL REFRACTING SURFACE

In this section, we discuss imaging by a spherical refracting surface. We give equations for Gaussian imaging and expressions for its primary aberrations for an arbitrary position of its aperture stop. The results given here form the cornerstone for imaging results for a spherical mirror, which can be obtained immediately, as indicated in Chapter 4. As illustrated in Figure 1-6, consider a spherical refracting surface *SS* of radius of curvature *R* separating media of refractive indices *n* and *n'*. The line joining its vertex V_0 and its center of curvature *C* is called the *optical axis*.

Consider a point object *P* at a distance *S* from the vertex and at a height *h* from the optical axis. Let *P'* be its Gaussian image at a distance S' and a height h'. The relationships between the distances and heights of the object and image points are given by Gaussian optics according to

$$\frac{n'}{S'} - \frac{n}{S} = \frac{n' - n}{R} \tag{1-17a}$$

$$= -\frac{n}{f} = \frac{n'}{f'} \tag{1-17b}$$

and

$$M_t = \frac{h'}{h} = \frac{S' - R}{S - R} \tag{1-18a}$$

$$= \frac{n S'}{n' S} \quad , \tag{1-18b}$$

where *f* and f' are the left and the right *focal lengths* of the refracting surface and *M* is the *transverse magnification* of the image. Here *f* represents the object distance *S* such that the image distance S' is infinity. Similarly, f' represents the image distance S' such that the object distance *S* is infinity. The height of an object or image below the optical axis is considered numerically negative. (See the Appendix for sign convention.)

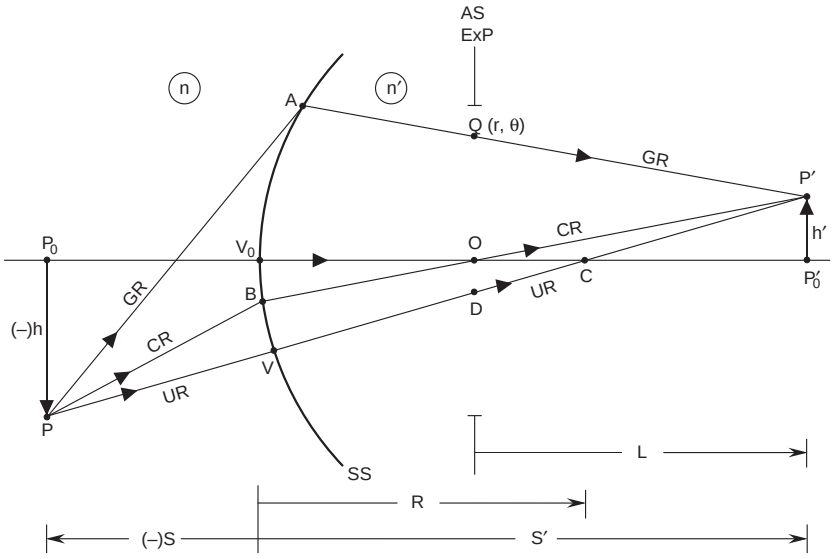


Figure 1-6. Imaging by a spherical refracting surface *SS* of radius of curvature *R* having its center of curvature at *C*, separating media of refractive indices n and *n′*. The Gaussian image plane lies at a distance *L* from the aperture stop and, therefore, its collocated exit pupil. The undeviated ray helps locate the image *P′* of a point object *P*.

In Figure 1-6, the aperture stop is also the exit pupil of the imaging system. The image lies at a distance *L* from the exit pupil. The ray PBP' passing through the center *O* of the aperture stop, which is also the exit pupil, is called the chief ray for the point object *P*. The aberration of a ray PAP' incident at a point *A* on the refracting surface and passing through a point *Q* in the plane of the exit pupil with polar coordinates (r, θ) with respect to the chief ray PBP' is given by

$$W(A) = [PAP'] - [PBP'] \quad ,$$

where the square brackets indicate an optical path length. It should be noted that the rays *PA* and *PB* from the point object *P* incident at points *A* and *B*, respectively, on the refracting surface, may not pass through the Gaussian image point P' after refraction unless the image at *P* is aberration free. It can be shown that, up to the fourth order in pupil and object or image coordinates, the aberration $W(A) \equiv W(Q)$ reduces to

$$\begin{aligned} W_s(r,\theta;h') &= a_{ss} r^4 + a_{cs} h' r^3 \cos\theta + a_{as} h'^2 r^2 \cos^2\theta \\ &\quad + a_{ds} h'^2 r^2 + a_{ts} h'^3 \cos\theta \quad , \end{aligned} \tag{1-19}$$

where

$$a_s = -\frac{n'(n'-n)}{8n^2}\left(\frac{1}{R} - \frac{1}{S'}\right)^2 \left(\frac{n'}{R} - \frac{n+n'}{S'}\right) \quad , \tag{1-20}$$

$$a_{ss} = (S'/L)^4 a_s \quad , \tag{1-21a}$$

$$a_{cs} = 4d a_{ss} \quad , \tag{1-21b}$$

$$a_{as} = 4d^2 a_{ss} \quad , \tag{1-21c}$$

$$a_{ds} = 2d^2 a_{ss} - \frac{n'(n'-n)}{4nRL^2} \quad , \tag{1-21d}$$

$$a_{ts} = 4d^3 a_{ss} - \frac{n'(n'-n)\,d}{2nRL^2} \quad , \tag{1-21e}$$

and

$$d = \frac{R - S' + L}{S' - R} \quad . \tag{1-22}$$

Note that L is (approximately) the radius of curvature of the reference sphere passing through the center of the exit pupil with its center of curvature at P'. Equation (1-19) gives the wave aberration at a point (r, θ) in the plane of the exit pupil for a point object whose Gaussian image height is h'.

The second term on the right-hand side of Eq. (1-21d) may be called the *coefficient of Petzval curvature*, and we denote it by a_p, i.e.,

$$a_p = -\frac{n'(n'-n)}{4nRL^2} \quad . \tag{1-23}$$

The corresponding wave aberration may be written

$$W_p(r) = a_p h'^2 r^2 \quad . \tag{1-24}$$

This aberration reduces to zero if the image is observed at a (longitudinal) distance Δ_L from the Gaussian image, where Δ_L is related to the aberration according to Eq.(1-3b), i.e.,

$$W_d(r) = \frac{n'}{2}\frac{\Delta_L}{L^2} r^2 = -W_p(r) \quad . \tag{1-25}$$

If the image is observed on a spherical surface of radius of curvature R_p passing through the axial image point P_0', the longitudinal defocus Δ_L for a Gaussian image at a height h' is given by its sag

$$\Delta_L = \frac{h'^2}{2R_p} \quad . \tag{1-26}$$

Comparing the values of Δ_L from Eqs. (1-25) and (1-26), and utilizing Eqs. (1-23) and Eq. (1-24), we obtain

$$R_p = \frac{nR}{n-n'} \quad . \tag{1-27}$$

We note that R_p, called the *Petzval radius of curvature*, is independent of the object position. The image surface under consideration is called the *Petzval image surface.* From Eqs. (1-21c), (1-21d), and (1-27), we may write

$$2a_{ds} - a_{as} = \frac{n'}{2R_p L^2} \quad . \tag{1-28}$$

We will utilize Eq. (1-28) in Section 7.7 where we relate the Petzval surface to the sagittal and tangential image surfaces that result from astigmatism.

Letting $h' = 0$ in Eq. (1-19), we note that the image P_0' of an axial point object P_0 suffers from spherical aberration only. The amount of spherical aberration does not change as we move from an on-axis to an off-axis point object. Note that when the aperture stop and, therefore, the exit pupil are located at the refracting surface, then $L = S'$ and Eqs. (1-21a) and (1-22) reduce to $a_{ss} = a_s$ and $d = R/(S' - R)$, respectively.

It is evident from Eq. (1-20) that $a_s = 0$ when $S' = (n + n')R/n$, which in turn corresponds to $S = (n + n')R/n'$. Accordingly, a_{ss}, a_{cs}, and a_{as} are all zero. Two conjugate points for which spherical aberration, coma, and astigmatism are zero are called *anastigmatic.* Depending on whether *R* is positive or negative, the object or the image point is virtual for these anastigmatic points. We note that spherical aberration is also zero when $S' = R$ and $S = R$. However, in this case, coma is also zero, but astigmatism is not due to the d_2 factor on the right-hand side of Eq. (1-21c). Two conjugates for which spherical aberration and coma are zero are called *aplanatic.* Thus, the points under consideration are aplanatic, and, once again, either the object or the image is virtual.

1.9 ABERRATION FUNCTION OF A MULTIELEMENT SYSTEM

Consider an optical system made up of a series of coaxial refracting and/or reflecting surfaces. Each surface produces primary aberrations with its own value of h' and *L*. The image of a point object formed by the first surface acts as an object for the second surface, and so on. The aberration is calculated surface by surface, and the aberration of the system is obtained by adding the aberration contributions of all the surfaces. Since the aberration of a surface is calculated at a point on its exit pupil, the coordinates of a pupil point must be transformed using *pupil magnification* of a surface to obtain the aberration contribution of a surface at a point on the exit pupil of the system. Similarly, image magnification of a surface can be used to obtain the system aberration in terms of the height of the image formed by the system.

For example, if $W_1(x_1, y_1; h_1')$ represents the aberration at a point (x_1, y_1) in the plane of the exit pupil of the first surface for an image of height h_1', it can be converted to an aberration contribution at a point (x_2, y_2) in the plane of the exit pupil of the second

surface and image height h_2' by letting $(x_1, y_1; h_1') = (x_2/m_2, y_2/m_2; h_2'/M_2)$, where m_2 and M_2 represent the pupil and image magnifications, respectively, for the second surface. Thus if $W_2(x_2, y_2; h_2')$ represents the aberration contribution of the second surface at a point (x_2, y_2) in the plane of its exit pupil corresponding to an image height of h_2', the total aberration for the two surfaces will be given by

$$W_s(x_2, y_2; h_2') = W_1\left(\frac{x_2}{m_2}, \frac{y_2}{m_2}; \frac{h_2'}{M_2}\right) + W_2(x_2, y_2; h_2') \quad . \tag{1-29}$$

This process can be continued to obtain the system aberration $W(x, y, h')$ at a point (x, y) in the plane of the exit pupil of the system corresponding to a height h' of the image of a point object formed by the system. It is utilized, for example, to calculate the aberrations of a thin lens in Chapter 2 and a plane-parallel plate in Chapter 3.

Since the refractive index of a transparent substance varies with optical wavelength, the angle of refraction of a ray also varies with it. Hence, even the Gaussian image of a multiwavelength point object formed by a refracting system is generally not a point. The distance and height of the image vary with the wavelength. The axial and transverse extents of the image are called *longitudinal* and *transverse chromatic aberrations*, respectively. They describe the chromatic change in position and magnification of the image, respectively. The monochromatic aberrations of a refracting system also vary with the wavelength, but such a variation is small for a small change in the wavelength and is usually negligible.

1.10 SUMMARY

The Gaussian image of a point object formed by an optical system is obtained by using Gaussian optics. The images of the aperture stop of the system by the system elements that precede and follow it are called its entrance and exit pupils, respectively. Whereas the entrance pupil determines the amount of light entering the system, the exit pupil determines how this light is distributed in the diffraction image. The optical wavefront (i.e., a surface of constant phase) exiting from the exit pupil is determined by tracing rays from the point object such that they all travel exactly the same optical path length as a ray, called the chief ray, travels in reaching the center of the pupil. If the wavefront is spherical with its center of curvature at the Gaussian image point, an aberration-free or diffraction-limited image is formed. For a circular exit pupil, the image is called the Airy pattern. Based on geometrical optics, the image formed is a point, as all of the rays pass through the Gaussian image point.

If the wavefront is not spherical, then its deviations along the rays from a corresponding spherical surface, called the Gaussian reference sphere, are called the wave aberrations of the rays. The wave aberration associated with a ray is numerically positive if it travels a longer optical path length than the chief ray to reach the reference sphere. An aberrated image is obtained in this case, the rays do not all pass through the Gaussian image point, and their distribution in the image plane is called a spot diagram. The

distance of a ray from the Gaussian image point is called the transverse ray aberration. The actual light distribution is referred to as the diffraction image. The wave and ray aberrations are related to each other according to Eq. (1-1). Whereas a wavefront tilt aberration varies linearly with the coordinates of a pupil point (in the plane of the exit pupil), a wavefront defocus aberration varies quadratically with the distance of the pupil point from its center.

The optical path length of a ray does not change when the location of the aperture stop is changed, but, since the chief ray changes, their aberrations with respect to it also change. The position of the aperture stop also affects which and how many of the rays are transmitted by the system. The size of the aperture stop is adjusted so that the amount of light from an axial point object remains unchanged. While the peak value of spherical aberration does not change, the coefficients of the other aberrations can and do change. Indeed, a lens designer chooses the position of the aperture stop judiciously so that the rays with large aberrations are blocked by it, without a substantial loss in the amount of transmitted light.

APPENDIX: SIGN CONVENTION

Although there is no universally accepted standard sign convention, we will use the Cartesian sign convention. It has the advantage that there are no special rules to remember other than those of a right-handed Cartesian coordinate system. Our sign convention is the same as that used by Mouroulis and Macdonald, but it is slightly different in its implementation from those of Born and Wolf, Welford, and Schroeder. It is different from the sign convention used, for example, by Jenkins and White, Klein and Furtak, and Hecht and Zajac. The rules of our sign convention are listed below.

1. Light is incident on a system from left to right.

2. Distances to the right of and above (left of and below) a reference point are positive (negative).

3. The radius of curvature of a surface is treated as the distance of its center of curvature from its vertex. Thus, it is positive (negative) when the center of curvature lies to the right (left) of the vertex.

4. The acute angle of a ray from the optical axis or from the surface normal is positive (negative) if it is counterclockwise (clockwise).

5. When light travels from right to left, as when it is reflected by an odd number of mirrors, then the refractive index and the spacing between two adjacent surfaces are given a negative sign.

Throughout the book, any quantities that are numerically negative are indicated in the figures by a parenthetical negative sign $(-)$.

CHAPTER 2
Thin Lens

2.1 INTRODUCTION

Among the simple optical imaging systems, a thin lens consisting of two spherical surfaces is the most common as well as practical. By applying the results of Section 1.8 and the procedure of Section 1.9, we give the imaging equations and expressions for the primary aberrations of a thin lens with aperture stop located at the lens. Its aberrations for other locations of the aperture stop may be obtained by applying the results of Section 1.7 to those given here. It is shown that when both an object and its image are real, the spherical aberration of a thin lens cannot be zero (unless its surfaces are made nonspherical). We illustrate by a numerical example, however, that it is possible to design a two-lens combination such that its spherical aberration and coma are both zero. In such a combination, these aberrations associated with one lens cancel the corresponding aberrations of the other. This cancellation is illustrated with a numerical example.

2.2 GAUSSIAN IMAGING

Consider a thin lens of refractive index n and focal length f' consisting of two spherical surfaces of radii of curvature R_1 and R_2 as illustrated in Figure 2-1. A lens is considered *thin* if its thickness is negligible compared to f', R_1, and R_2. Its optical axis *OA* is the line joining the centers of curvature C_1 and C_2 of its surfaces. Since the lens is thin, we neglect the spacing between its surfaces. We assume that its aperture stop *AS* is located at the lens, so that its entrance and exit pupils *EnP* and *ExP*, respectively, are also located there. The lens is located in air; therefore, the refractive index of the surrounding medium is 1.

Consider a point object *P* located at a distance *S* from the lens and at a height *h* from its axis. The first surface forms the image of *P* at P' and the second surface forms the image of P' at P''. Applying the results of Section 1.8 to imaging by the two surfaces of the lens, where $n = 1$ and $n' = n$ for the first surface and $n = n$ and $n' = 1$ for the second surface, we can show that the image distance S' and its height h' are given by the relations

$$\frac{1}{S'} - \frac{1}{S} = (n-1)\left(\frac{1}{R_1} - \frac{1}{R_2}\right) \tag{2-1a}$$

$$= \frac{1}{f'} \tag{2-1b}$$

and

$$M = \frac{h'}{h} = \frac{S'}{S}\ , \tag{2-2}$$

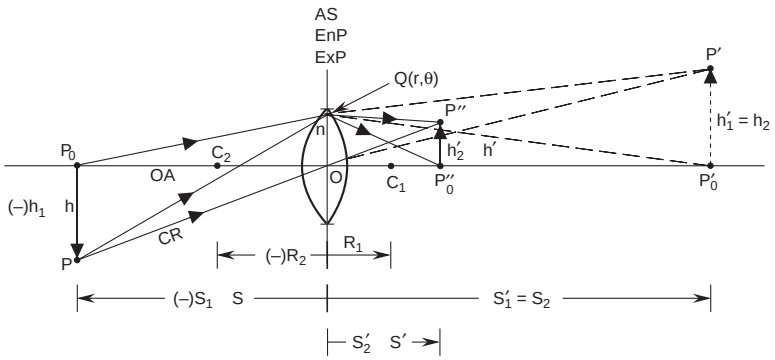


Figure 2-1. Imaging by a thin lens of refractive index *n* formed by two surfaces of radii of curvature R_1 and R_2 with their centers of curvature at C_1 and C_2. Whereas R_1 is numerically positive, R_2 is negative. $P_0'P'$ is the Gaussian image of the object P_0P formed by the first surface. $P_0''P''$ is the image of the virtual object $P_0'P'$ formed by the second surface. The aperture stop *AS*, entrance pupil *EnP*, and the exit pupil *ExP* are all located at the lens.

respectively, where *M* is the magnification of the image. Note that we are able to write Eq. (2-1b) because, by definition, the *focal length* f' is the image distance when the object is at infinity.

2.3 PRIMARY ABERRATIONS

The aberration of an object ray PQP'' passing through a point Q in the plane of the exit pupil with polar coordinates (r,θ) with respect to the chief ray POP'' passing through the center 0 of the exit pupil is given by

$$W(Q) = [PQP''] - [POP''] \quad .$$

Noting that the optical path lengths $[P'Q]$ and $[P'O]$ are numerically negative, since they are *virtual*, the aberration of the ray can be written in terms of the aberrations produced by the two surfaces, i.e.,

$$W(Q) = \{[PQP'] - [POP']\} + \{[P'QP''] - [P'OP'']\} \quad .$$

By applying the results of Section 1.8 to the two surfaces of the thin lens and following the procedure of Section 1.9, it can be shown that the primary aberration function of the thin lens is given by

$$W(r,\theta;h') = a_s r^4 + a_c h' r^3 \cos\theta + a_a h'^2 r^2 \cos^2\theta + a_d h'^2 r^2 + a_t h'^3 r \cos\theta \quad , \qquad (2\text{-}3)$$

where

$$a_s = -\frac{1}{32n(n-1)f'^3}\left[\frac{n^3}{n-1} + (3n+2)(n-1)p^2 + \frac{n+2}{n-1}q^2 + 4(n+1)pq\right] \quad , \quad (2\text{-}4a)$$

$$a_c = -\frac{1}{4nf'^2S'}\left[(2n+1)p+\frac{n+1}{n-1}q\right] \quad , \tag{2-4b}$$

$$a_a = -\frac{1}{2f'S'^2} \quad . \tag{2-4c}$$

and

$$a_d = \frac{1}{2}a_a - \frac{1}{4nf'S'^2} \quad . \tag{2-4d}$$

Note that there is no distortion term in Eq. (2-3); i.e., a thin lens with an aperture stop at the lens does not produce any distortion. The quantities p and q are called the *position* and *shape factors* of a thin lens, respectively. They are given by

$$p = -\frac{2f'}{S_1}-1 \tag{2-5a}$$

$$= 1-\frac{2f'}{S_2'} \tag{2-5b}$$

and

$$q = \frac{R_2+R_1}{R_2-R_1} \quad . \tag{2-6}$$

Several examples of the position and shape factors are illustrated in Figures 2-2 and 2-3, respectively. Both positive and negative lenses (in the sense of the sign of their focal length) are considered in these figures. The names associated with the different lens shapes are also noted in Figure 2-3.

We note from Eqs. (2-4c) and (2-4d) that astigmatism and field curvature coefficients of a thin lens do not depend on its position and shape factors. Moreover, the astigmatism coefficient does not depend on the refractive index of the lens, and the field curvature coefficient is smaller than the astigmatism coefficient by a factor of $(n+1)/2n$.

2.4 SPHERICAL ABERRATION AND COMA

From Eqs. (2-4a) and (2-4b) we note that the spherical aberration and coma of a thin lens depend on its position and shape factors. For a given position factor, the value of the shape factor that minimizes the spherical aberration is given by the condition

$$\frac{\partial a_s}{\partial q} = 0 \quad . \tag{2-7}$$

Thus, we obtain

$$q_{min} = -2p\frac{n^2-1}{n+2} \quad . \tag{2-8}$$

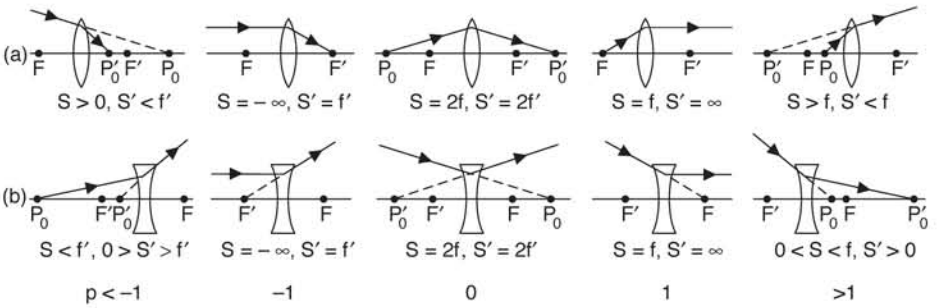


Figure 2-2. ***Position factor*** **$1 < p < -1$ of a thin lens. (a) Positive lens, i.e., $f' > 0$. (b) Negative lens, i.e., $f' < 0$. F and F' are the object- and image-space focal points of a lens of image-space focal length f'. P_0 and P_0' represent an axial point object and its point image, respectively. S and S' are the object and image distances from the center of the lens. Note that $f = -f'$.**

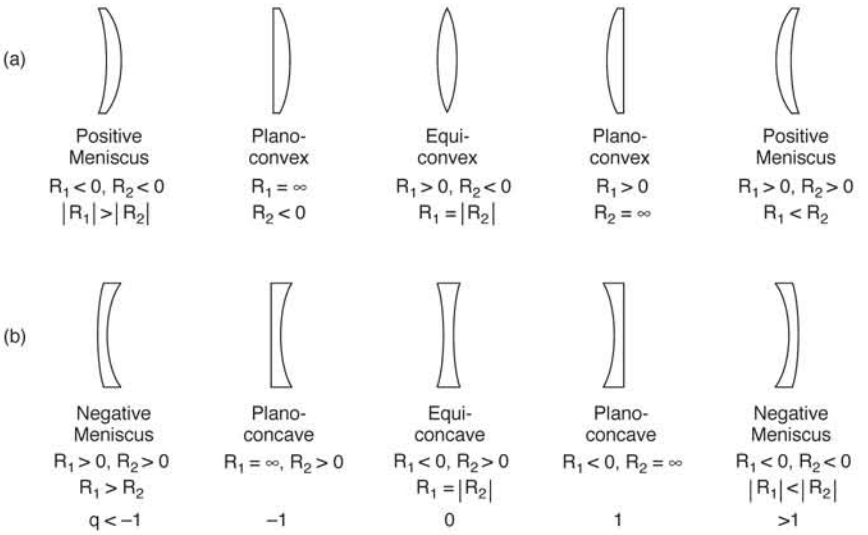


Figure 2-3. ***Shape factor*** **$1 < q < -1$ of a thin lens with spherical surfaces of radii of curvature R_1 and R_2. (a) Positive lens. (b) Negative lens.**

Substituting Eq. (2-8) into Eq. (2-4a), we obtain the corresponding *minimum spherical aberration*:

$$a_{smin} = -\frac{1}{32 f'^3}\left[\left(\frac{n}{n-1}\right)^2 - \frac{n}{n+2}p^2\right] \quad . \tag{2-9}$$

Thus, following Eq. (2-4a), we note that, for a given value of p, a_s as a function of q follows a parabola with a vertex lying at (q_{min}, a_{smin}). For different values of p, the parabolas have the same shape but different vertices. It is evident from Eqs. (2-5a) and (2-5b) that when both an object and its image are real,

$$-1 \le p \le 1 \quad , \qquad \text{or} \qquad p^2 \le 1 \quad . \tag{2-10}$$

As indicated in Figure 2-2, the case $p = -1$ corresponds to an object at infinity and the image at the focal plane of the lens. Similarly, $p = 1$ corresponds to an object at the focal plane and the image at infinity. The case $p = 0$ corresponds to object and image lying at distances of $2f$ and $2f'$, respectively. For spherical aberration to be zero, Eq. (2-9) yields

$$p^2 = \frac{n(n+2)}{(n-1)^2} > 1 \quad . \tag{2-11}$$

Hence, spherical aberration of a thin lens cannot be zero when both the object and its image are real.

For a thin lens with a refractive index $n = 1.5$, Eqs. (2-4a), (2-8), and (2-9) reduce to

$$a_s = -\frac{1}{24f'^3}\left(6.75 + 3.25p^2 + 7q^2 + 10pq\right) \quad , \tag{2-12a}$$

$$q_{min} = -(5/7)p \quad , \tag{2-12b}$$

and

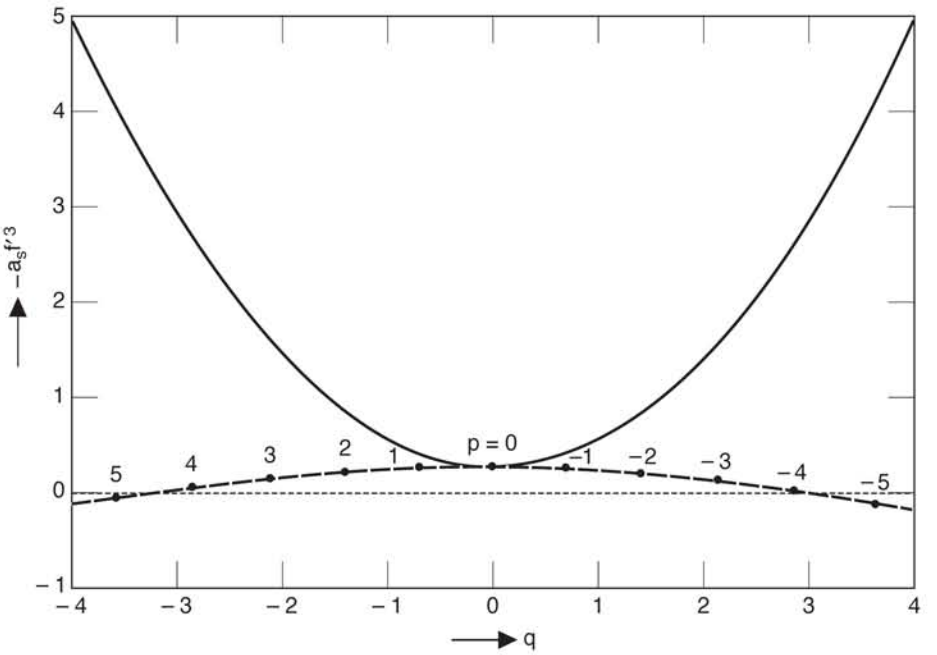


Figure 2-4. Parabolic variation of spherical aberration of a thin lens with its shape factor q for $p = 0$. How its minimum value varies with q is indicated by the lower parabolic curve. Several values of p are indicated on this curve.

$$a_{smin} = -\frac{1}{32 f'^3}\left(9 - \frac{3}{7}p^2\right) \quad , \tag{2-12c}$$

respectively. Figure 2-4 shows the parabolic variation of spherical aberration with q for $p = 0$. The minimum value of spherical aberration corresponds to $q = 0$, i.e., an equiconvex lens. As pointed out earlier, the variation of spherical aberration with q for other values of p follows the same parabola except that the location of its vertex (q_{min}, a_{smin}) depends on p. The vertices of the parabolas follow the lower parabolic curve in Figure 2-4, which represents a_{smin} as a function of q, obtained by substituting Eq. (2-12b) into Eq. (2-12c). The solid dots on this curve indicate various values of p. The minimum value of spherical aberration approaches zero for $|p| = \sqrt{21}$. It changes its sign for larger values of $|p|$.

It follows from Eq. (2-4b) that the coma of a thin lens is zero if its position and shape factors are related to each other according to

$$q = -\frac{(2n+1)(n-1)}{n+1}p \quad . \tag{2-13}$$

For n = 1.5, Eqs. (2-4b) and (2-13) reduce to

$$a_c = -\frac{1}{6 f'^2 S'}(4p + 5q) \tag{2-14}$$

and

$$q = -0.8p \quad , \tag{2-15}$$

respectively. For $p = -1$, the values of q giving minimum spherical aberration $(q_{min} = 0.71)$ and zero coma $(q = -0.8)$ are approximately the same. Thus, a lens designed for zero coma for parallel incident light will have practically the minimum amount of spherical aberration. It is also possible to design and combine two thin lenses such that the spherical aberration and coma of one cancel the corresponding aberrations of the other, as illustrated by a numerical example in the next section.

2.5 NUMERICAL PROBLEMS

2.5.1 Thin Lens Focusing a Parallel Beam of Light

As a numerical example, we determine the radii of curvature of the surfaces of a thin lens of refractive index 1.5 focusing a parallel beam of light at a distance of 15 cm from it with minimum spherical aberration. According to Eq. (2-5), $p = -1$ for a parallel beam. Substituting in Eq. (2-12b), we obtain $q_{min} = 5/7$ for minimum spherical aberration. Equation (2-6), therefore, gives

$$\frac{R_2}{R_1} = \frac{q+1}{q-1} = -6 \quad . \tag{2-16}$$

Since $f' = 15$ cm, Eqs. (2-1a) and (2-1b) yield $R_1 = 8.75$ cm and $R_2 = -52.50$ cm, corresponding to a nearly *plano-convex* lens with its convex side facing the incident light. For a lens of diameter 2 cm, the peak value of spherical aberration, according to Eqs. (2-3) and (2-12c), is given by $A_s = -0.79$ μm. The other primary aberrations of the focused beam can be obtained from Eqs. (2-3) and (2-4). Thus, it can be shown, for example, that the peak values of coma, astigmatism, and field curvature for a parallel beam incident on the lens at an angle of 5° from its axis, corresponding to an image height of 1.31 cm, are given by $A_c = 0.28\,\mu$m, $A_a = -2.5\,\mu$m, and $A_d = -2.1\,\mu$m. A thin lens with aperture stop located at the lens does not produce any distortion. It may be noted that if the lens is turned around so that its (relatively) planar side faces the incident light, its focal length does not change. However, its shape factor changes sign, thereby changing both its spherical aberration as well as its coma.

2.5.2 Aplanatic Doublet Focusing a Parallel Beam of Light

Since spherical aberration of a thin lens varies as f'^{-3}, it is possible to make it zero for a combination of lenses having focal lengths of different signs. *A doublet* designed to correct for spherical aberration can at the same time be corrected for coma. For example, we now show that two thin lenses of refractive index 1.5 focusing a parallel beam of light with radii of curvature 9.2444 cm and –15.5197 cm for the first lens, and –9.5618 cm and – 15.3120 cm for the second lens, give zero spherical aberration and coma with a focal length of 15 cm when placed in contact with each other. Substituting for the refractive index and the radii of curvature of the lens surfaces into Eqs. (2-1), we find that the focal lengths of the two lenses are given by $f_1' = 11.5870$ cm and $f_2' = -50.9235$ cm. Hence, the focal length of the doublet given by $f'^{-1} = f_1'^{-1} + f_2'^{-1}$ is $f' = 15$ cm. The shape factors of the lenses are given by $q_1 = 0.2534$ and $q_2 = 4.3257$. For a parallel beam of incident light, the position factor for the first lens is given by $p_1 = -1$. Substituting for n, p_1, and q_1 into Eq. (2-12a), we find that the spherical aberration coefficient for the first lens is $a_{s1} = -2.1201 \times 10^{-4}\,\text{cm}^{-3}$. Since the second lens focuses the beam at a distance of $S_2' = 15$ cm, its position factor is given by $p_2 = 1 - 2f_2'/S_2'$ or $p_2 = 7.7898$. Substituting for n, p_2, and q_2 into Eq. (2-12a), we find that the spherical aberration coefficient for the second lens is $a_{s2} = 2.1201 \times 10^{-4}\,\text{cm}^{-3}$, which is equal in magnitude but opposite in sign to the corresponding coefficient for the first lens. Hence, spherical aberration of the doublet is zero.

Now we consider the coma aberrations produced by the two lenses and the lens doublet. The first lens focuses the incident parallel beam at a distance $S_1' = f_1'$. Equation (2-14) yields the coma coefficient for the first lens, $a_{c1} = 2.9281 \times 10^{-4}\,\text{cm}^{-3}$. Similarly, for the second lens, $a_{c2} = -2.2618 \times 10^{-4}\,\text{cm}^{-3}$. Now, for a beam incident at an angle β from the axis of the thin-lens doublet, the first lens focuses it at a height of $h_1' = \beta f_1'$. The second lens forms the image of this focus at a height h_2' given by Eq. (2-2), i.e., $h_2'/h_1' = -S_2'/S_2 = S_2'/S_1' = f'/f_1' = 1.2946$. If (r, θ) represent the polar coordinates of a point in the plane of the thin-lens doublet, the coma aberrations produced by the two lenses are given by

$$W_{c1}(r,\theta) = a_{c1}h_1'r^3\cos\theta$$

$$= 2.2618\times 10^{-4}h_2'r^3\cos\theta \qquad (2\text{-}17a)$$

and

$$W_{c2}(r,\theta) = a_{c2}h_2'r^3\cos\theta$$

$$= -2.2618\times 10^{-4}h_2'r^3\cos\theta \quad . \qquad (2\text{-}17b)$$

The coma aberration of the lens doublet is given by

$$W_c(r,\theta) = W_{c1}(r,\theta)+W_{c1}(r,\theta) = 0 \quad . \qquad (2\text{-}18)$$

Thus, both spherical aberration and coma of the doublet are zero. Such a system is called *aplanatic*.

Finally, we consider the astigmatism and field curvature aberrations of the lens doublet. Substituting for the focal length and the image distance for the two lenses in Eq. (2-4c), we obtain their astigmatism coefficients $a_{a1} = -3.2141\times 10^{-4}\,\text{cm}^{-3}$ and $a_{a2} = 4.3638\times 10^{-5}\,\text{cm}^{-3}$. Hence, astigmatism aberration of the doublet at a point (r,θ) in its plane may be written

$$W_a(r,\theta;h_2') = W_{a1}(r,\theta;h_1')+W_{a2}(r,\theta;h_2')$$

$$= a_{a1}h_1'^2r^2\cos^2\theta + a_{a2}h_2'^2r^2\cos^2\theta$$

$$= (0.5967a_{a1}+a_{a2})h_2'^2r^2\cos^2\theta$$

$$= -1.4815\times 10^{-4}h_2'^2r^2\cos^2\theta \quad . \qquad (2\text{-}19)$$

For a beam incident on the doublet at an angle of 5° from its axis, we obtain $h_2' = 1.31$ cm. Hence, for a beam of diameter 2 cm, the peak value of astigmatism aberration is approximately given by $A_a = -2.54$ μm. Comparing Eqs. (2-4c) and (2-4d), the corresponding field curvature aberration may be obtained from A_a by multiplying it by $(n+1)/2n$. Thus, we find that $A_d = -2.12\,\mu\text{m}$.

2.6 SUMMARY

A thin lens generally consists of two spherical surfaces with negligible thickness. The spherical aberration of such a thin lens cannot be zero when an object and its image are both real. However, since this aberration varies as the cube of the lens focal length, it can be made zero by combining two lenses of focal lengths with opposite signs. The doublet, as it is called, can also be made coma free, and the system is then referred to as being aplanatic.

CHAPTER 3
Aberrations of a Plane-Parallel Plate

3.1 INTRODUCTION

In Chapter 2, we considered the imaging properties of a thin lens consisting of two spherical surfaces. Now, we consider "imaging" by a *plane-parallel plate*, i.e., a plate whose two surfaces are parallel to each other, and each with a radius of curvature of infinity. Unlike a lens, such a plate is not used for imaging per se, but it is often used in imaging systems, for example, as a beam splitter or a window. The imaging relations and aberrations of a plane-parallel plate cannot be obtained from those for a thin lens in Chapter 2 by letting the radii of curvature of its surfaces approach infinity, since we neglected its thickness. However, as discussed below, they can be obtained by applying the results of Section 1.8 to its two surfaces and combining the results obtained according to the discussion of Section 1.9. It is shown that the distance between an object and its image formed by the plate, called the *image displacement*, is independent of the object position, and the aberration produced by it approaches zero as the object distance approaches infinity. Thus, a plane-parallel plate placed in the path of a converging beam not only displaces its focus by a certain amount but also introduces aberrations into it. In the case of a collimated beam, it only shifts the beam without introducing any aberrations.

3.2 GAUSSIAN IMAGING

Consider, as indicated in Figure 3-1, a plane-parallel plate of thickness t and refractive index n forming an image of a point object lying at a distance S from its front surface and at a height h from its axis. Let the aperture stop of the plate be of radius a located at its front surface. First, we determine the location of the image formed by the plate. Using Eqs. (1-17) and (1-18) we determine the location and height of the image. For the first surface, $n_1 = 1$, $n_1' = n$ and $R_1 = \infty$. Accordingly, it forms the image of P at P' such that

$$S_1' = nS_1 \equiv nS \tag{3-1}$$

and

$$M_1 = h_1'/h_1 = n_1 S_1'/n_1' S_1 = 1 \quad , \tag{3-2}$$

where $h_1 \equiv h$. For the second surface, $n_2 = n$, $n_2' = 1$, $R_2 = \infty$, and $S_2 = S_1' - t$. Hence, it forms the image of P' at P'' such that

$$S_2' = S_2/n = (S_1' - t)/n \tag{3-3}$$

and

$$M_2 = h_2'/h_1' = n_2 S_2'/n_2' S_2 = 1 \quad . \tag{3-4}$$

Substituting for S_1' from Eq. (3-1) into Eq. (3-3) and noting that S_2' is numerically

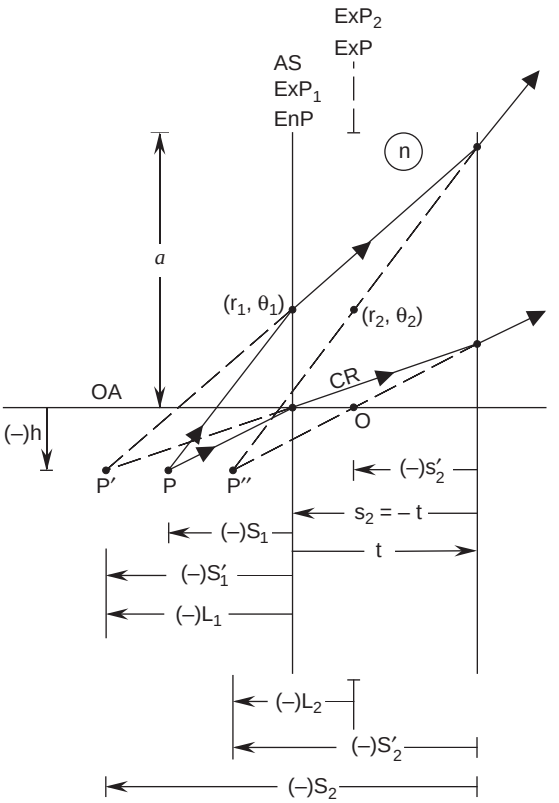


Figure 3-1. Imaging of a point object *P* by a plane-parallel plate of refractive index *n* of *P′* formed by the second surface of the plate. The aperture stop *AS* and, therefore, the entrance pupil *EnP* of the plate are located at the first surface. A negative sign in parentheses indicates a numerically negative quantity.

negative, the displacement PP'' of the final image from the object may be written

$$PP'' = -S_1 - (-S_2' - t)$$

$$= t\left(1 - 1/n\right) \quad . \tag{3-5}$$

Thus, the image displacement is independent of the object distance *S*. It depends only on the thickness and the refractive index of the plate.

Next, we determine the locations and magnifications of the pupils for the two surfaces of the plate. Since the aperture stop is located at the first surface, the entrance pupil *EnP* of the system is also located there. Moreover, the entrance and exit pupils EnP_1 and ExP_1 for this surface are also located at the surface. The entrance pupil EnP_2 for the second surface is ExP_1. The exit pupil ExP_2 for this surface is the image of EnP_2 formed by it. Thus, letting $n_2 = n$, $n_2' = 1$, $s_2 = -t$, and $R_2 = \infty$, we find from Eqs. (1-17) and (1-18) that ExP_2 is located at a distance $s_2' = -t/n$ from the second surface and its magnification $m_2 = 1$. Of course, ExP_2 is also the exit pupil *ExP* of the system. It is

evident that, for the first surface, the distance L_1 of the image P' from ExP_1 is equal to its distance S_1' from the surface. For the second surface, distance L_2 of the image P'' from ExP_2 is given by

$$L_2 = S_2' - s_2' \quad , \tag{3-6a}$$

since L_2, S_2', and s_2' are all numerically negative. Substituting for S_2' and s_2', we find that

$$L_2 = S \quad . \tag{3-6b}$$

Now we use the results obtained above to determine the aberrations produced by the plate.

3.3 PRIMARY ABERRATIONS

First, we determine the aberration $W_1(r_1, \theta_1; h_1')$ contributed by the first surface at a point (r_1, θ_1) in the plane of ExP_1. Letting $n_1 = 1$, $n_1' = n$, and $R_1 = \infty$, Eq. (1-20) yields

$$a_{s1} = \frac{n(n^2 - 1)}{8S_1'^3} \quad . \tag{3-7}$$

Moreover, Eq. (1-22) reduces to $d_1 = -1$, and since $S_1' = L_1$, Eq. (1-21a) reduces to $a_{ss1} = a_{s1}$. The Petzval contributions to field curvature and distortion represented by the second term on the right-hand side of Eqs. (1-21d) and (1-21e) are zero. Hence, for the first surface, Eq. (1-19) may be written

$$W_1(r_1, \theta_1; h_1') = a_{s1}(r_1^4 - 4h_1' r_1^3 \cos\theta_1 + 4h_1' r_1^2 \cos^2\theta_1 + 2h_1'^2 r_1^2 - 4h_1'^3 r_1 \cos\theta_1) \quad . \tag{3--8}$$

Next, we determine the aberration $W_2(r_2, \theta_2; h_2')$ contributed by the second surface at a point (r_2, θ_2) in the plane of ExP_2. Letting $n_2 = n$, $n_2' = 1$, and $R_2 = \infty$, Eq. (1-20) yields for this surface

$$a_{s2} = -\frac{n^2 - 1}{8n^2 S_2'^3} \quad . \tag{3-9}$$

Once again, Eq. (1-22) reduces to $d_2 = -1$, and the Petzval contributions to field curvature and distortion are zero. Hence, for the second surface, Eq. (1-19) may be written

$$W_2(r_2, \theta_2; h_2') = a_{ss2}(r_2^4 - 4h_2' r_2^3 \cos\theta_2 + 4h_2'^2 r_2^2 \cos^2\theta_2 + 2h_2'^2 r_2^2 - 4h_2'^3 r_2 \cos\theta_2) \ , \tag{3-10}$$

where

$$a_{ss2} = (S_2'/L_2)^4 a_{s2} \quad . \tag{3-11}$$

Finally, we combine the aberrations introduced by the two surfaces to obtain the aberration introduced by the plate. Since m_2 and M_2 are both unity, $(r_1, \theta_1) = (r_2, \theta_2)$ and $h_2' = h_1' = h$, respectively. Hence, following Eq. (1-29), the aberration of the plane-parallel plate at a point (r, θ) in the plane of its exit pupil can be written

$$W(r,\theta;h) = W_1(r,\theta;h) + W_2(r,\theta;h) \quad . \tag{3-12}$$

Substituting Eqs. (3-8) and (3-10) into Eq. (3-12), we may write the primary aberration function

$$W(r,\theta;h) = a_s\left(r^4 - 4hr^3\cos\theta + 4h^2r^2\cos^2\theta + 2h^2r^2 - 4h^3r\cos\theta\right) \quad , \tag{3-13}$$

where

$$a_s = a_{s1} + \left(S_2'/L_2\right)^4 a_{s2} \quad . \tag{3-14}$$

Substituting Eqs. (3-1), (3-3), (3-6b), (3-7), and (3-9) into Eq. (3-14), we obtain

$$a_s = \frac{\left(n^2-1\right)t}{8n^3S^4} \quad . \tag{3-15}$$

Note that the aberration increases linearly with the plate thickness *t*. Moreover, as expected, the aberration reduces to zero for a collimated incident beam $(S \to -\infty)$. This is indeed why a lens designer places beam splitters and windows in an imaging system in its collimated spaces wherever possible.

3.4 NUMERICAL PROBLEM

As a numerical example we determine the aberrations of a plane-parallel plate placed in the path of a converging beam as shown in Figure 3-2. The plate has a refractive index of 1.5. Its thickness is 1 cm and its diameter is 4 cm. In the absence of the plate, the beam comes to a focus at *P* at a distance of 8 cm from its front surface at a height of 0.5 cm from its axis. From Eq. (3-5), we find that the plate displaces the image from *P* to P' which is at the same height as *P* but at a distance of 8.33 cm from its front surface. Substituting for *n*, *t*, and *S* = 8 cm in Eq. (3-15), we obtain $a_s = 1.13 \times 10^{-5}\,\text{cm}^{-3}$. Noting that the maximum value of *r* is 2 cm, we obtain the peak values of the primary aberrations introduced by the plate from Eq. (3-13); $A_s = 1.81\,\mu\text{m}$, $A_c = -1.81\,\mu\text{m}$, $A_a = 0.45\,\mu\text{m}$, $A_d = 0.23\,\mu\text{m}$, and $A_t = -0.11\,\mu\text{m}$.

3.5 SUMMARY

A plane-parallel plate is often used in imaging systems as a beam splitter or a window. It introduces no aberrations when placed in a collimated space. However, it introduces aberrations when placed in a converging or a diverging beam of light. The primary aberrations thus introduced are given by Eq. (3-13). They increase linearly with the thickness of the plate.

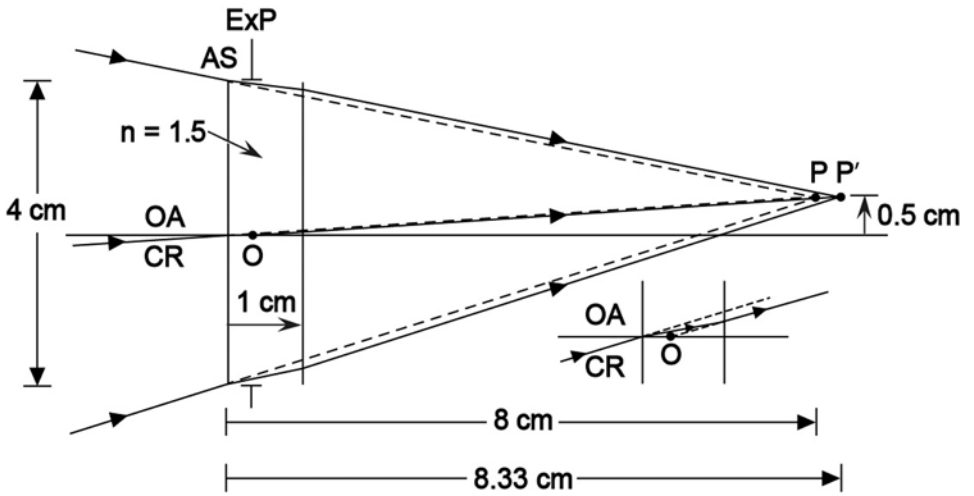


Figure 3-2. Plane-parallel plate placed in a converging beam of light. Rays incident on the plate converging toward P converge toward P' after refraction by it.

CHAPTER 4
Aberrations of a Spherical Mirror

4.1 INTRODUCTION

So far, we have considered refracting imaging systems: a spherical refracting surface in Chapter 1, a thin lens in Chapter 2, and a plane-parallel plate in Chapter 3. Now we consider the imaging properties of a spherical reflecting surface, i.e., a *spherical mirror*. These properties can be obtained in a manner similar to that for a spherical refracting surface. However, the geometry of the problem is different since now a ray incident on the surface is reflected back into the same medium containing the incident ray, instead of being refracted into another medium. Accordingly, it is instructive to draw object and image rays and not blindly use the imaging and aberration relations appropriate for a reflecting surface. In this chapter, we give the relations describing the primary aberrations of a spherical mirror for an arbitrary position of the aperture stop. These relations are applied to specific cases, one when the aperture stop is located at the mirror and the other when it is located at its center of curvature. It is shown that, in the first case, field curvature and distortion are zero. In the second case, coma, astigmatism, and distortion are zero. A numerical problem illustrates these results.

4.2 PRIMARY ABERRATION FUNCTION

Consider an imaging system consisting of a spherical mirror of radius of curvature R and focal length f'. Let the aperture stop and the corresponding exit pupil of the system be located as indicated in Figure 4-1. The line joining the center of curvature C of the mirror and the center of the aperture stop (and, therefore, the center O of the exit pupil) defines the optical axis of the system. Consider an object lying at a distance S from the vertex V_0 of the mirror. Let the height of an object point P from the optical axis be h. The distance S' and the height h' of its Gaussian image P' are given by

$$\frac{1}{S'} + \frac{1}{S} = \frac{2}{R} = \frac{1}{f'} \quad , \tag{4-1}$$

and

$$M = \frac{h'}{h} = \frac{S' - R}{S - R} \tag{4-2a}$$

$$= -S'/S \quad , \tag{4-2b}$$

respectively, where M is the magnification of the image.

The aberration $W(Q)$ of an object ray incident at a point A on the mirror passing through a point Q in the plane of the exit pupil with polar coordinates (r, θ) with respect to the chief ray passing through the center O of the exit pupil is given by

$$W(Q) = [PAP'] - [PBP'] \quad .$$

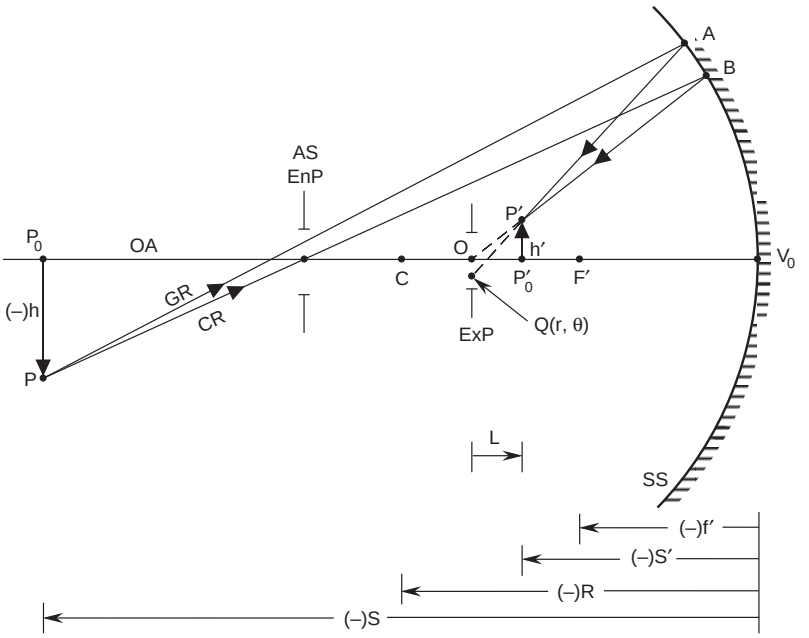


Figure 4-1. Imaging by a spherical mirror of radius of curvature *R*. The line joining the center of the aperture stop *AS* and the center of curvature *C* of the mirror defines the optical axis *OA*. The chief ray *CR* from a point object *P* passes through the center of the aperture stop.

It can be shown that, up to the fourth order in pupil and object or image coordinates, the aberration $W(A) \equiv W(Q)$ reduces to

$$W_s(r,\theta;h') = a_{ss}r^4 + a_{cs}h'r^3\cos\theta + a_{as}h'^2r^2\cos^2\theta + a_{ds}h'^2r^2 + a_{ts}h'^3r\cos\theta \ , \qquad (4\text{-}3)$$

where

$$a_s = \frac{n}{4R}\left(\frac{1}{R}-\frac{1}{S'}\right)^2 \ , \qquad (4\text{-}4a)$$

$$= \frac{1}{4R}\left(\frac{1}{R}-\frac{1}{S'}\right)^2 \ , \qquad (4\text{-}4b)$$

$$a_{ss} = (S'/L)^4\, a_s \ , \qquad (4\text{-}5a)$$

$$a_{cs} = 4da_{ss} \ , \qquad (4\text{-}5b)$$

$$a_{as} = 4d^2a_{ss} \ , \qquad (4\text{-}5c)$$

$$a_{ds} = 2d^2a_{ss} - \frac{n}{2RL^2} \ , \qquad (4\text{-}5d)$$

$$a_{ts} = 4d^3 a_{ss} - \frac{d}{RL^2} \quad , \tag{4-5e}$$

$$d = \frac{R - S' + L}{S' - R} \quad , \tag{4-6}$$

and L is the distance of the Gaussian image plane from the plane of the exit pupil. Thus, it is numerically positive in Figure 4-1. We note from the equations given above that, unless a_s is zero, coma, astigmatism, and distortion of a spherical mirror are zero when $d = 0$. As discussed in Section 4.4, this happens when the aperture stop of the mirror is located at its center of curvature. As in the case of a spherical refracting surface, spherical aberration and coma are zero when the object is located at the center of curvature of the mirror, i.e., when $S = -R$

Comparing Eqs. (4-1)–(4-6) with Eqs. (1-17)–(1-22), we note that the results for a reflecting surface can be obtained from those for a refracting surface if we let $n = 1$ (since the mirror is in air), $n' = -1$ (minus sign representing reflection).

4.3 APERTURE STOP AT THE MIRROR

If the aperture stop is located at the mirror as in Figure 4-2, then the entrance and exit pupils are also located there. Accordingly, $L = S'$ and $a_{ss} \to a_s$ and $d \to R/(S' - R)$. The primary aberration function given by Eq. (4-3) becomes

$$W_s(r, \theta; h') = \frac{1}{4R}\left(\frac{1}{R} - \frac{1}{S'}\right)^2 r^4 + \frac{S' - R}{R^2 S'^2} h' r^3 \cos\theta + \frac{1}{RS'^2} h'^2 r^2 \cos^2\theta \quad . \tag{4-7}$$

It represents the optical path difference of a ray such as PQP' with respect to the chief ray PV_0P' in Figure 4-2 up to the fourth order in pupil and object (or image) coordinates. Note that the field curvature and distortion coefficients are zero.

If the object is located at infinity, as in astronomical observations, then

$$S' = R/2 = f' \tag{4-8}$$

and

$$d = -2 \quad . \tag{4-9}$$

If it lies at an angle β from the optical axis, then

$$h' = -\beta f' \quad . \tag{4-10}$$

Substituting Eqs. (4-8)–(4-10) into Eq. (4-7), we obtain the primary aberration function for a spherical mirror for an object at infinity at an angle β from its optical axis:

$$W_s(r, \theta; \beta) = \frac{1}{32 f'^3} r^4 + \frac{1}{4 f'^2} \beta r^3 \cos\theta + \frac{1}{2f'} \beta^2 r^2 \cos^2\theta \quad . \tag{4-11}$$

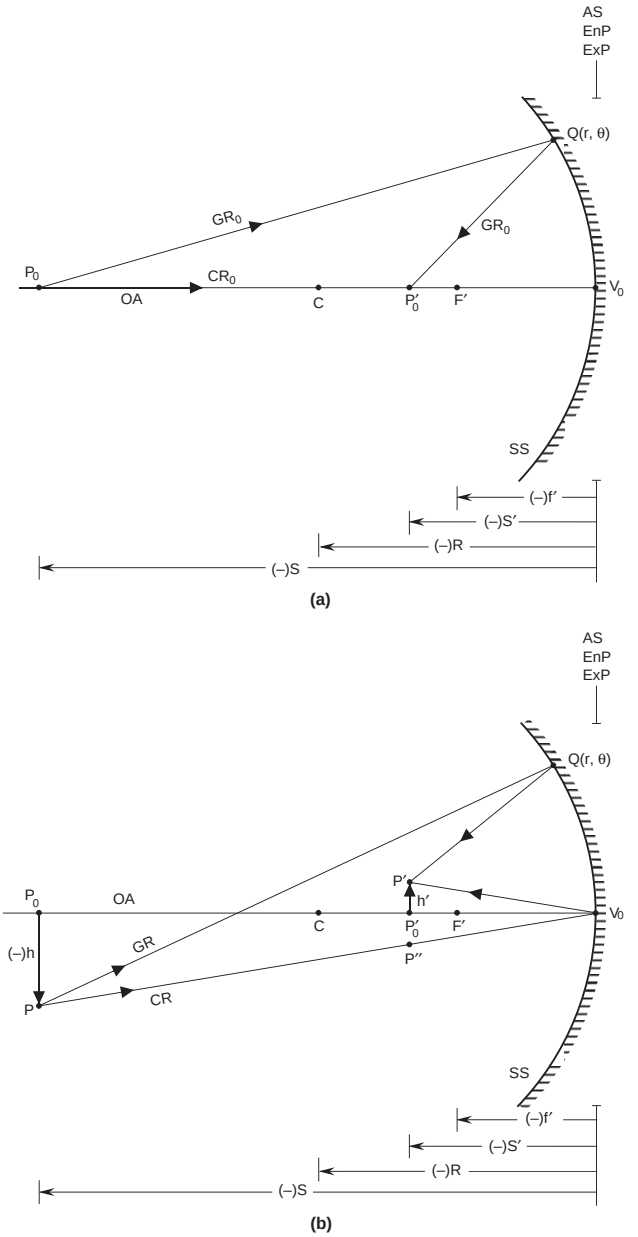


Figure 4-2. Same as Figure 4-1 except that the aperture stop is located at the mirror surface and axial imaging is also shown.

4.4 APERTURE STOP AT THE CENTER OF CURVATURE OF THE MIRROR

If the aperture stop is located at the center of curvature of the mirror as indicated in Figure 4-3, then the entrance pupil is also located there. The exit pupil, which is the image of the aperture stop by the mirror, is also located there, as may be seen by letting $S = R$ in Eq. (4-1). The distance L of the image from the exit pupil is numerically negative, since it lies to the right of the exit pupil. Accordingly,

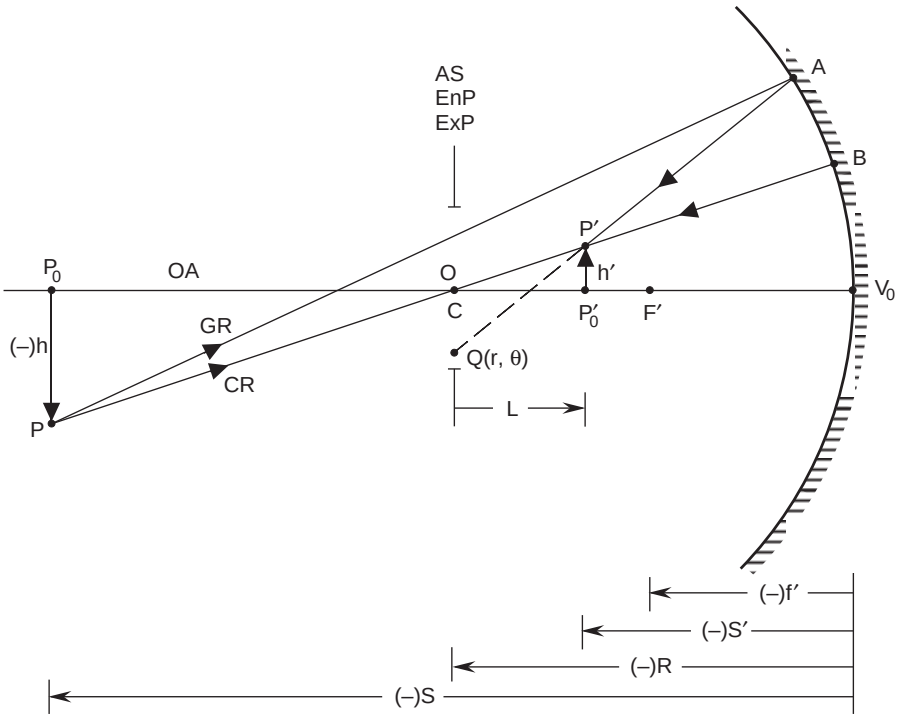


Figure 4-3. Same as Figure 4-1, except that the aperture stop is located at the center of curvature *C* of the mirror.

$$L = S' - R \quad . \tag{4-12}$$

and, therefore, Eqs. (4-5a) and (4-6) become

$$a_{ss} = \frac{S'^2}{4R^3(S'-R)^2} \quad . \tag{4-13}$$

and

$$d = 0 \quad . \tag{4-14}$$

respectively. Letting $d = 0$ in Eqs. (4-5b)–(4-5e), and substituting the results obtained into Eq. (4-3), we obtain the primary aberration function

$$W_s(r;h') = \frac{S'^2 r^4}{4R^3(S'-R)^2} - \frac{h'^2 r^2}{2R(S'-R)^2} \quad . \tag{4-15}$$

Thus, coma, astigmatism, and distortion of a spherical mirror with aperture stop at its center of curvature are zero. A concave mirror has negative spherical aberration but positive field curvature. If the image is observed on a spherical surface of radius of curvature $R/2$ at a distance S' from the mirror, then the second term on the right-hand side of Eq. (4-15) representing the field curvature also vanishes. The spherical image

surface is, of course, the Petzval image surface. It should be noted that, in going from Eq. (4-7) to Eq. (4-15), the maximum value of r, i.e., the radius of the exit pupil, has been multiplied by a factor of $(S-R)/S$ or $-(S'-R)/S'$. Hence, the peak value of spherical aberration has not changed, as expected, owing to a change in the position of the aperture stop.

For a point object at infinity

$$S' = R/2 \tag{4-16}$$

and, therefore,

$$L = -R/2 \quad , \tag{4-17}$$

and the spherical image surface of radius of curvature $R/2$ is concentric with the mirror. The spherical aberration is given by

$$a_{ss} = 1/4R^3 = 1/32f'^3 \quad ; \tag{4-18}$$

i.e., it is the same as for a mirror with aperture stop at its surface, as expected. It can be eliminated by placing, at the center of curvature of the mirror, a glass plate whose thickness varies as r^4. This indeed is the principle of the Schmidt system, which will be discussed in Chapter 5.

It is not difficult to see why all primary aberrations, except spherical, vanish when the aperture stop is located at the center of curvature of a spherical mirror and the image is observed on the Petzval surface. Since the exit pupil is also located at the center of curvature, the chief ray corresponding to an off-axis point object passes through it. Moreover, since the mirror is spherical, any line passing through its center of curvature forms the optical axis. Hence, every point object is like an on-axis object; therefore, the only aberration that arises (with respect to its Petzval image) is spherical aberration. The Petzval curvature, corresponding to the second term on the right-hand side of Eq. (4-5d), is nonzero. It has the implication that an image aberrated by spherical aberration alone is formed on a spherical surface of radius of curvature $R/2$. This, of course, is the Petzval image surface passing through the axial image point P_0'. It is concentric with the mirror when the object is at infinity.

4.5 NUMERICAL PROBLEMS

Now we consider simple numerical problems in which a spherical mirror of diameter 4 cm and a radius of curvature 10 cm images an object 2 cm high located at a distance of 15 cm from it. We assume that the aperture stop is located at the mirror and the object lies below the optical axis. Table 4-1 gives the Gaussian as well as the aberration parameters for this problem. Both concave and convex mirrors are considered in this table. The concave mirror forms a real image, but the convex mirror forms a virtual image. We note that whereas astigmatism is the dominant primary aberration in the case of the concave

Table 4-1. Gaussian and aberration parameters for a spherical mirror of radius *a* imaging an object lying at a finite distance from it. The aperture stop is located at the mirror.

Gaussian Parameters

Mirror	R (cm)	S' (cm)	h' (cm)	F	d
Concave	–10	–7.5	1	$7.5/4$	–4
Convex	10	3.75	-0.5	3.75/4	–1.6

Aberration Parameters

Mirror	a_{ss} $\left(\text{cm}^{-3}\right)$	A_{ss} (μm)	A_{cs} (μm)	A_{as} (μm)
Concave	-2.78×10^{5}	-4.4	35.56	–71.1
Convex	6.94×10^{4}	111	178	71.1

S = –15 cm, h = –2 cm, a = 2 cm, $S' = -RS/(2S-R)$

$F = |S'|/2a$, $d = R/(S'-R)$

mirror, it is coma that dominates in the case of the convex mirror. Field curvature and distortion are zero in both cases, since the aperture stop lies at the mirror surface.

Table 4-2 lists the Gaussian and aberration parameters for an object lying at infinity at an angle of 1 milliradian from the optical axis of the mirror. The magnitude of a primary aberration is independent of whether the mirror is concave or convex, but its sign depends on its type. Spherical aberration is the dominant aberration in Table 4-2. Of course, the field curvature and distortion are zero once again.

If the aperture stop of the mirror is moved to its center of curvature, the peak value A_s of its spherical aberration does not change. However, its coma and astigmatism reduce to zero, but its field curvature becomes nonzero. The radius of the exit pupil a_{ex}, the field curvature coefficient a_d, and the peak value of field curvature for the problems under consideration are given in Table 4-3. The numbers without parentheses are for an object at S = 15 cm, and those with parentheses are for an object at infinity at 1 milliradian from the optical axis of the system. As a reminder, we add that

Table 4-2. Gaussian and aberration parameters for a spherical mirror imaging an object lying at infinity at an angle of 1 milliradian from its optical axis. The aperture stop is located at the mirror.

Gaussian Parameters					
Mirror	R **(cm)**	S' **(cm)**	h' **(cm)**	F	d
Concave	–10	–5	5×10^{-3}	1.25	–2
Convex	10	5	-5×10^{-3}	1.25	–2

Aberration Parameters				
Mirror	a_{ss} $\left(\text{cm}^{-3}\right)$	A_{ss} (μm)	A_{cs} (μm)	A_{as} (μm)
Concave	-2.5×10^{-4}	-40	0.8	-4×10^{-3}
Convex	-2.5×10^{-4}	40	0.8	4×10^{-3}

Table 4-3. Radius of the exit pupil and field curvature parameters for a spherical mirror when the aperture stop is located at its center of curvature.*

Mirror	a_{ex} **(cm)**	a_d $\left(\text{cm}^{-3}\right)$	A_d (μm)
Concave	$2/3$	8×10^{-3}	35.6
	(2)	$\left(2\times10^{-3}\right)$	$\left(2\times10^{-3}\right)$
Convex	$10/3$	-1.28×10^{-3}	–35.6
	(2)	$\left(-2\times10^{-3}\right)$	$\left(-2\times10^{-3}\right)$

* The numbers without parentheses are for an object at S = 15 cm and those with parentheses are for an object at infinity at 1 milliradian from the optical axis of the system.

$$a_{ex} = a\left|(S-R)/S\right| = a\left|(S'-R)/S'\right| \quad , \tag{4-19}$$

$$a_d = -1/2R(S'-R)^2 \quad , \tag{4-20}$$

and

$$A_d = a_d h'^2 a_{ex}^2 \quad , \tag{4-21}$$

where $a = 2$ cm is the radius of the mirror. The field curvature as an aberration disappears when the image is observed on a spherical surface of radius of curvature –5 cm for the concave mirror and 5 cm for the convex mirror, located at the image plane. When the object is located at infinity, this surface is concentric with the mirror.

4.6 SUMMARY

The imaging properties of a spherical mirror can be obtained from those for a corresponding spherical refracting surface by letting the refractive index of the object space be 1 and that of the image space be –1 (representing reflection). It is worth remembering that the spherical aberration of a spherical mirror is not zero. A paraboloidal mirror, for example, gives zero spherical aberration for an axial point object at infinity. An ellipsoidal mirror, however, is required for an aberration-free image of a point object at a finite distance. As in the case of lenses, two mirrors are required for aplanatic imaging.

How the aberrations of a system change as the position of its aperture stop is changed is illustrated in this chapter by considering the images formed by a spherical mirror. Only the field curvature and distortion are zero when the aperture stop lies at the mirror. However, only spherical aberration and field curvature are nonzero when the stop lies at the center of curvature of the mirror.

CHAPTER 5
Schmidt Camera

5.1 INTRODUCTION

We have seen in Chapter 4 that a spherical mirror gives spherical aberration, which we know from Section 1.7 to be independent of the location of its aperture stop. When the aperture stop is located at the center of curvature of the mirror, it also produces field curvature, although coma, astigmatism, and distortion are all zero. As we will discuss in Chapter 6, a paraboloidal mirror forms an aberration-free image of a point object only when it lies on its axis at an infinite distance from it. In order to utilize the simplicity of fabrication of a spherical mirror, we need a way to compensate its spherical aberration. An optical system consisting of a spherical mirror and a transparent plate of nonuniform thickness placed at its center of curvature to compensate for its spherical aberration is called a *Schmidt camera*. The plate is appropriately called the *Schmidt plate.* With the exception of field curvature, the image formed is free of primary aberrations.

As discussed in Section 4.4, the field curvature is such that an aberration-free image is formed on a spherical surface of radius of curvature equal to half that of the mirror. For an object at infinity, this surface is concentric with the mirror. In this chapter, we determine the shape of the Schmidt plate and discuss the chromatic aberrations associated with it. A numerical problem illustrates the results obtained.

5.2 SCHMIDT PLATE

Consider a spherical mirror with its aperture stop located at its center of curvature *C*, as shown in Figure 5-1, imaging an object lying at infinity. From Eq. (4-18), the optical path difference between a ray of zone *r* and the chief ray from an axial point object is given by

$$W(r) = \frac{r^4}{32 f'^3} \quad , \tag{5-1}$$

where f' is the focal length of the mirror. It is negative, implying that the optical path length of the ray under consideration to the focus F' is shorter than that of the chief ray. It also means that the ray intersects the axis after reflection at an axial point F'', which is slightly closer to the mirror vertex than the paraxial focus F'. This may be seen independently from the isosceles triangle CAF'' in which $CF'' + AF'' > CA = 2|f'|$ or $CF'' > |f'| = CF'$, since $CF'' = AF''$. In order that the path length of the ray be equal to that of the chief ray, its path length must be increased.

If a plate of refractive index n and a thickness $t(r)$ is placed at the center of curvature with a flat surface normal to the axis of the mirror, the additional optical path length introduced by the plate is given by $(n-1)\,t(r)$. All object rays transmitted by the

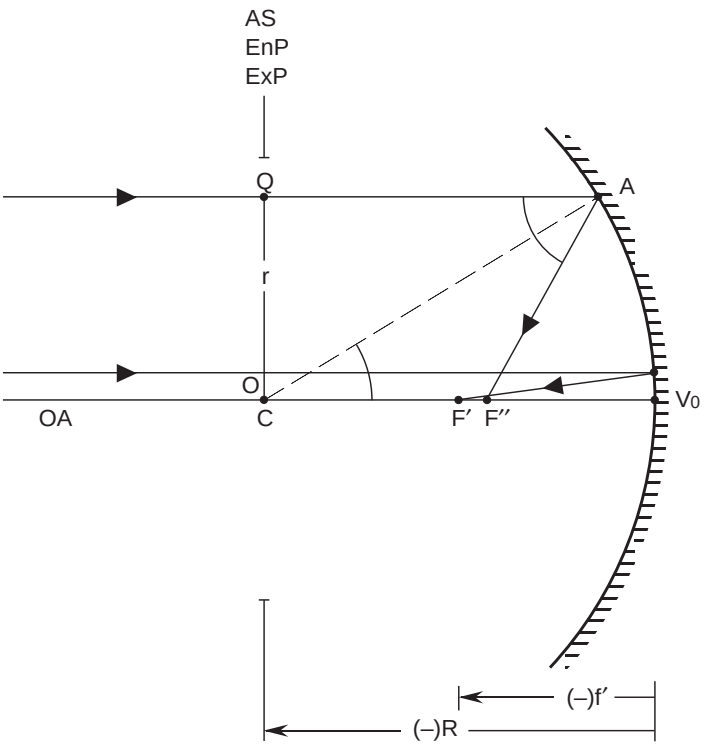


Figure 5-1. Imaging by a spherical mirror with aperture stop located at its center of curvature. Rays of different zones from an axial object at infinity intersect the axis of the mirror after reflection at different points, such as F' and F'', thus forming an image aberrated by spherical aberration. The ray shown intersecting the axis at F'' has a zone of $\sqrt{3}a/2$, where a is the radius of the aperture stop.

system travel equal optical path lengths and converge to a common focus F' if $t(r)$ is given by

$$W(r) + (n-1)t(r) = 0 \quad . \tag{5-2}$$

Substituting Eq. (5-1), we find that the plate thickness is given by

$$t(r) = -\frac{r^4}{32(n-1)f'^3} \quad . \tag{5-3}$$

It increases from a value of zero at its center to values proportional to the fourth power of the zonal radius. In practice, a plane-parallel plate of constant thickness t_0 would be added to it so that it can be fabricated. The shape of the plate is shown in Figure 5-2, where it is shown to slightly tilt a nonaxial ray so that, after reflection by the mirror, it passes through F'.

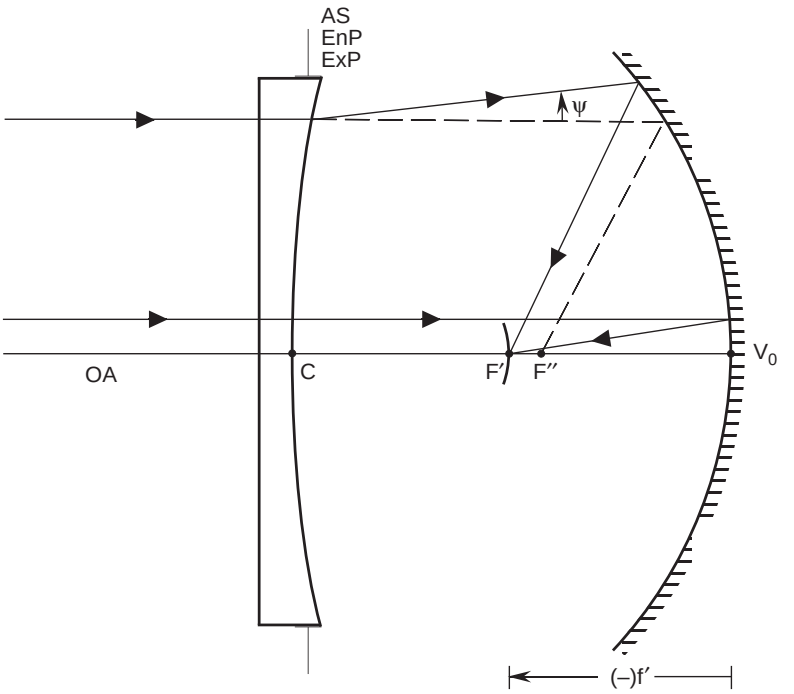


Figure 5-2. Imaging by a Schmidt camera consisting of a spherical mirror and a transparent plate placed at its center of curvature *C*. The spherical aberration of the mirror is precorrected by the plate so that the system forms an image free of spherical aberration. Dashed lines indicate the path of a ray in the absence of a Schmidt plate.

Although spherical aberration is corrected by the use of such a plate, it does introduce *chromatic aberration*. Since the refractive index of the plate varies with the wavelength of object radiation, the angular deviation of a ray produced by the plate also varies with it. Consider a ray corresponding to a refractive index n and passing through the plate at a zone r. Since the plate is located in air and the wave aberration produced by it is $(n-1)\,t(r)$, following Eq. (1-1), the angular deviation of the ray produced by it is given by

$$\psi = (n-1)\frac{dt}{dr} \quad . \tag{5-4}$$

Substituting Eq. (5-3) into Eq. (5-4), we obtain

$$\psi = -\frac{r^3}{8f'^3} \quad . \tag{5-5}$$

From Eq. (5-4), the *angular dispersion* of any ray is given by

$$\Delta\psi = \Delta n\frac{dt}{dr} \quad , \tag{5-6a}$$

where Δn is the variation in the refractive index of the plate across the spectral bandwidth of the object radiation. Substituting for dt/dr from Eq. (5-4) into Eq. (5-6a), we obtain

$$\Delta\psi = \frac{\Delta n}{n-1}\psi \quad . \tag{5-6b}$$

Thus, the angular dispersion $\Delta\psi$ of a ray produced by the plate is proportional to its angular deviation ψ. The value of ψ is maximum and equal to $-a^3/8f'^3$ for the marginal rays, i.e., for $r = a$, where a is the radius of the plate.

To reduce the chromatic aberration, we must reduce the maximum value of ψ. To do so, we add to the plate a very thin plano-convex lens. Such a lens will reduce the focus distance such that the rays are now focused at a point F'' instead of F', as in Figure 5-3.

A plano-convex lens introduces thickness to the plate varying as r^2. Thus, the plate thickness may be written

$$t(r) = t_0 - \frac{r^4}{32(n-1)f'^3} + \frac{br^2}{n-1} \quad , \tag{5-7}$$

where b is a constant chosen to minimize the chromatic aberration. Comparing the defocus aberration br^2 introduced by the plate with Eq. (1-3b), we find that the distance between F' and F'' is given by $2bf'^2$. F'' lies on the right-hand side of F', as in Figure 5-3, if b is numerically negative. The thickness variations of plates with different values of b are shown in Figure 5-4. We note that the depth of *material removal*, starting with a plane-parallel plate, is minimum when $b = a^2/32f'^3$ (corresponding to $c = 1$ in the figure). However, we are interested in minimizing the maximum value of the angular deviation of a ray. As shown below, this requires that b be equal to $3a^2/64f'^3$ (or $c = 1.5$).

Substituting Eq. (5-7) into Eq. (5-4), we find that the angular deviation of a ray is now given by

$$\psi = \frac{r^3}{8f'^3} + 2br \quad . \tag{5-8}$$

Its maximum value in the range $0 \le r \le a$ occurs either at its stationary point $r = \left(16f'^3b/3\right)^{1/2}$ obtained by letting $\partial\psi/\partial r = 0$ or at $r = a$. At the former its absolute value is $(4|b|/3)\left(16f'^3b/3\right)^{1/2}$ and at the latter it is $\left(-a^3/8f'^3\right) + 2ba$. These two values are both equal to $a^3/32f'^3$ if $b = 3a^2/64f'^3$ (or $c = 1.5$), thus reducing the angular deviation as well as the chromatic aberration by a factor of 4 compared to their values if $b = 0$.

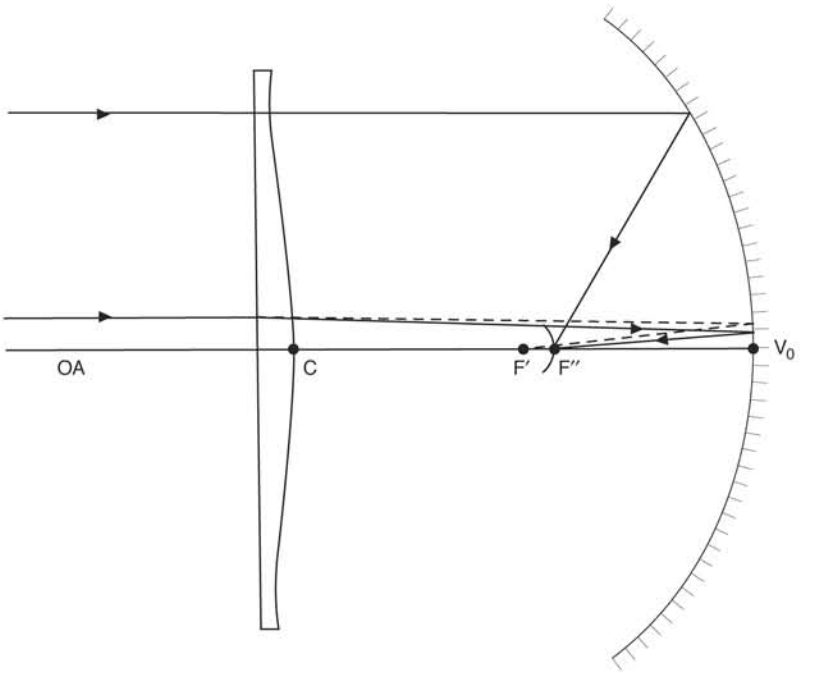


Figure 5-3. Schmidt camera with a plate introducing minimum chromatic aberration. The dashed lines indicate the path of a ray in the absence of the Schmidt plate. All rays passing through the plate and reflected by the mirror are focused at F'', where the ray passing through the neutral zone of the plate is focused by the mirror.

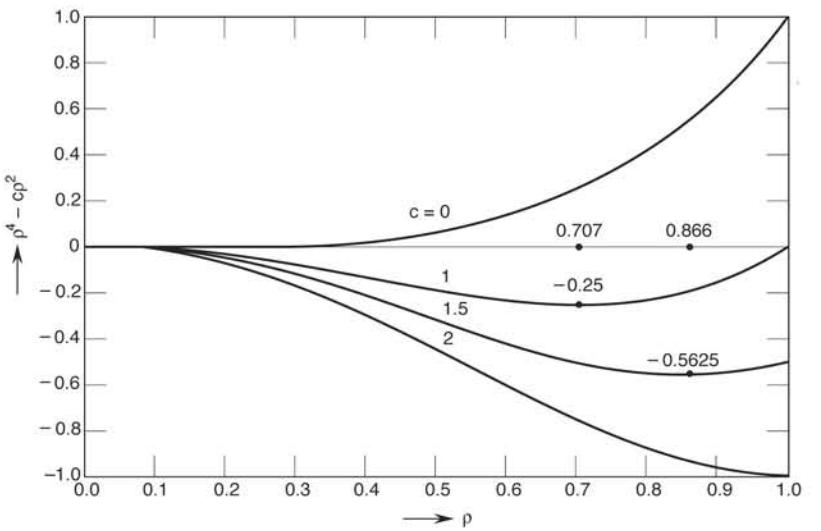


Figure 5-4. Thickness variation of a Schmidt plate for different values of b, where $c = 32f'^3b/a^2$ and $\rho = r/a$. The variation is minimum when $b = a^2/32f'^3$, corresponding to $c = 1$.

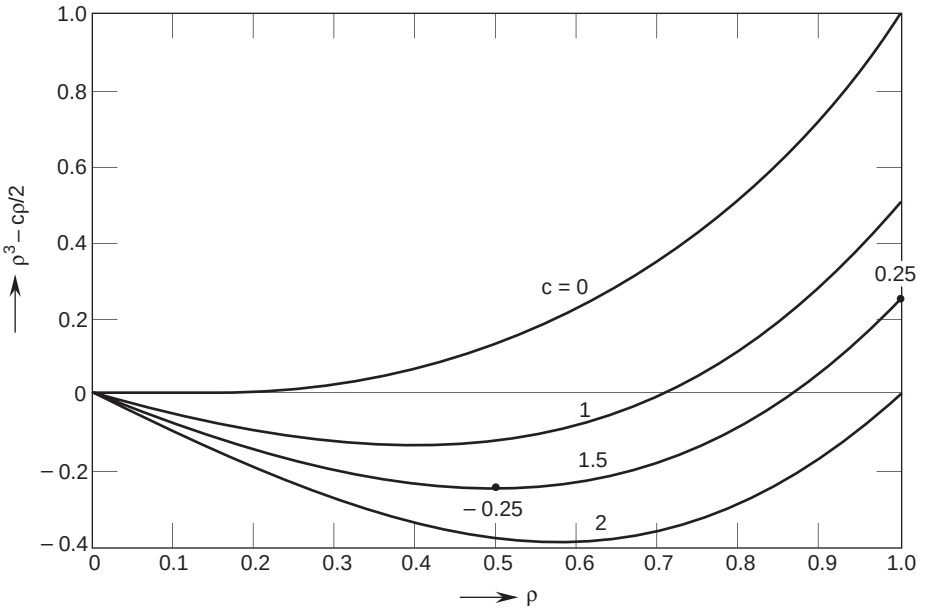


Figure 5-5. Dependence of angular dispersion on the value of *b*. It is minimum when $b = 3a^2/64f'^3$ corresponding to $c = 1.5$.

Substituting this value of b into Eq. (5-7) we find that the plate thickness required for eliminating spherical aberration introduced by the mirror and minimizing the chromatic aberration introduced by the plate is given by

$$t(r) = t_0 - \frac{1}{32(n-1)f'^3}\left(r^4 - \frac{3}{2}a^2r^2\right) \quad . \tag{5-9}$$

We note that $\partial\psi/\partial r = 0$ for $r = \sqrt{3}a/2$. This value of r is called the *neutral zone* of the plate since a ray incident normal to it passes through it undeviated as in Figure 5-3. As may be seen from Figure 5-4, the thickness variation of the plate and the material removal are maximum at this zone. This variation is more than twice the variation for a minimum-thickness-variation plate; compare the numbers –0.5625 and –0.25 in the figure that occur at zones of $r = 0.707a$ and $r = 0.866a$, respectively.

The lens component of the Schmidt plate has a focal length of $f_l = -32\,f'^3/3a^2$. The vertex radius of curvature of the plate is equal to $(1-n)f_l$. This, of course, is the radius of curvature of the second surface of the lens. The angular dispersion of the rays is now given by

$$\Delta\psi = -\frac{\Delta n}{8(n-1)f'^3}\left(r^3 - \frac{3}{4}a^2r\right) \quad . \tag{5-10}$$

Its maximum value occurs for rays with $r = a/2$ and a. It is given by

$$[\Delta\psi]_{max} = -\frac{\Delta n}{32(n-1)}\frac{a^3}{f'^3} \quad . \tag{5-11}$$

The dependence of angular dispersion on the value of *b* is illustrated in Figure 5-5. We shall see in Chapter 7 that the value of *b* giving minimum chromatic aberration also gives the position of the defocused image plane in which the rays forming an image of a point object aberrated by spherical aberration have a minimum spot radius (circle of least confusion).

It has been shown in Section 4.4 that a spherical mirror with an aperture stop located at its center of curvature gives only spherical aberration and field curvature. The Schmidt plate compensates for the spherical aberration and, therefore, the image of an extended object observed on a spherical surface concentric with the mirror is free of primary aberrations. Strictly speaking, the lens component of the plate also introduces small amounts of primary aberrations. The spherical aberration contributed by it can be made zero by slightly adjusting the value of the r^4 term in the plate thickness $t(r)$. The mirror also contributes some *secondary* or sixth-order spherical aberration. It can be made zero by introducing an r^6 term in the plate thickness.

It should be noted that as the field angle β increases, the size of the focal surface also increases, which, in turn, obscures the ray bundle incident on the mirror. For a *field of view* of radius β, the linear obscuration of the on-axis beam incident on the mirror is given by $\epsilon = 2\beta F$, where F is the focal ratio of the system.

5.3 NUMERICAL PROBLEMS

As a numerical example, we consider a spherical mirror with a radius a = 5 cm and a focal length $f' = -40$ cm so that F = 4. According to Eq. (5-1), the peak value of spherical aberration introduced by it for an object at infinity is equal to $3.05\,\mu m$. If a Schmidt plate of refractive index $n = 1.5$ is used to compensate for this spherical aberration, the difference in its maximum and minimum thickens is 6.10 μm according to Eq. (5-3). Thus, starting with a plate of uniform thickness, as much as $6.1\,\mu m$ deep material must be removed at its center, reducing to a value of zero at its edge. This would be satisfactory for operation in monochromatic light for which the refractive index is 1.5. The image is formed at a distance of 40 cm from the mirror. The image of an extended object lying at infinity is free of primary aberrations when observed on a spherical surface of radius of curvature 40 cm concentric with the mirror passing through its focal point F'.

For white-light operation, the thickness variation of the plate for minimum chromatic aberration is given by Eq. (5-9). Thus, the plate has a certain thickness at the center, and its variation is maximum and equal to 3.43 μm at its neutral zone of $\sqrt{3}a/2$ = 4.33 cm. Its variation at its edge is 3.05 μm. We note that the depth of material removal is less for this plate than that for the monochromatic operation. The image is now formed at a distance that is 0.586 mm closer to the mirror than its focal plane. If Δn = 0.025 across the spectral bandwidth of object radiation, then, according to Eq. (5-11), the minimum

radius of the chromatic image will be 1.22 μm. In practice, the image will be larger than this due to diffraction by, say, the diameter of the Airy disc (discussed in Chapter 8). For visible light, the diameter of the Airy disc (neglecting the effect of obscuration due to the focal surface) is approximately 6.83 μm, where we have used a visible wavelength of 0.7 μm.

5.4 SUMMARY

In a Schmidt camera, the spherical aberration of a spherical mirror is made zero by placing a corrector plate (called the Schmidt plate) at its center of curvature. The aperture stop and the entrance and exit pupils also lie at this location, yielding zero astigmatism and coma, and thus providing an anastigmatic system. The plate thickness described by Eq. (5-3) and illustrated in Figure 5-2 introduces dispersion, which is minimized by modifying the plate thickness to that described by Eq. (5-9) and shown in Figure 5-3.

CHAPTER 6
Aberrations of a Conic Surface

6.1 INTRODUCTION

So far, we have considered the aberrations of spherical surfaces, which are conic surfaces of zero eccentricity. In this chapter, we discuss the aberrations of a *conic surface* with an arbitrary value of *eccentricity*. Our starting point is imaging by and aberrations of a spherical surface discussed in Sections 1.8 and 4.2. It should be noted that the Gaussian imaging equations for a conic surface of a certain *vertex radius of curvature* are the same as those for a spherical surface of the same radius of curvature. Given the aberrations of a spherical surface, we determine the additional aberrations introduced by a corresponding conic surface. In particular, we show that if the aperture stop is located at the conic surface, the only additional aberration is spherical aberration. The other (primary) aberrations of the conic surface are identical to those of the spherical surface. The aberrations of a conic surface are further generalized to obtain the aberrations of a *general aspherical* (*nonconic*) *surface.* The aberrations of a *paraboloidal mirror* are briefly discussed and compared with those of a spherical mirror. Finally, we outline a procedure to determine the aberrations of a multimirror system.

6.2 CONIC SURFACE

A conic surface of eccentricity e and vertex radius of curvature R is described by its *sag* according to

$$z_c = \frac{r_c^2 / R}{1 + \left[1 - \left(1 - e^2\right) r_c^2 / R^2\right]^{1/2}}, \tag{6-1}$$

where, as illustrated in Figure 6-1, $\left(x_c, y_c, z_c\right)$ are the coordinates of a point on it and

$$r_c = \left(x_c^2 + y_c^2\right)^{1/2} \tag{6-2}$$

is the distance of the point from the z axis. The origin of the coordinate system is at the vertex of the conic, and the z axis is along its axis of rotational symmetry. The various conic surfaces are described by their values of e according to

e	$= 1$	Paraboloid
	< 1	Ellipsoid
	> 1	Hyperboloid
	$= 0$	Sphere .

If we neglect the terms in r_c of an order higher than four, Eq. (6-1) becomes

$$z_c = \frac{r_c^2}{2R} + \left(1 - e^2\right) \frac{r_c^4}{8R^3} \quad . \tag{6-3}$$

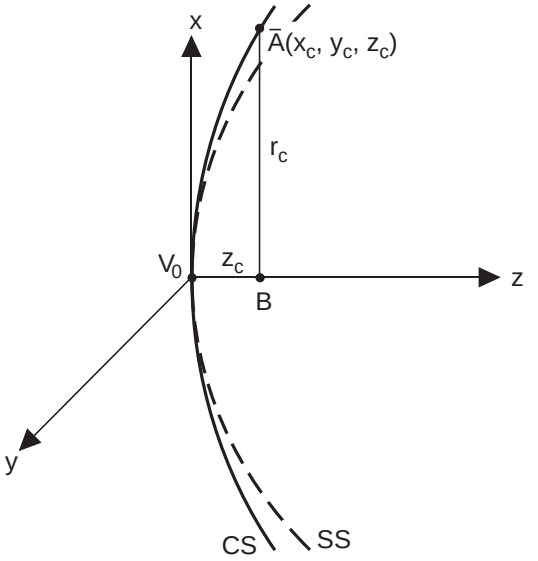


Figure 6-1. Sag of a conic surface. The origin of the coordinate system lies at the vertex V_0 of the conic. The axis about which the conic is rotationally symmetric is the *z* axis of the coordinate system. z_c is the sag of a point *A* on the conic.

Thus, up to the fourth order in r_c, the sag of a spherical $(e = 0)$ surface is larger than that of a conic surface by $e^2 r_c^4 / 8R^3$. Up to this order, the chord $V_0 A \simeq r_c$.

6.3 CONIC REFRACTING SURFACE

6.3.1 On-Axis Point Object

Consider a conic surface separating media of refractive indices *n* and *n'*. Compared to a spherical surface, a conic surface introduces an *additional* aberration for an axial point object P_0, which for a ray passing through a point *A* on the spherical surface in Figure 6-2 is given by

$$\Delta W_c\left(\overline{A}_0\right) \simeq (n' - n)\overline{A}_0 A \quad , \tag{6-4a}$$

where

$$\overline{A}_0 A \simeq e^2 r_c^4 / 8R^3 \tag{6-4b}$$

is approximately equal to the sag difference between a sphere and a conic of the same vertex radius of curvature. Since $V_0 A \simeq r_c$, we may write

$$\Delta W_c\left(\overline{A}_0\right) = \sigma V_0 \overline{A}_0^4 \quad , \tag{6-5a}$$

where

$$\sigma = (n' - n) e^2 / 8R^3 \quad . \tag{6-5b}$$

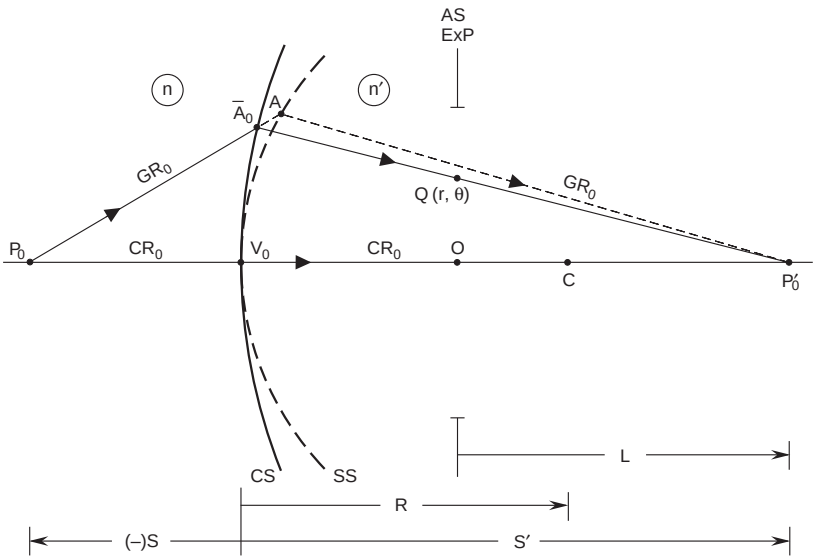


Figure 6-2. Imaging of an on-axis point object P_0 by a conic refracting surface *CS* of vertex radius of curvature *R* and center of curvature *C*. The Gaussian image is located at P_0'.

We note from (approximate) triangles $V_0\bar{A}_0P'$ and OQP_0' in the figure that $V_0A/OQ = S'/L$. Hence, the aberration at a point Q in the plane of the exit pupil at a distance r from the optical axis is given by

$$\Delta W_c(Q) = \sigma(S'/L)^4 OQ^4 \quad ,$$

or

$$\Delta W_c(r) = \sigma(S'/L)^4 r^4 \quad . \tag{6-6}$$

6.3.2 Off-Axis Point Object

For an off-axis point object such as *P* in Figure 6-3, the optical path length of the chief ray for a conic surface is also different from that for a spherical surface. Accordingly, the conic contribution to the aberration of a ray from the point object *P* and passing through a point $\bar{A}$ on the conic surface is given by

$$\begin{aligned}\Delta W_c(\bar{A}) &\simeq (n'-n)(\bar{A}A - \bar{B}B) \\ &= \sigma(V_0\bar{A}^4 - V_0\bar{B}^4) \quad .\end{aligned} \tag{6-7}$$

Let (r, θ) be the polar coordinates of a point *Q*, where the ray under consideration intersects the plane of the exit pupil, with respect to *O* as the origin. From Figure 6-4, which represents the projection of the exit pupil on the refracting surface, we note that

$$V_0\bar{A}^2 = \bar{A}\bar{B}^2 + V_0\bar{B}^2 - 2\bar{A}\bar{B}\,V_0\bar{B}\cos\theta \quad . \tag{6-8}$$

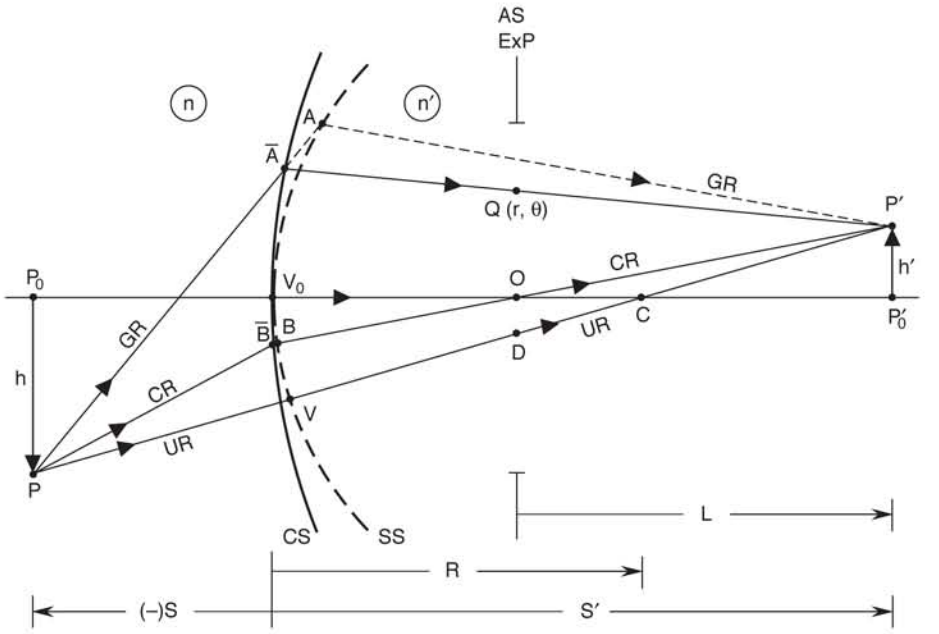


Figure 6-3. Imaging of an off-axis point object *P* by a conic refracting surface of vertex radius of curvature *R* and center of curvature *C*. The Gaussian image is located at *P′*.

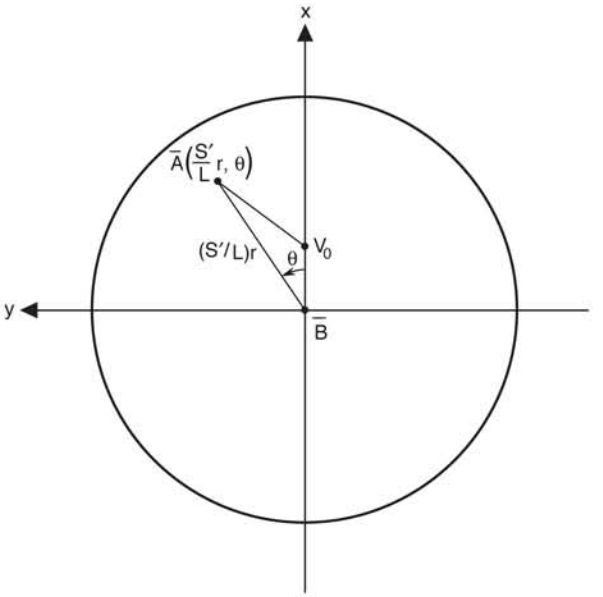


Figure 6-4. Projection of the exit pupil on the refracting surface in Figure 6-3. Point *B*, which lies on the chief ray, forms the center of the projected pupil.

Also, from (approximate) triangles $B\overline{A}P'$ and OQP' in Figure 6-3, we note that

$$\overline{A}\,\overline{B} \simeq (S'/L)r \tag{6-9a}$$

Similarly, from (approximate) triangles $OV_0\overline{B}$ and $OP_0'P'$,

$$V_0\overline{B} \simeq gh' \quad , \tag{6-9b}$$

where

$$g = \frac{S'-L}{L} \quad . \tag{6-10}$$

Substituting Eqs. (6-9a), (6-9b), and (6-10) into Eq. (6-8), squaring the result, and then substituting into Eq. (6-7), we obtain

$$\begin{aligned}\Delta W_c(Q) = \sigma\big[&(S'/L)^4 r^4 - 4(S'/L)^3 gh'r^3\cos\theta + 4(S'/L)^2 g^2h'^2r^2\cos^2\theta \\ &+ 2(S'/L)^2 g^2h'^2r^2 - 4(S'/L)g^3h'^3r\cos\theta\big] \quad .\end{aligned} \tag{6-11}$$

Adding Eqs. (6-11) and (1-19), we obtain the primary aberrations of a conic surface. Note that if the aperture stop is located at the conic surface so that $L = S'$ and, in turn, $g = 0$, then its aberrations differ from those of a spherical surface only in spherical aberration. The other primary aberrations are identical for the two surfaces.

6.4 GENERAL ASPHERICAL REFRACTING SURFACE

Consider a *general* rotationally symmetric *aspherical surface* with a vertex radius of curvature R and described by its sag according to

$$z_g = \frac{r_g^{\,2}}{2R} + \left(1 - e^2 + s_g\right)\frac{r_g^{\,4}}{8R^3} \quad , \tag{6-12}$$

where s_g represents the fourth-order sag contribution over and above that of a conic surface of the same vertex radius of curvature. Compared to a spherical surface of radius of curvature R, the general aspherical surface contributes an additional optical path length to a zonal ray from an axial point object given by

$$\Delta W_g(A) = \sigma' V_0 A^4 \quad , \tag{6-13a}$$

where

$$\sigma' = (n'-n)\left(e^2 - s_g\right)\Big/8R^3 \quad . \tag{6-13b}$$

Comparing Eqs. (6-5) and (6-13), we note that if we replace σ with σ' in the results

obtained for a conic surface, we obtain the aberrations for a general aspherical surface. This conclusion applies to the aberrations for off-axis point objects as well.

6.5 CONIC REFLECTING SURFACE

The additional primary aberrations introduced by a conic or a general aspherical *reflecting* surface, compared to a spherical one (discussed in Chapter 4), can be obtained from those for a corresponding refracting surface by letting $n' = -n = 1$. Thus, for example, the additional aberration of a *conic mirror* is given by Eq. (6-11), where

$$\sigma = -e^2/4R^3 \quad . \tag{6-14}$$

6.6 PARABOLOIDAL MIRROR

For a paraboloidal (e = 1) mirror, we note from Eqs. (4-4) and (6-14) that for an object at infinity,

$$a_s = -\sigma = 1/R^3 \quad . \tag{6-15}$$

Hence, following Eqs. (4-3) and (4-11), we find that the spherical aberration of a paraboloidal mirror is zero when the object lies at infinity, i.e.,

$$a_{sc} = (S'/L)^4 (a_s + \sigma) = 0 \quad . \tag{6-16}$$

If, in addition, the aperture stop is located at the mirror, then $L = S'$ and, therefore, g = 0. Hence, Eq. (6-11) shows that the other primary aberrations of a paraboloidal mirror are identical with those for a spherical mirror. Accordingly, the image of an off-axis object at infinity formed by a paraboloidal mirror with stop at the mirror surface suffers only from coma and astigmatism given by the corresponding terms in Eq. (4-11). Thus, for example, the image of an object lying at infinity at an angle of 1 milliradian from the axis of a paraboloidal mirror of vertex radius of curvature of 10 cm suffers from coma aberration with a peak value of 0.8 μm but a negligible value of astigmatism. These values are the same as those of the corresponding aberrations in Table 4-2 where imaging by a spherical mirror was considered. Thus, the difference between imaging by paraboloidal and spherical mirrors lies in their spherical aberrations: zero in the case of a paraboloidal mirror and -40 μm peak aberration, for example, in the case of a concave spherical mirror. Of course, the image of an axial object at infinity by a paraboloidal mirror is aberration free.

6.7 MULTIMIRROR SYSTEMS

The aberrations of a multielement system can be calculated by determining the aberrations of each element at its respective exit pupil and then combining them according to the procedure described in Section 1.9. Thus, for example, we can show that an afocal system consisting of two confocal paraboloidal mirrors acting as a beam expander is anastigmatic, introducing only field curvature and distortion aberrations. Similarly, we can investigate the aberrations of two-mirror systems such as the

Cassegrain and *Ritchey-Chrétien* telescopes. As is sometimes the case in practice, this is easier said than done.

6.8 SUMMARY

The Gaussian imaging equations for a conic surface of a certain vertex radius of curvature are the same as those for a spherical surface of the same radius of curvature. The aberrations of a conic surface with its aperture stop located at the surface are the same as those for a corresponding spherical surface, except for spherical aberration. The aberrations of a general aspherical (nonconic) surface can be obtained from those of a conic surface by simply modifying its sag contribution, as in Eq. (6-13b). The aberrations of a multielement system can be obtained by the procedure described in Section 1.9. As a reminder, note that only the primary aberrations are discussed in this book, and the actual aberrations of a system are obtained by tracing rays or by measurement.

CHAPTER 7

Ray Spot Sizes and Diagrams

7.1 INTRODUCTION

In Chapters 2–6, we have determined the primary wave aberrations of simple optical imaging systems. In this chapter, we use the relationship between the wave and ray aberrations given in Section 1.2 to determine the ray distribution for a point object, called the *ray spot diagrams*, in the Gaussian image plane. For each primary aberration, we determine the extent or the size of the image spot in terms of its peak value and the focal ratio of the image-forming light cone. In the case of spherical aberration and astigmatism, we consider the ray distributions in defocused image planes as well, and we determine the plane in which the spot size is minimum. These minimum-size spots are referred to as the *circles of least confusion* and represent the best aberrated images based on geometrical optics.

We define the centroid and the standard deviation of the ray distribution and calculate them for the primary aberrations. In lens design, one often tries to minimize the spot sigma rather than the spot radius. However, we will see in Chapter 8 that, in reality, which is based on diffraction of light at the exit pupil of the system, an image distribution is not given by the corresponding ray spot diagrams. For example, the aberration-free image of a point object is a point according to geometrical optics, but its diffraction image for a circular pupil consists of a bright spot surrounded by concentric dark and bright rings. Even so, it is quite common practice in lens design to look at the spot diagrams in the early stages of a design, as discussed in Section 7.10.

7.2 WAVE AND RAY ABERRATIONS

Consider an optical system consisting of a series of rotationally symmetric coaxial refracting and/or reflecting surfaces imaging a point object. We have discussed in Chapter 1 that the primary aberration function representing the *wave aberration* at its exit pupil can be written

$$W(r,\theta;h') = a_s r^4 + a_c h' r^3 \cos\theta + a_a h'^2 r^2 \cos^2\theta + a_d h'^2 r^2 + a_t h'^3 r \cos\theta \quad , \qquad (7\text{-}1)$$

where (r,θ) are the polar coordinates of a point in the plane xy of the exit pupil of the system, h' is the height of the Gaussian image point P', and a_s, a_c, a_a, a_d, and a_t represent the coefficients of *spherical aberration*, *coma*, *astigmatism*, *field curvature*, and *distortion*, respectively. The angle θ is equal to zero or π for points lying in the *tangential* or *meridional plane* (i.e., the zx plane containing the optical axis and the point object and, therefore, its Gaussian image). The *chief ray*, which, by definition, passes through the center of the exit pupil, always lies in this plane. The plane normal to the tangential plane but containing the chief ray is called the *sagittal plane*. As the chief ray bends when it is refracted or reflected by a surface, so does the sagittal plane.

For an optical system with a circular exit pupil, say, of radius a, it is convenient to use normalized coordinates (ρ, θ) where $\rho = r/a,\ 0 \le \rho \le 1, 0 \le \theta \le 2\pi$, suppress the explicit dependence on h', and write the aberration function in the form

$$W(\rho,\theta) = A_s\rho^4 + A_c\rho^3\cos\theta + A_a\rho^2\cos^2\theta + A_d\rho^2 + A_t\rho\cos\theta \quad , \tag{7-2}$$

where the new aberration coefficients A_i, representing the peak or maximum values of the aberrations, are related to those used in Eq. (7-1) according to

$$A_s = a_s a^4,\ A_c = a_c h' a^3,\ A_a = a_a h'^2 a^2,\ A_d = a_d h'^2 a^2,\ A_t = a_t h'^3 a \quad . \tag{7-3}$$

Although we will discuss the spot diagrams in terms of these *peak aberration coefficients,* it is necessary to know their dependence on the image height h' when discussing images of extended objects.

If (x, y) represent the rectangular coordinates of a pupil point, the corresponding normalized coordinates (ξ, η) are given by

$$(\xi,\eta) = \frac{1}{a}(x,y) \tag{7-4a}$$

$$= \rho(\cos\theta, \sin\theta) \quad , \tag{7-4b}$$

where $-1 \le \xi \le 1, -1 \le \eta \le 1,$ and $\xi^2 + \eta^2 = \rho^2 \le 1$. The aberration function defined in the form of Eq. (7-2) has the advantage that the aberration coefficients A_i have the dimensions of length (i.e., dimensions of the wave aberration), and they represent the peak or maximum values of the corresponding primary aberrations. For example, if $A_s = 1\lambda$, where λ is the wavelength of the object radiation, we speak of one wave of spherical aberration.

The distribution of rays in an image plane is called the ray spot diagram. Their density (i.e., the number of rays per unit area) is called the *geometrical point-spread function* (PSF). If the system is aberration free, then the wavefront is spherical and all the object rays transmitted by the system converge to the Gaussian image point. When the wavefront is aberrated, a ray passing through a point (r, θ) in the plane of the exit pupil intersects the Gaussian image plane at a point (x_i, y_i) which, following Eq. (1-1), may be written

$$(x_i, y_i) = 2F\left(\frac{\partial W}{\partial \xi},\ \frac{\partial W}{\partial \eta}\right) \tag{7-5a}$$

$$= 2F\left(\cos\theta\frac{\partial W}{\partial \rho} - \frac{\sin\theta}{\rho}\frac{\partial W}{\partial \theta},\ \sin\theta\frac{\partial W}{\partial \rho} + \frac{\cos\theta}{\rho}\frac{\partial W}{\partial \theta}\right) \quad , \tag{7-5b}$$

where $F = R/2a$ is the focal ratio or the f-number of the image-forming light cone. Here, (x_i, y_i) represent the *ray aberrations*, i.e., the coordinates of the point of intersection of

the ray in the Gaussian image plane with respect to the Gaussian image point, and R is the radius of curvature of the *Gaussian reference sphere* with respect to which the aberration $W(\rho, \theta)$ is defined. The reference sphere is centered at the Gaussian image point (0,0) and, like the aberrated wavefront, passes through the center of the exit pupil. In Eqs. (7-5), we have assumed that the refractive index of the medium in which the image is formed is unity since it is often the case in practice. Substituting Eq. (7-2) into Eq. (7-5b), we find that, in the absence of distortion, the chief ray intersects the Gaussian image plane at the Gaussian image point.

For a radially symmetric aberration, i.e., one for which $W(\rho, \theta) = W(\rho)$, we note from Eq. (7-5b) that the PSF is also radially symmetric. The radial distance r_i of a ray from the Gaussian image point in that case is given by

$$\begin{aligned} r_i &= \left(x_i^2 + y_i^2\right)^{1/2} \\ &= 2F \left| \frac{\partial W(\rho)}{\partial \rho} \right| \quad . \end{aligned} \tag{7-6}$$

For a *uniformly illuminated pupil*, the location of the centroid of a PSF can be obtained from the aberration function according to

$$(x_c, y_c) = \left(\frac{2F}{\pi}\right) \iint \left(\frac{\partial W}{\partial \xi}, \frac{\partial W}{\partial \eta} \right) d\xi \, d\eta \quad . \tag{7-7}$$

The standard deviation of the ray aberrations or the *spot sigma* is given by

$$\sigma_s = \left\langle (x_i - x_c)^2 + (y_i - y_c)^2 \right\rangle^{1/2} \tag{7-8a}$$

$$= 2F \left\{ \frac{1}{\pi} \iint \left[\left(\frac{\partial W}{\partial \xi} - x_c \right)^2 + \left(\frac{\partial W}{\partial \eta} - y_c \right)^2 \right] d\xi \, d\eta \right\}^{1/2} \quad . \tag{7-8b}$$

For a symmetric aberration such as astigmatism, the PSF is symmetric and the centroid lies at the origin, i.e., $(x_c, y_c) = (0, 0)$. The spot sigma in such cases is equal to the *root mean square* (*rms*) *radius*. Substituting Eq. (7-6) for a radially symmetric aberration, Eq. (7-8b) reduces to

$$\sigma_s = 2\sqrt{2}F \left[\int \left(\frac{\partial W}{\partial \rho} \right)^2 \rho \, d\rho \right]^{1/2} \quad . \tag{7-9}$$

Now we discuss the characteristics of an image aberrated by a primary aberration. To be definite, we assume that each of the aberration coefficients A_i is positive, unless stated otherwise.

7.3 SPHERICAL ABERRATION

Figure 7-1 illustrates the relationship between a wavefront aberrated by spherical aberration

$$W(\rho) = A_s \rho^4 \tag{7-10}$$

and the reference sphere centered at a Gaussian image point P_0'. Substituting Eq. (7-10) into Eq. (7-6), we find that a ray of zone ρ intersects the Gaussian image plane at a distance

$$r_i = 8FA_s \rho^3 \tag{7-11}$$

from P_0'. Thus, the rays lying on a circle of radius ρ in the exit pupil lie on a circle of radius r_i given by Eq. (7-11) in the Gaussian image plane. The maximum value of r_i is $8FA_s$ and corresponds to rays with $\rho = 1$; i.e., it corresponds to the marginal rays. We shall refer to the maximum value of r_i as the radius of the image spot. Note that since A_s

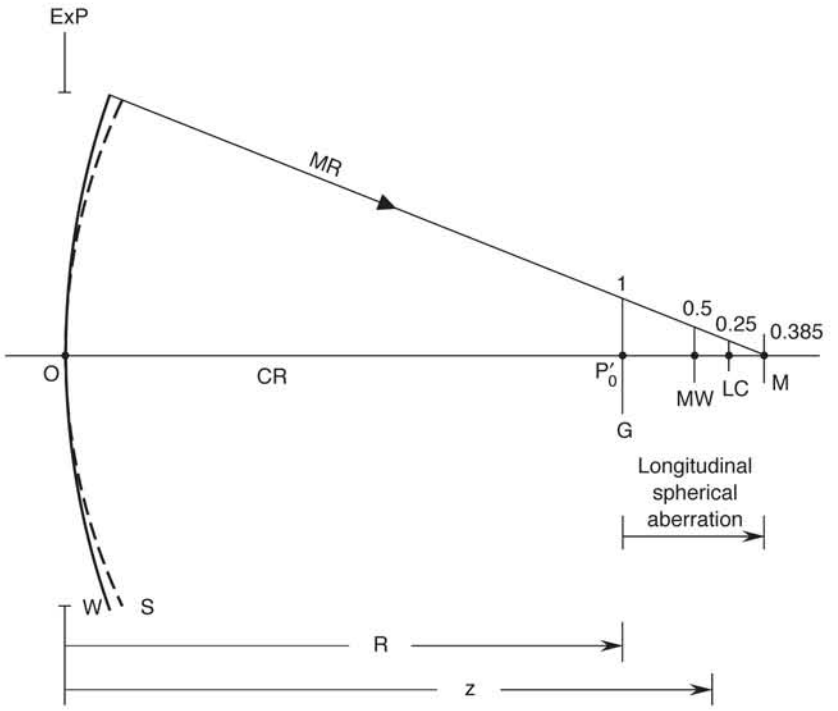


Figure 7-1. Ray spot radii in various image planes for a wavefront *W* aberrated by spherical aberration. *P*-paraxial, *M*-marginal, *MW*-midway, *LC*-least confusion. The reference sphere *S* is centered at the Gaussian image point P_0'. The values of spot radii indicated in the figure are in units of $8FA_s$, where *F* is the focal ratio of the image-forming light cone and A_s is the peak value of the spherical aberration.

is independent of the height *h* of the point object from the optical axis, the ray distribution owing to spherical aberration alone is also independent of *h*.

Let us consider the ray distribution in a slightly defocused image plane by introducing a defocus aberration B_d. The aberration with respect to a reference sphere centered at a defocused point may be written

$$W(\rho) = A_s\rho^4 + B_d\rho^2 \quad , \tag{7-12}$$

The rays of zone ρ now lie in the defocused image plane on a circle of radius

$$r_i = 8FA_s\left|\rho^3 + (B_d/2A_s)\rho\right| \quad . \tag{7-13}$$

For the *marginal rays*, corresponding to $\rho = 1$, $r_i \to 0$ if $B_d = -2A_s$, Following Eqs. (1-3c) and (1-3d), we find that the marginal rays intersect the axis at a distance

$$\Delta = -8F^2B_d \tag{7-14a}$$

$$= 16F^2A_s \tag{7-15b}$$

from P_0'. This distance shown as $P_0'M$ in Figure 7-1 is called the *longitudinal spherical aberration.* A negative value of Δ implies that, compared to the old reference sphere, the new reference sphere is centered at a point that is farther from the center of the exit pupil. Hence, the point of intersection *M* of the marginal rays lies to the right of P_0' as shown in Figure 7-1. This is to be expected for positive values of A_s. The points P_0' and *M* are called the *Gaussian* and the *marginal image points.* Substituting $B_d = -2A_s$ in Eq. (7-10), we find that the maximum value of r_i in the *marginal image plane* occurs for rays of zone $\rho = 1/\sqrt{3}$. This maximum value $r_{i\max}$ is $2/3\sqrt{3}$ (or 0.385) times the corresponding value in the *Gaussian image plane*. Thus, the *marginal spot radius* is considerably smaller than the *paraxial spot radius* .

The image plane lying midway between the paraxial and marginal planes corresponds to $B_d = -A_s$. The spot radius in this plane is half of that in the paraxial plane and corresponds to marginal rays. Comparing Eq. (7-13) with Eq. (5-8), we find that the spot radius is minimum in a plane corresponding to $B_d = -3A_s/2$, i.e., a plane which is 3/4 of the way from the paraxial plane to the marginal plane. The spot radius in this case is 1/4 of the paraxial spot radius and corresponds to rays of zone $\rho = 1/2$ and 1. This spot is called the *circle of least* (*spherical*) *confusion.* The spot radii in various image planes are listed in Table 7-1.

The deliberate mixing of one aberration with one or more other aberrations is called *aberration balancing*. Here, we have balanced spherical aberration with defocus in order to minimize the spot radius or its sigma value. The amount of defocus that gives the

Table 7-1. Spot radius and sigma for spherical aberration A_s.

Image Plane	Balancing Defocus B_d/A_s	Spot Radius $r_{imax}/8FA_s$	Spot Sigma $\sigma_s/8FA_s$
Gaussian	0	1	0.5
Marginal	– 2	0.385	0.289
Midway	– 1	0.5	0.204
Minimum spot sigma	– 4/3	1/3	0.167
Least confusion	– 3/2	0.25	0.177

smallest ray spot or sigma may be called the *optimum defocus* based on geometrical optics. The balanced aberration giving the smallest ray spot is $A_s[\rho^4 - (3/2)\rho^2]$. Similarly, the balanced aberration that gives the smallest spot sigma is $A_s[\rho^4 - (4/3)\rho^2]$. Based on *diffraction*, the optimum amount of defocus corresponds to the midway plane, since in that case it is used to reduce the *variance* of the aberration across the exit pupil, i.e., the balanced aberration giving minimum variance is $A_s\left(\rho^4 - \rho^2\right)$, similar to the Zernike polynomial $Z_4^0(\rho)$ (see Table 8-2).

7.4 COMA

The coma wave aberration is given by

$$W(\rho,\theta) = A_c\rho^3\cos\theta = A_c\xi\left(\xi^2+\eta^2\right) \quad . \tag{7-16}$$

Substituting Eq. (7-13) into Eq. (7-55), we obtain the corresponding ray aberrations in the Gaussian image plane with respect to the Gaussian image point. They are given by

$$(x_i, y_i) = 2FA_c\rho^2(2+\cos2\theta, \sin2\theta) \tag{7-17a}$$

$$= 2FA_c\left(\rho^2 + 2\xi^2, 2\xi\eta\right) \quad . \tag{7-17b}$$

We note that the rays coming from a circle of radius ρ in the exit pupil lie on a circle of radius $2FA_c\rho^2$ in the image plane, which is centered at $\left(4FA_c\rho^2, 0\right)$. The circle in the image plane is traced out twice as 0 varies from 0 to 2π to complete a circle of rays in the exit pupil. Figure 7-2 illustrates these circles in the image plane for $\rho = 1/2$ and 1. For $\rho = 1$, the rays in the image plane lie on a circle of radius $2FA_c$ centered at $(4FA_c, 0)$. Accordingly, $CB/CP' = 1/2$ where P' is the Gaussian image point, so that the angle $CP'B$ is equal to 30°. Hence, all of the rays in the image plane are contained in a cone of semiangle 30° bounded by a circle of radius $2FA_c$ centered at $(4FA_c, 0)$ corresponding to the marginal rays. The vertex of the cone, of course, coincides with the Gaussian image point P'. Since the spot diagram has the shape of a comet, the aberration is appropriately

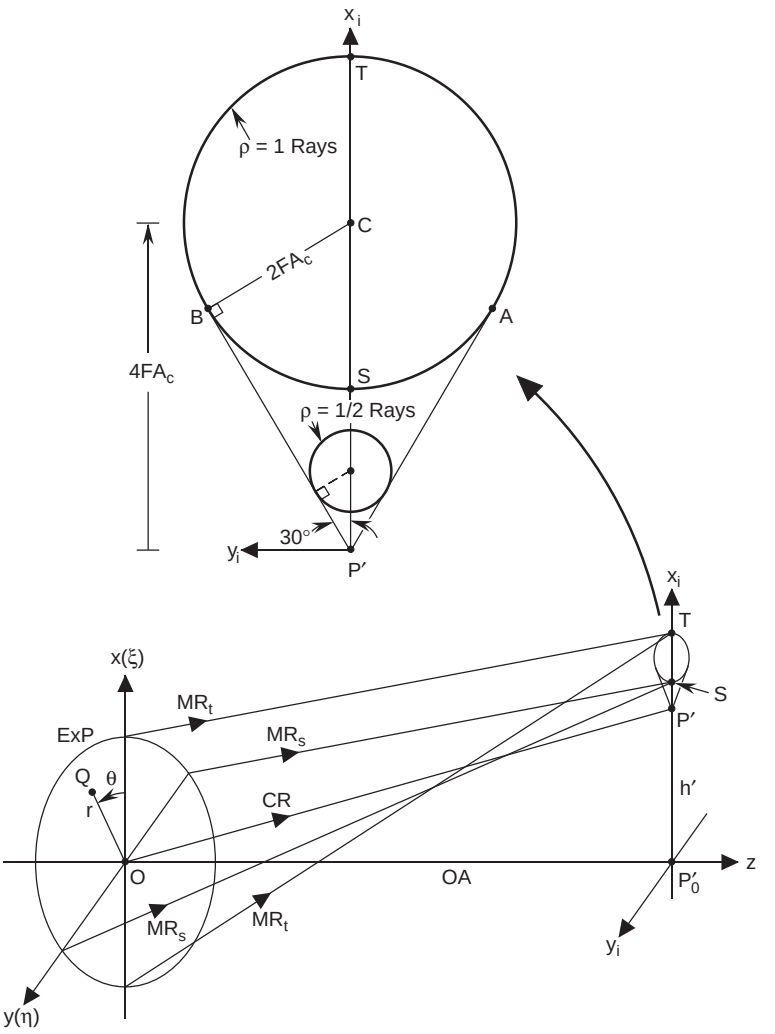


Figure 7-2. Ray spot diagram for coma. The tangential marginal rays MR_t are focused at the point T and the sagittal marginal rays MR_s are focused at the point S. All rays in the image plane lie in a cone of semiangle 30° with its vertex at the Gaussian image point P' bounded by the upper arc of a circle of radius $2FA_c$ centered at $(4FA_c, 0)$. The cone angle is 30° because $CB/CP' = 1/2$.

called coma. Note that the two tangential marginal rays $MR_t\,(\rho = 1,\ \theta = 0,\ \pi)$ intersect this plane at T at a distance $6FA_c$ from P', and the two sagittal marginal rays $MR_s\,(\rho = 1,\ \theta = \pi/2,\ 3\pi/2)$ intersect the image plane at S at a distance $2FA_c$ from P'. Accordingly, the length $6FA_c$ and half-width $2FA_c$ of the coma pattern are called *tangential* and *sagittal coma*, respectively.

Since the PSF is highly asymmetric about the Gaussian image point P', its centroid does not lie at it. Substituting Eq. (7-16) into Eq. (7-7), we obtain the location of the centroid:

$$\left(x_c, y_c\right) = \left(2FA_c, 0\right) \quad . \tag{7-18}$$

Thus, the centroid lies at the point S in Figure 7-2 where the sagittal marginal rays intersect the image plane. Substituting Eqs. (7-17) and (7-18) into Eq. (7-8a), we obtain the ray spot sigma:

$$\begin{aligned}\sigma_s &= 2FA_c \left\langle \left[\rho^2\left(2 + \cos 2\theta\right) - 1\right]^2 + \rho^4 \sin^2 2\theta \right\rangle^{1/2} \\ &= 2\sqrt{2/3}FA_c \quad . \end{aligned} \tag{7-19}$$

Measuring the ray coordinates in the image plane with respect to a point other than the Gaussian image point is equivalent to introducing a wavefront tilt aberration in the aberration function. A tilt aberration with a peak value of A_t is equivalent to measuring the wave aberration with respect to a reference sphere centered at a point in the image plane with coordinates $\left(-2FA_t, 0\right)$. Hence, measuring the ray aberrations with respect to the centroid is equivalent to a tilt aberration of $-A_c\rho\cos\theta$ or $A_t = -A_c$. Accordingly, the aberration function with respect to the centroid can be written

$$W\left(\rho, \theta\right) = A_c\left(\rho^3 - \rho\right)\cos\theta \quad . \tag{7-20}$$

It should be evident that if the ray aberrations are measured with respect to any other point in the image plane, including the Gaussian image point, the spot sigma will increase. The aberration function given by Eq. (7-20) represents coma aberration balanced optimally with tilt aberration to yield minimum sigma value or bring the centroid at the Gaussian image point. However, the variance of the wave aberration is minimum when $A_t = -\left(2/3\right)A_c$, i.e., if the balanced aberration is $A_c\left[\rho^3 - (2/3)\rho\right]\cos\theta$, similar to the Zernike polynomial $Z_3^1\left(\rho, \theta\right)$.

It is worth mentioning that the centroid of a PSF is associated with the line of sight of an imaging system, as discussed in Chapter 10. Moreover, the centroid of a geometrical PSF is identically the same as that of the diffraction PSF.

7.5 ASTIGMATISM

The astigmatism wave aberration is given by

$$W\left(\rho, \theta\right) = A_a\rho^2\cos^2\theta = A_a\xi^2 \quad . \tag{7-21}$$

The corresponding ray aberrations are given by

$$\left(x_i, y_i\right) = 4F\left(A_a\rho\cos\theta, 0\right) = 4F\left(A_a\xi, 0\right) \quad . \tag{7-22}$$

The point of intersection of a ray with the Gaussian image plane depends only on its ξ coordinate in the exit pupil. Thus, as indicated in Figure 7-3, all the rays transmitted by the exit pupil intersect the Gaussian image plane on a line along the *x* axis centered on the

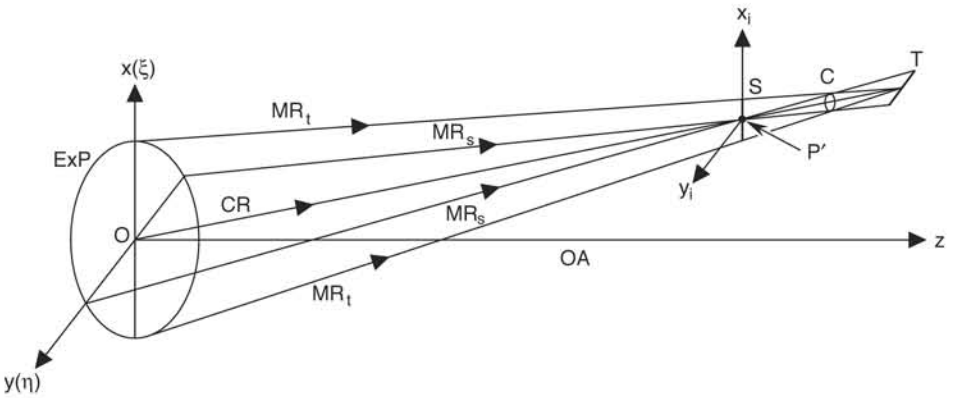


Figure 7-3. Astigmatic focal lines when only astigmatism is present. The tangential marginal rays MR_t are focused at a point on the tangential focal line *T*. Similarly, the sagittal marginal rays MR_s are focused at the Gaussian image point P' on the sagittal focal line *S*. The focal lines *S* and *T* lie in the tangential and sagittal planes, respectively. The circle of least confusion *C* lies in a plane midway between the planes of line images *S* and *T*.

Gaussian image point. Since $-1 \le \xi \le 1$, the full width of this line is $8FA_a$. If we add a small amount of defocus B_d to the astigmatism given by Eq. (7-21) by observing the image in a slightly defocused plane, the wave aberration becomes

$$W(\rho,\theta) = A_a\rho^2\cos^2\theta + B_d\rho^2 = (A_a + B_d)\xi^2 + B_d\eta^2 \quad . \tag{7-23}$$

The corresponding ray aberrations are given by

$$(x_i, y_i) = 4F\rho\left[(A_a + B_d)\cos\theta, B_d \sin\theta\right] \tag{7-24a}$$

$$= 4F\left[(A_a + B_d)\xi, B_d\,\eta\right] \quad . \tag{7-24b}$$

For a given value of ρ, the locus of the points of intersection of the rays in the defocused image plane is given by

$$\left(\frac{x_i}{A}\right)^2 + \left(\frac{y_i}{B}\right)^2 = 1 \quad , \tag{7-25}$$

where

$$A = 4F(A_a + B_d)\rho \text{ and } B = 4FB_d\,\rho \quad . \tag{7-26}$$

Thus, the rays lying on a circle of radius ρ in the exit pupil, in general, lie in a defocused image plane on an ellipse whose semiaxes are given by *A* and *B*, respectively. The largest ellipse is obtained for the marginal rays. The relationship defocus wave aberration B_d and the corresponding longitudinal defocus is given by Eq. (1-3d).

We note that if $B_d = 0$, the ellipse reduces to a line of full width of $8FA_a$ along the *x* axis. Thus, as discussed above, the image in the Gaussian image plane is a line *S* along the *x* axis centered on the Gaussian image point. If, however, $B_d = -A_a$, corresponding to $\Delta = 8F^2A_a$, then the ellipse reduces to a line *T* along the *y* axis. The full width of this line image is the same as that of the line image *S*. The line image along the *x* axis is called the *sagittal* (or *radial*) *image* and lies in the tangential (or meridional) plane *zx*, containing the point object (which lies along the *x* axis in the object plane) and the optical axis. Similarly, the line image *T* along the *y* axis is called the *tangential image* and lies in the sagittal plane *yz*. The distance $8F^2A_a$ between the two line images is called *longitudinal astigmatism.* The two line images are called the *astigmatic focal lines.*

If $B_d = -A_a/2$, corresponding to $\Delta = 4F^2A_a$, the ellipse reduces to a circle of maximum diameter of $4FA_a$,which is half the full width of the two line images. Since this circle is the smallest of all the possible images, Gaussian or defocused, it is called the *circle of least* (*astigmatic*) *confusion.* The spot sigma is minimum and equal to $\sqrt{2}FA_a$ in this plane.

Since $A_a \sim h'^2$, the width of the line images of a point object increases quadratically with the height h' of the Gaussian image point. Similarly, longitudinal astigmatism $8F^2A_a$ increases as h'^2 Thus, if we consider a line object, its sagittal image will also be a line, which is slightly longer (by an amount $8FA_a$) than but coincident with its Gaussian image. However, its tangential image will be parabolic with a vertex radius of curvature of $h'^2/16F^2A_a$ or $1/4R^2a_a$. Similarly, the sagittal image of a planar object will be planar, but its tangential image will be paraboloidal. Note that longitudinal astigmatism corresponding to a Gaussian image at a height h' represents the sag of the tangential image surface at that height.

Figure 7-4 illustrates the effect of astigmatism and field curvature on the image of a spoked wheel where the images formed on the sagittal and tangential surfaces are shown. A magnification of -1 is assumed in the figure. As discussed earlier, a point object *P* is imaged as a sagittal or radial line P_s' on the sagittal surface and as a tangential line P_t' on the tangential surface. Each point on the object is imaged in this manner, so that the sagittal image consists of sharp radial lines and diffuse circles while the tangential image consists of sharp circles and diffuse radial lines. If the object contains lines that are neither radial nor tangential, they will not be sharply imaged on any surface.

It should be understood that the astigmatism discussed here is for a system that is rotationally symmetric about its optical axis, and its value reduces to zero for an axial point object. It is different from the *astigmatism of the eye*, which is caused by one or more of its refracting surfaces, usually the cornea, that is curved more in one plane than another. The refracting surface that is normally spherical acquires a small cylindrical component, i.e., it becomes toric. Such a surface forms a line image of a point object even when it lies on its axis. Hence, a person afflicted with astigmatism sees points as lines. If

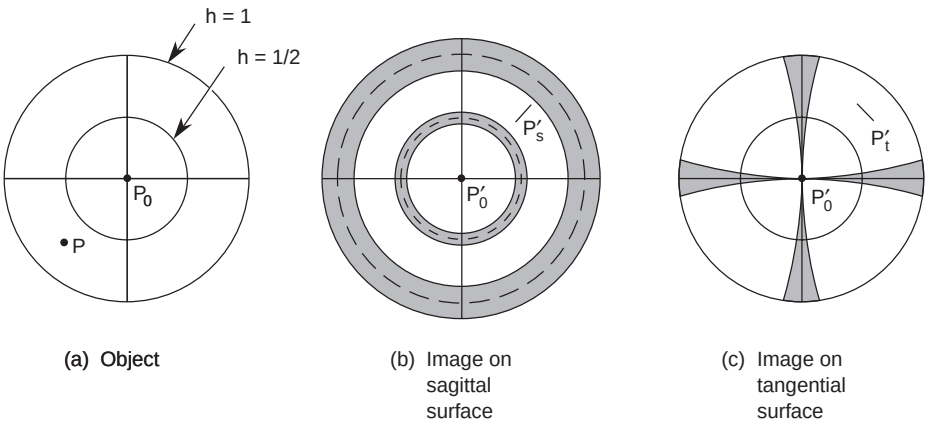


Figure 7-4. Astigmatic images of a spoked wheel. Gaussian magnification of the image is assumed to be – 1. The sagittal and tangential images P_s' and P_t' of a point object P are shown very much exaggerated. The dashed circles in (b) are the Gaussian images of the object circles.

the object consists of vertical and horizontal lines as in the wires of a window screen, such a person can focus (by accommodation) only on the vertical or the horizontal lines at a time. This is analogous to the spoked wheel example where the rim is in focus in one observation plane and the spokes are in focus in another.

7.6 FIELD CURVATURE

The wave aberration corresponding to field curvature is given by

$$W(\rho) = A_d \rho^2 = A_d\left(\xi^2 + \eta^2\right) \quad . \tag{7-27}$$

Since the wave aberration is radially symmetric, the distribution of rays in the Gaussian image plane is also radially symmetric. For rays lying on a circle of radius ρ in the exit pupil, the radius of the circle of corresponding rays in the image plane, following Eq. (7-6), is given by

$$r_i = 4FA_d \rho \quad . \tag{7-28}$$

Its maximum value is $4FA_d$ and corresponds to the marginal rays. The spot sigma value is $2\sqrt{2}FA_d$.

From the discussion in Section 1.4, we note that a defocus aberration represented by Eq. (7-21) implies that the wavefront is spherical, but it is not centered at the Gaussian image point. Instead, it is centered at a distance

$$\Delta = -8F^2 A_d \tag{7-29}$$

from the Gaussian image point along the optical axis. (Strictly speaking, it is centered on the line joining the center of the exit pupil and the Gaussian image point.) Since the aberration coefficient $A_a \sim h'^2$, the sagittal image of a line object will be parabolic with a vertex radius of curvature of $h'^2/16F^2A_d$, or $1/4R^2a_d$. Similarly, the image of a planar object will be paraboloidal. The paraboloidal surface for a system with zero astigmatism is called the *Petzval image surface.*

7.7 ASTIGMATISM AND FIELD CURVATURE

Now we consider the combined effect of astigmatism and field curvature. Thus, the aberration with respect to the Gaussian image point is now given by

$$W(\rho,\theta) = A_a\rho^2\cos^2\theta + A_d\rho^2 = (A_a + A_d)\xi^2 + A_d\eta^2 \quad . \tag{7-30}$$

Note that whereas in Eq. (7-23) the defocus coefficient was a variable, here it is fixed for a given point object. Since both A_a and A_d are proportional to h'^2, we find, following the discussion of Sections 7.5 and 7.6, that the sagittal and tangential images of a line object are formed on parabolic curves with vertex radii of curvature given by

$$R_s = h'^2/16F^2A_d = 1/4R^2a_d \tag{7-31}$$

and

$$R_t = h'^2/16F^2(A_a + A_d) = 1/4R^2(a_a + a_d) \quad , \tag{7-32}$$

respectively. Similarly, the images of a planar object centered on the optical axis will be the corresponding paraboloids symmetric about the optical axis.

Combining Eqs. (7-31) and (7-32) with Eq. (1-28), for imaging by a spherical refracting surface, where L is the same as R here, we find that

$$\frac{3}{R_s} - \frac{1}{R_t} = \frac{2}{R_p} \quad . \tag{7-33}$$

It has the consequence that the Petzval surface is three times as far from the tangential surface as it is from the sagittal surface, as may be seen by comparing the sags of the three surfaces. Moreover, the sagittal surface always lies between the tangential and the Petzval surfaces. When astigmatism is zero, the sagittal and tangential surfaces coincide with the Petzval surface. Although Eq. (7-33) and its consequences have been obtained for a single spherical refracting surface, they hold for any rotationally symmetric imaging system.

7.8 DISTORTION

The distortion wave aberration is given by

$$W(\rho,\theta) = A_t\rho\cos\theta = A_t\xi \quad , \tag{7-34}$$

where the aberration coefficient A_t is proportional to h'^3. The corresponding ray aberrations are given by

$$(x_i, y_i) = (2FA_t, 0) = (Ra_t h'^3, 0) \ . \tag{7-35}$$

Since the ray aberrations are independent of the coordinates (ρ, θ) of a ray in the exit pupil, all the rays converge at the image point $(2FA_t, 0)$, which lies along the x axis at a distance $2FA_t$ from the Gaussian image point. Thus, a wavefront aberrated by distortion is tilted with respect to the Gaussian reference sphere by an angle

$$\beta = A_t/a \ . \tag{7-36}$$

This angle is proportional to h'^3. Similarly, the distance $2FA_t$ of the perfect image point from the Gaussian image point is proportional to h'^3. This gives rise to the familiar *pincushion-* or *barrel*-distorted image of a square grid, shown in Figure 7-5, depending on whether A_t is positive or negative, respectively. Distortion is often measured as a fraction of the image height. Thus, for example, the percent distortion is $100Ra_t h'^2$.

7.9 SPOT DIAGRAMS

If an optical system is aberration free, the wavefront at its exit pupil corresponding to a certain point object is spherical, and all the object rays lying in the pupil plane converge at the Gaussian image point. For an aberrated system, the wavefront is nonspherical and the rays distributed in a *spot diagram* give a qualitative description of the effects of an aberration. A lens designer generally starts with rays that are distributed in a certain grid pattern in the plane of the entrance pupil of the system. Figure 7-6 shows the ray grid patterns in the pupil plane that are commonly used in practice. In Figure 7-6a, the rays are

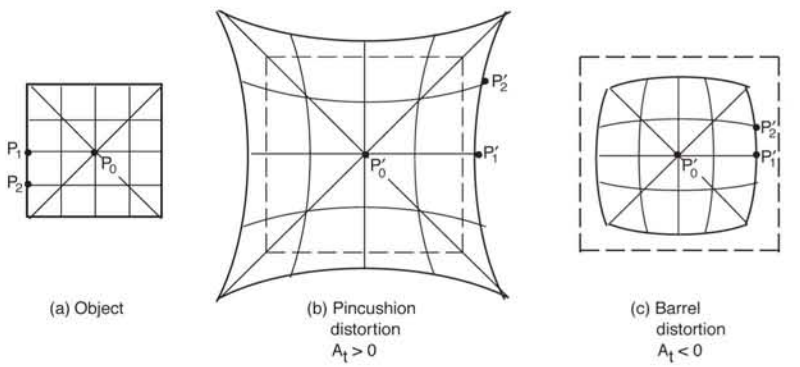


Figure 7-5. Images of a square grid in the presence of distortion. When the distortion aberration coefficient A_t is positive, we obtain pincushion distortion as in (b). When A_t is negative, we obtain barrel distribution as in (c). The dashed squares represent the Gaussian image of the square object with a magnification of – 1.5. Thus, whereas P_0' is the Gaussian image of an axial point object P_0, P_1', and P_2 are the images of the off-axis point objects P_1 and P_2 displaced from their Gaussian positions because of distortion.

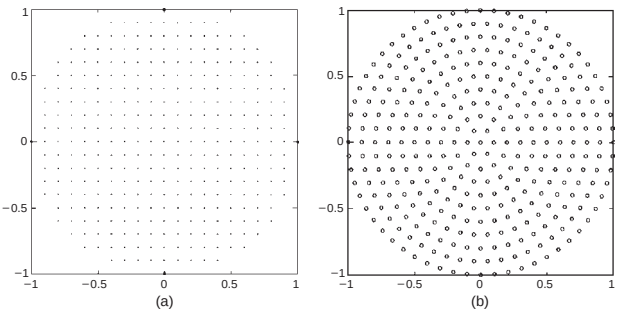


Figure 7-6. Ray grid pattern in the pupil plane normalized by the pupil radius. (a) Square grid of uniformly spaced points. (b) Hexa-polar grid of concentric rings.

distributed in a uniformly spaced square array, while in Figure 7-6b they are distributed in a hexa-polar array.

In the absence of any aberration, the spot diagram in a defocused image plane looks exactly like the one in the pupil plane, except for its scale. The spot diagrams for spherical aberration in various image planes considered above are shown in Figure 7-7. It is evident that, instead of the expected radial symmetry of the PSFs, a four-fold symmetry is obtained in the case of the square grid of rays in the pupil plane, and hexagonal symmetry in the case of the hexa-polar grid. This is simply an artifact of the ray grid used in the pupil plane. As in the case of defocus, the PSF for astigmatism is also uniform. Hence, the spot diagram for it also looks like the input array across an elliptical spot, which reduces to a circle or a line depending on the amount of balancing defocus. The spot diagrams for coma are shown in Figure 7-8. Only the chief ray passes through the Gaussian image point, which is shown with coordinates (0, 0) in the figure. Note that the two grids yield different results, as may be seen from near the top of the spot.

7.10 ABERRATION TOLERANCE AND A GOLDEN RULE OF OPTICAL DESIGN

It is common practice in lens design to look at the spot diagrams in the early stages of a design, in spite of the fact that they do not represent what is observed in reality. Optical designers consider a system to be close to its diffraction limit if the ray spot radius is less than or equal to the radius $1.22\lambda F$ of the Airy disc, discussed in Chapter 8. We note, for example, that this holds for spherical aberration in the Gaussian image plane if $A_s \leq 0.15\lambda$, although a larger value of A_s is obtained in the other image planes. Considering that the long dimension of the coma spot is $6FA_c$ and the line image for astigmatism is $8FA_a$ long, the aberration tolerance for the spot size to be smaller than the Airy disc is $A_c < 0.4\,\lambda$ and $A_a < 0.3\,\lambda$, respectively. The aberration tolerances based on the spot size are summarized in Table 7-2. These tolerances are roughly consistent with the Rayleigh's $\lambda/4$ (quarter-wave) rule (see Section 8.3.6) that the peak peak-to-valley

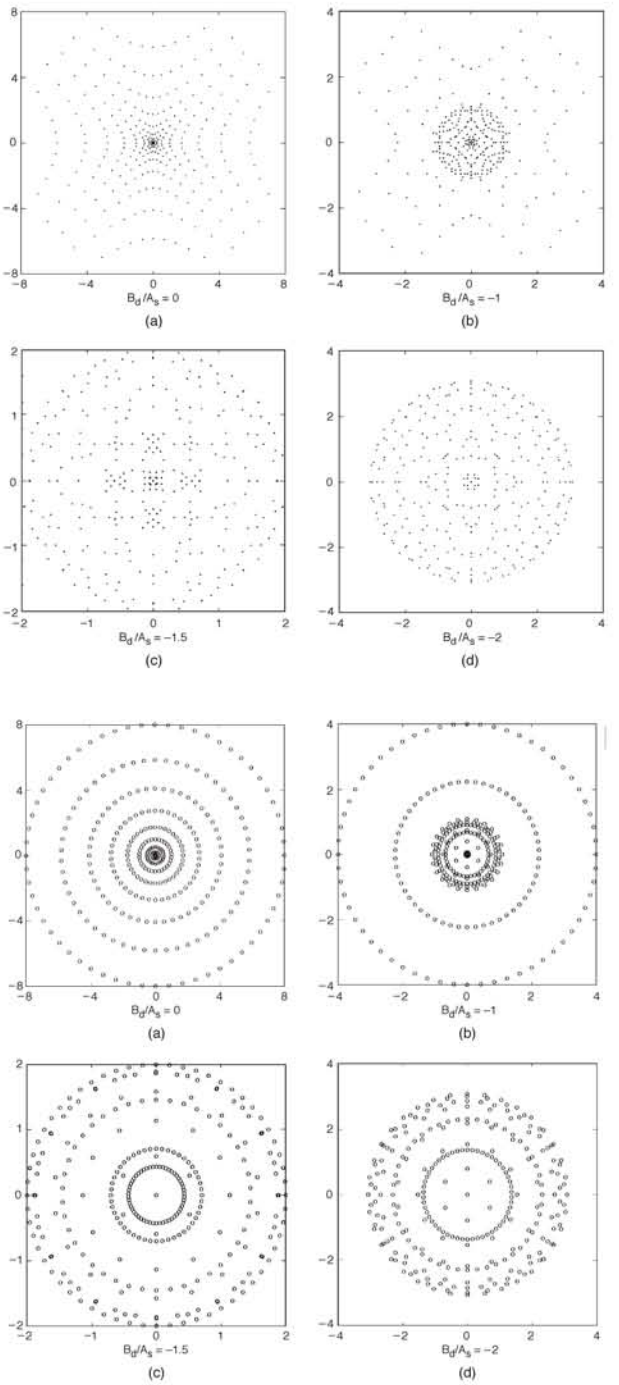


Figure 7-7. Spot diagrams for spherical aberration A_s in various image planes, as indicated by the value of B_d, for square and hexa-polar grids in the pupil plane: (a) Gaussian, (b) midway, (c) least confusion, and (d) marginal. The spot sizes are in units of FA_s. The PSFs are four- or six-fold symmetric, instead of being radially symmetric, depending on the grid used.

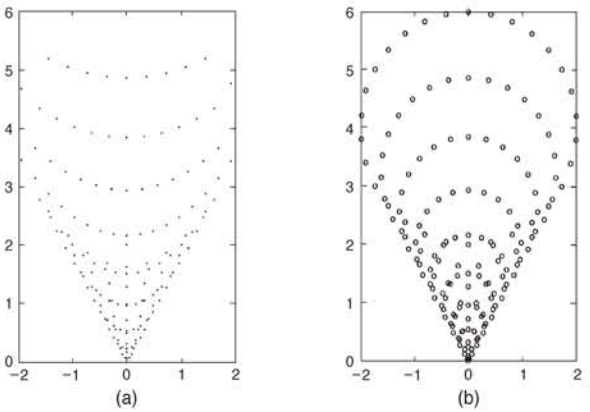


Figure 7-8. Spot diagrams for coma A_c in units of FA_c for (a) square and (b) hexapolar grid of rays in the pupil plane. Only the chief ray passes through the Gaussian image point P' indicated with coordinates (0, 0).

aberration be less than $\lambda/4$. This yields a *golden rule of optical design* in that a designer strives for a small spot until its size is nearly equal to that of the Airy disc, and then analyzes the system by its aberration variance and diffraction characteristics such as the PSF or the modulation transfer function.

The *depth of focus* (giving the tolerance on the location of the plane for observing the image) can be determined from Eqs. (7-28) and (7-29). Thus, the defocus aberration tolerance is $B_d \lesssim 0.3\,\lambda$ for a spot radius smaller than or equal to that of the Airy disc, which, in turn, implies a depth of focus of $2.4\lambda F^2$. Alternatively, the *depth of field* (giving the tolerance on the object location for a fixed observation plane) can be determined from the depth of focus by dividing by the longitudinal magnification. Similarly, distortion tolerance for a certain amount of *line-of-sight error* can be obtained from Eq. (7-36) .

Table 7-2. Aberration tolerance based on the ray spot size.

Aberration	Spot 'radius' in Gaussian image plane	Tolerance for near diffraction limit
Spherical A_s	$8FA_s$	$A_s \le 0.15\lambda$
Coma A_c	$3FA_c$	$A_c \le 0.4\lambda$
Astigmatism A_a	$4FA_a$	$A_a \le 0.3\lambda$
Defocus B_d	$4FB_d$	$B_d \le 0.3\lambda$

7.11 SUMMARY

The aberration-free image of a point object based on geometrical optics is a point. All of the object rays transmitted by the system pass through the Gaussian image point. For an aberrated system, the rays are distributed in the image plane in the vicinity of the Gaussian image point as a spot diagram. The quality of the aberrated image is determined by the size and the nature of the diagram, which, in turn, is described by its standard deviation and centroid. In the case of spherical aberration and astigmatism, the spot size can be reduced by introducing defocus, i.e., by observing the image in an appropriate plane other than the Gaussian image plane. The minimum spot thus obtained is referred to as the circle of least confusion, and reducing the spot size in this manner is called aberration balancing. A lens designer looks at the spot diagrams in the early stages of a design and considers diffraction images as the size of the spot approaches that of the Airy disc (discussed in Chapter 8).

CHAPTER 8

Systems With Circular Pupils

8.1 INTRODUCTION

In this chapter, we consider optical systems with *circular exit pupils* and discuss imaging by them based on the diffraction of object radiation at the exit pupil. Our starting point is an equation for the distribution of light in the image of a point object called the *diffraction point-spread function* (PSF) of the system. This equation is equally suitable for calculating the *diffraction pattern* of a circular aperture. Since, under certain conditions, the *diffraction image* of an *incoherent object* is given by the convolution of its Gaussian image (which is a scaled replica of the object) and the system PSF,[1-3] the PSF calculations are fundamental to the theory of optical imaging. To understand the effect of aberrations on images, it is essential that we first understand the aberration-free PSF. Accordingly, we give briefly the characteristics of the aberration-free image of a point object.

Our discussion on aberrated images is built slowly. First, we discuss *defocused images* and irradiance along the axis of the pupil. Next, approximate relationships between the ratio of the PSF values at its center with and without aberration, called the *Strehl ratio,* and the variance of the aberration across the pupil are developed. The approximate results for *primary aberrations* are compared with the corresponding exact results to determine the range of validity of the simple Strehl ratio formulas. The concept of *aberration balancing* is introduced in which an aberration of a certain order in pupil coordinates is mixed or balanced with one or more aberrations of lower order to minimize its variance, and thereby maximize the Strehl ratio of the system. *Aberration tolerances* based on a Strehl ratio of 0.8 are given for primary and *balanced primary aberrations.* Rayleigh's *quarter-wave rule* is briefly discussed, and balanced aberrations are identified with *Zernike circle polynomials.* The aberrated PSFs for various amounts of primary aberrations are given, and their *symmetry properties* in and about the Gaussian image plane are illustrated.

Since the diffraction image of an incoherent isoplanatic object is given by the convolution of its Gaussian image and the PSF of the system forming the image, the *spatial frequency spectrum* of the diffraction image is given by the product of the spectrum of the Gaussian image and the *optical transfer function* (OTF) of the system.[1-3] Thus, the OTF of the system, which is equal to the Fourier transform of its PSF, is equally fundamental to the theory of optical imaging. The OTF of an aberration-free system with a circular pupil is given, and how it is affected by an aberration is discussed. The concept of a *Hopkins ratio,* representing the ratio of the magnitudes of the OTFs at a certain spatial frequency, with and without aberration is introduced. Aberration tolerances based on a Hopkins ratio of 0.8 are given for primary aberrations. Finally, *contrast reversal* of certain spatial frequencies in an object is illustrated by considering their imaging by a defocused system.

8.2 POINT-SPREAD FUNCTION (PSF)

In this section we give a general equation for evaluating the aberrated PSF of a system with a circular exit pupil. We give closed-form analytical solutions for the aberration-free PSF and encircled power giving the fraction of total power in the image of a point object in a circle of certain radius centered at the Gaussian image point. Defocused PSFs and axial irradiance of the convergent image-forming beam are considered next. It is shown, for example, that the irradiance distribution even for an aberration-free system is not symmetric about the Gaussian image plane unless the Fresnel number of the pupil (defined below) as observed from the Gaussian image point is very large. The content of this section forms the basis from which to study the effects of aberrations on the images.

8.2.1 Aberrated PSF

Consider an aberrated optical system with a circular exit pupil of radius *a* imaging a point object radiating at a wavelength λ. Let *R* be the distance between the planes of the exit pupil and the Gaussian image and let $\Phi(\rho,\theta)$ be the *phase aberration* at a point (ρ,θ) in the pupil plane, where ρ is in units of *a*. The phase aberration Φ is related to the *wave aberration* $W(\rho,\theta)$ considered in earlier chapters according to $\Phi = (2\pi/\lambda)W$. The diffraction PSF or the irradiance distribution of the image in a plane normal to its optical or the *z* axis at a distance *z* from the plane of its exit pupil may be written[1]

$$I(r,\theta_i; z) = \frac{PS_p}{\pi^2\lambda^2 z^2}\left|\int_0^1\int_0^{2\pi} \exp\left[i\Phi(\rho,\theta)\right]\exp\left[-\pi i\frac{R}{z}\rho r\cos(\theta-\theta_i)\right]\rho\, d\rho\, d\theta\right|^2, \qquad (8\text{-}1)$$

where (r,θ_i) are the polar coordinates of the observation point with respect to the point where the line joining the center of the exit pupil and the Gaussian image point intersects the observation plane, *r* is in units of λF ($F = R/2a$ being the *focal ratio* or the *f-number* of the image-forming light cone), *P* is the total power in the exit pupil and, therefore, in the image, and $S_p = \pi a^2$ is the area of the exit pupil of the system. Strictly speaking, the PSF of a system represents the irradiance distribution of the image of a point object per total power in the image. Accordingly, whereas the irradiance is in units of W/m^2, the PSF is in units of m^{-2}. The angles θ and θ_i are zero for pupil and observation points lying in the tangential plane on the positive side of the *x* axis. As in earlier chapters, we assume that the point object lies along the *x* axis so that the *zx* plane represents the tangential plane.

The function $\exp\left[i\Phi(\rho,\theta)\right]$ is called the *pupil function* of the system. A system whose aberration function $\Phi(\rho,\theta)$ is (approximately) the same for all points on an extended object is called *isoplanatic*. The image of an incoherent object formed by such a system is obtained by convolving its Gaussian image with the PSF of the system, i.e., by adding the irradiance distributions of its image elements. Similarly, the complex amplitude distribution of the image of a *coherent object* formed by such a system is obtained by adding the complex amplitude distributions of its image elements.

The PSF of a system depends on the optical wavelength λ in several ways. Referring to Eq. (8-1), first, the power P in the exit pupil at a certain wavelength (strictly speaking, across a narrow spectral band with a certain mean wavelength) may be different from that at another wavelength. This variation will depend on the spectral radiance distribution of the object and the spectral transmission of the system. Second, there is an inverse-square-law dependence on the wavelength. It affects the "brightness" of the PSF: the shorter the wavelength, the brighter the PSF. Third, the wave aberration may depend on the wavelength if the system has one or more dispersive elements. Even if the system is nondispersive, the phase aberration is inversely proportional to it. Hence, the effect of an aberration on the PSF is different at two different wavelengths. Fourth, the variable r is normalized by the wavelength. It affects the "size" of the PSF: the shorter the wavelength, the narrower the PSF. However, a shorter wavelength also means larger phase aberration and, therefore, more spreading of the PSF due to the aberration. The white light or polychromatic PSF may be determined by integrating the monochromatic PSF across the spectral distribution of the image-forming radiation.

8.2.2 Aberration-Free PSF

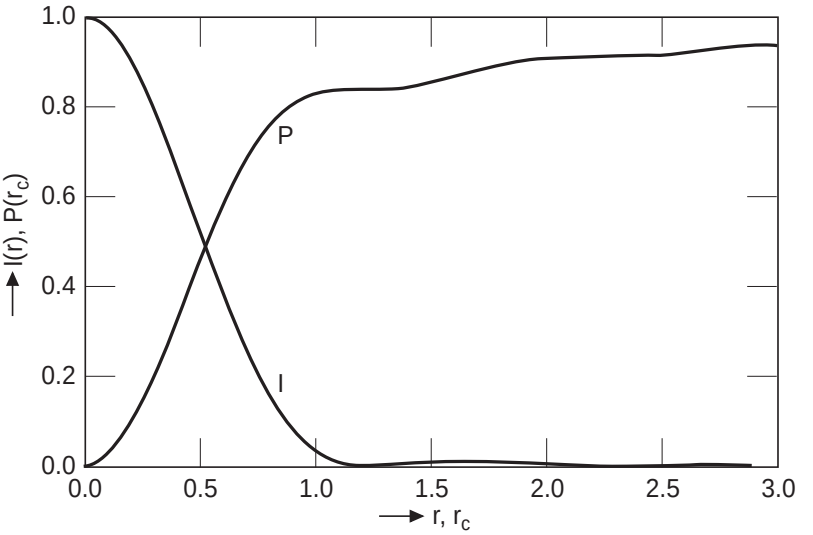
It can be shown that the aberration-free irradiance distribution, obtained from Eq. (8-1) by letting $\Phi(\rho, \theta) = 0$ and $z = R$, is given by

$$I(r; R) = \frac{PS_p}{\lambda^2 R^2}\left[\frac{2J_1(\pi r)}{\pi r}\right]^2 \quad , \tag{8-2}$$

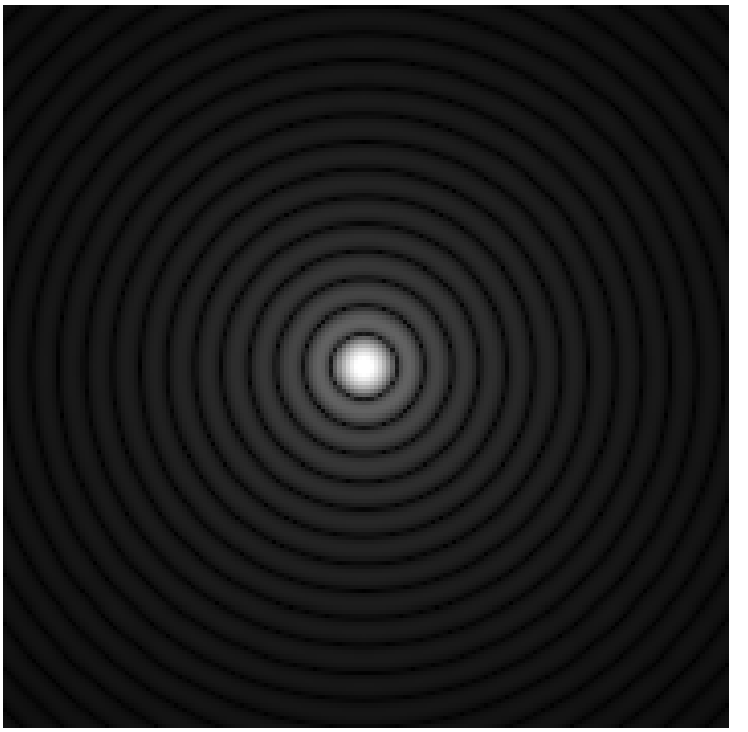
where $J_1(\cdot)$ is the first-order Bessel function of the first kind. This distribution, normalized by its central value $PS_p/\lambda^2 R^2$, is shown in Figure 8-1a. It is called the *Airy pattern*,[4] and it is illustrated in 2D in Figure 8-1b. It consists of a bright spot, called the *Airy disc*, surrounded by dark and bright rings. The power contained in a circle of radius r_c (in units of λF) centered at the Gaussian image point $r = 0$ is given by

$$P(r_c) = 1 - J_0^2(\pi r_c) - J_1^2(\pi r_c) \quad . \tag{8-3}$$

The encircled-power distribution normalized by the total power P is also shown in Figure 8-1a. The location of the *maxima* and *minima* of the irradiance distribution, the values of irradiance at these points, and the corresponding encircled powers are given in Table 8-1. The minima and maxima correspond to the roots of $J_1(\pi r) = 0$ and $J_2(\pi r) = 0$, respectively, where $J_2(\cdot)$ is the second-order Bessel function of the first kind. It should be evident that the encircled power corresponding to minima is given by $1 - J_0^2(\pi r_m)$, where r_m represents the value of r for a minimum. The central bright spot of radius 1.22 contains 83.8% of the total power. Note that the principal maximum of the irradiance distribution lies at $r = 0$ where Huygens' spherical wavelets originating at the exit pupil interfere constructively. The aberration-free image of an object is also called its *diffraction-limited image* (since the quality of the image is limited only by diffraction of



(a)



(b)

Figure 8-1. (a) Irradiance and encircled power distributions for an aberration-free system with a circular pupil. (b) 2D PSF, called the Airy pattern.

Table 8-1. Irradiance and encircled power corresponding to the maxima and minima of the PSF. The irradiance is normalized by the central value $I_i(0) = PS_p/\lambda^2 R^2$, and the encircled power is normalized by the total power P in the exit pupil and, therefore, in the image. r and r_c are in units of λF.

Max/Min	r, r_c	$I(r)$	$P(r_c)$
Max	0	1	0
Min	1.22	0	0.838
Max	1.64	0.0175	0.867
Min	2.23	0	0.910
Max	2.68	0.0042	0.922
Min	3.24	0	0.938
Max	3.70	0.0016	0.944
Min	4.24	0	0.952
Max	4.71	0.0008	0.957

the object radiation at the exit pupil of the system). It should be noted that the radius of the Airy disc increases linearly with the wavelength and the irradiance at its center decreases quadratically with it.

8.2.3 Rotationally Symmetric PSF

For a radially symmetric aberration $\Phi(\rho)$, carrying out the angular integration, Eq. (8-1) reduces to

$$I(r; z) = \frac{4PS_p}{\lambda^2 z^2}\left|\int_0^1 \exp\left[i\Phi(\rho)\right] J_0(\pi r\rho R/z)\, \rho\, d\rho\right|^2 . \tag{8-4}$$

It is clear from Eq. (8-4) that the irradiance distribution is rotationally symmetric about the z axis. Hence, it is radially symmetric in any observation plane normal to it. Moreover, it does not depend on the sign of $\Phi(\rho)$ (as may be seen by changing i to $-i$).

8.2.4 Defocused PSF

If the imaging system is aberration free but the image is observed in a plane $z \neq R$ then the image suffers from defocus aberration given by (see Section 1.4)

$$\Phi(\rho; z) = B_d(z)\rho^2 , \tag{8-5}$$

where

$$B_d(z) = \frac{\pi a^2}{\lambda}\left(\frac{1}{z} - \frac{1}{R}\right) \tag{8-6a}$$

$$= \pi N\left(\frac{R}{z}-1\right) \quad , \tag{8-6b}$$

is the peak defocus aberration. In Eq. (8-6b), $N = a^2/\lambda R$ is the *Fresnel number* of the exit pupil as observed from the center of the Gaussian image plane. Thus, the edge of the exit pupil is farther than its center by approximately $N\lambda/2$ from the center of the Gaussian image plane so that N is the number of Fresnel's *half-wave zones* in the exit pupil.

From Eq. (8-1) we note that the irradiance distribution is asymmetric about the plane $z = R$; i.e., the distributions in two planes located at $z = R \pm \Delta$, where Δ is a longitudinal defocus, are not identical, even if the system is otherwise aberration free. There are three reasons for this asymmetry. First, the inverse-square law dependence on z increases the irradiance for $z < R$ and decreases it for $z > R$. Second, B_d is asymmetric since the defocus coefficients for these two planes are different, as may be seen from Eq. (8-6). Third, the exponent in Eq. (8-1), which determines the scale of the image, depends on z.

For systems with a small Fresnel number $N \lesssim 5$, z can be much different from R for B_d to achieve a significant value. Accordingly, all of the three factors mentioned above contribute to the asymmetry of the irradiance distribution about the plane $z = R$. One consequence of this is that the irradiance at points on and near the z axis can be higher for $z < R$ than for $z = R$. For example, a beam of diameter 25 cm and a wavelength of 10.6 μm focused at a distance of 1.5 km corresponds to $N = 1$. A Strehl ratio [discussed in the next section and whose exact value is given by the square of the quantity in parenthesis in Eq. (8-7)] of 0.8 is obtained at two z values: 1 km and 3 km. The principal maximum of axial irradiance occurs at a distance of $0.6R = 0.9$ km.

If the Fresnel number of a system is very large ($N >> 10$), B_d becomes large even for very small differences in z and R. For example, a photographic system with $a = 1$ cm, $\lambda = 0.5$ μm, and $R = 10$ cm corresponds to $N = 2000$, and a Strehl ratio of 0.8 is obtained for $z = R \pm 25$ μm. Accordingly, the defocus tolerance for such systems dictates that z be practically equal to R. Hence, following Eq. (8-6a), we note that two observation planes $z = R \pm \Delta$ correspond to defocus coefficients of $B_d = \mp \pi\Delta/4\lambda F^2$. Since these coefficients are equal in magnitude but opposite in sign, letting $\Phi(\rho) = B_d \rho^2$ in Eq. (8-4), we find that the irradiance distribution for an unaberrated system with a large Fresnel number is symmetric about the Gaussian image plane $z = R$.

8.2.5 Axial Irradiance

If we let $r = 0$ in Eq. (8-1), we obtain the irradiance along the z axis (strictly speaking, along the line joining the center of the exit pupil and the Gaussian image point). For an aberration-free system, the axial irradiance is given by

$$I(0;z) = \frac{PS_p}{\lambda^2 z^2}\left(\frac{\sin B_d/2}{B_d/2}\right)^2 \quad . \tag{8-7}$$

Figure 8-2 shows how this irradiance varies with z for systems with $N = 1$, 10, and 100.[5] We note that it is highly asymmetric about the Gaussian image point ($z = R$) when $N = 1$, but it becomes more and more symmetric as N increases. The axial irradiance for a Gaussian pupil (with a truncation parameter of unity discussed in Chapter 9) is also shown in this figure with a similar behavior. It should be noted that even though the principal maximum of axial irradiance does not lie at the focus, unless N is very large, maximum central irradiance on a target in a beam-focusing system always occurs when the beam is focused on it.[5] It is evident from Figure 8-2 that the *depth of focus* decreases as N increases.

8.3 STREHL RATIO

Now we consider an aberrated system and discuss how the value at the center of the PSF is affected by the aberration in the system, thereby introducing the concept of the Strehl ratio. We obtain simple but approximate expressions for the Strehl ratio in terms of the variance of the aberration across the exit pupil of the system. We introduce the concept of aberration balancing in which an aberration of a certain order in pupil coordinates is balanced by one or more aberrations of lower order to minimize its variance and thus maximize the Strehl ratio. We give aberration tolerances for primary and balanced primary aberrations corresponding to a Strehl ratio of 0.8. A brief discussion of Rayleigh's quarter-wave rule is given, and balanced primary aberrations are identified with the corresponding Zernike circle polynomials.

8.3.1 General Expressions

The *Strehl irradiance ratio* of an image or a system, defined as the ratio of the irradiance values at the center of the image in a plane with and without the aberration, according to Eq. (8-1) is given by

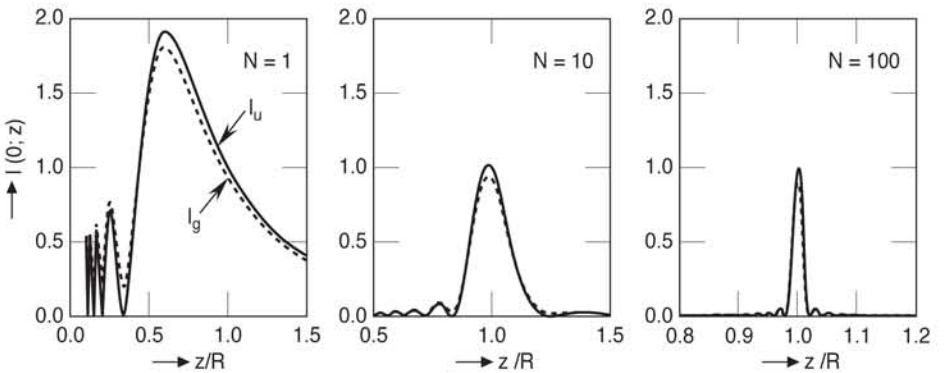


Figure 8-2. Axial irradiance of a circular beam focused at a distance R with a Fresnel number $N = 1$, 10, and 100. The irradiance is in units of the focal-point irradiance $PS_p/\lambda R^2$. The subscripts u and g refer to uniform and (truncated) Gaussian beams, respectively. Gaussian beams are discussed in Chapter 9.

$$S = \pi^2 \left| \int_0^1 \int_0^{2\pi} \exp\left[i\Phi(\rho,\theta)\right] \rho \, d\rho \, d\theta \right|^2 \tag{8-8}$$

$$= \left| \left\langle \exp\left[i \left(\Phi - \langle \Phi \rangle \right) \right] \right\rangle \right|^2$$

$$= \left\langle \cos\left(\Phi - \langle \Phi \rangle \right) \right\rangle^2 + \left\langle \sin\left(\Phi - \langle \Phi \rangle \right) \right\rangle^2$$

$$\geq \left\langle \cos\left(\Phi - \langle \Phi \rangle \right) \right\rangle^2 \quad , \tag{8-9}$$

where the angular brackets indicate an average across the pupil. Expanding the cosine function in a power series and retaining the first two terms for small aberrations yields the Maréchal result[6]

$$S \gtrapprox \left(1 - \sigma_\Phi^2 / 2 \right)^2 \quad , \tag{8-10}$$

where

$$\sigma_\Phi^2 = \left\langle \left(\Phi - \langle \Phi \rangle \right)^2 \right\rangle$$

$$= \left\langle \Phi^2 \right\rangle - \langle \Phi \rangle^2 \tag{8-11}$$

is the *variance* of the phase aberration across the pupil. Note that

$$\left\langle \Phi^n \right\rangle = \pi \int_0^1 \int_0^{2\pi} \Phi^n(\rho,\theta) \, \rho \, d\rho \, d\theta \quad . \tag{8-12}$$

For small values of *standard deviation* σ_Φ, three approximate expressions have been used in the literature:

$$S_1 \simeq \left(1 - \sigma_\Phi^2 / 2 \right)^2 \quad , \tag{8-13}$$

$$S_2 \simeq 1 - \sigma_\Phi^2 \quad , \tag{8-14}$$

and

$$S_3 \simeq \exp\left(- \sigma_\Phi^2 \right) \quad . \tag{8-15}$$

The first is the *Maréchal formula*, the second is the commonly used expression obtained when the term in σ_Φ^4 in the first is neglected,[7] and the third is an empirical expression giving a better fit to the actual numerical results for various aberrations.[8] We note that the Strehl ratio for a small aberration does not depend on its type but only on its variance across the pupil.

8.3.2 Primary Aberrations

Table 8-2 gives the form as well as the standard deviation σ_Φ of a primary aberration. Here, the aberration coefficient A_i represents the peak value of an aberration. (The balanced aberrations noted in the table are considered in the next section.) Comparing it with the aberration coefficients a_j considered in earlier chapters, we note, for example, that $A_c = a_c h' a^3$, where h' is the height of the Gaussian image point from the optical axis of the system. It also lists the tolerance for an aberration coefficient A_i for a Strehl ratio of 0.8. The *optical tolerances* listed in Table 8-2 are for the wave (as opposed to phase) aberration coefficient, as is customary in optics. A Strehl ratio of 0.8 corresponds to an aberration with a standard deviation of $\sigma_W = \lambda/14$.

8.3.3 Balanced Primary Aberrations

In Chapter 7, where we discussed ray aberrations, we mixed one aberration with another in order to minimize the size of the ray spot in an image plane. For example, in the case of spherical aberration, the circle of least confusion was determined to be in a plane 3/4 of the way from the Gaussian image plane to the marginal image plane. The radius of the circle of least confusion was found to be 1/4 of the spot radius in the Gaussian image plane. Similarly, in the case of astigmatism, it is in a plane lying midway between the planes containing the sagittal (Gaussian) and tangential line images with a diameter that is half the length of the line images.

Table 8-2. Primary aberrations, their standard deviations, and values of aberration coefficients, peak aberrations, and peak-to-valley aberrations for a Strehl ratio of 0.8.

					S = 0.8		
Aberration	$\Phi(\rho,\theta)$	σ_Φ	W_p	W_{p-v}	A_i	$\lvert W_p \rvert$	$W_{p\ v}$
Spherical	$A_s\rho^4$	$\frac{2A_s}{3\sqrt{5}}$	A_s	A_s	0.25	0.25	0.25
Balanced spherical	$A_s(\rho^4-\rho^2)$	$\frac{A_s}{6\sqrt{5}}$	$\frac{A_s}{4}$	$\frac{A_s}{4}$	1	0.25	0.25
Coma	$A_c\rho^3\cos\theta$	$\frac{A_c}{2\sqrt{2}}$	A_c	$2A_c$	0.21	0.21	0.42
Balanced coma	$A_c(\rho^3-2\rho/3)\cos\theta$	$\frac{A_c}{6\sqrt{2}}$	$\frac{A_c}{3}$	$\frac{2A_c}{3}$	0.63	0.21	0.42
Astigmatism	$A_a\rho^2\cos^2\theta$	$\frac{A_a}{4}$	A_a	A_a	0.30	0.30	0.30
Balanced astigmatism	$A_a\rho^2(\cos^2\theta-1/2)$ $= (A_a/2)\rho^2\cos2\theta$	$\frac{A_a}{2\sqrt{6}}$	$\frac{A_a}{2}$	A_a	0.37	0.18	0.37

For small aberrations, since the Strehl ratio is maximum when the aberration variance is minimum, the best image plane is one that corresponds to minimum variance. Thus, for example, we balance spherical aberration with defocus and write it as

$$\Phi(\rho) = A_s \rho^4 + B_d \rho^2 \quad . \tag{8-16}$$

We determine the amount of defocus B_d such that the variance σ_Φ^2, is minimized; i.e., we calculate σ_Φ and let

$$\frac{\partial \sigma_\Phi^2}{\partial B_d} = 0 \quad , \tag{8-17}$$

to determine B_d. Proceeding in this manner, we find that the optimum value is $B_d = -A_s$. The standard deviation of the optimally balanced aberration is $A_s/6\sqrt{5}$, which is a factor of 4 smaller than the standard deviation for $B_d = 0$. Since the standard deviation has been reduced by a factor of 4 by balancing spherical aberration with defocus, the optical tolerance has been increased by the same factor. Following Section 1.4, a defocus of $B_d = -A_s$ is introduced by observing the image in a plane that is farther from the exit pupil than the Gaussian image plane by $8F^2A_s$. Moreover, since $B_d = 0$ and $B_d = -2A_s$ correspond to Gaussian and marginal image planes, respectively, we note that, based on diffraction, the best image is obtained in a plane lying midway between them. This is different from the plane containing the circle of least confusion that corresponds to $B_d = -1.5A_s$.

Coma and astigmatism can be treated similarly. Table 8-2 lists the form of a balanced primary aberration, its standard deviation, and its tolerance for a Strehl ratio of 0.8. We note that in the case of coma, the balancing aberration is a wavefront tilt with a coefficient that is minus two-thirds of the coma coefficient. Thus, maximum Strehl ratio is obtained at a point that is displaced from the Gaussian image point by $(4FA_c/3, 0)$ but lies in the Gaussian image plane. By balancing coma with an appropriate amount of tilt, its standard deviation is reduced by a factor of 3. In the case of astigmatism, the best Strehl ratio is obtained in a plane that is farther than the Gaussian image plane by $4F^2A_a$. As discussed in Chapter 7, this is also the plane in which the circle of least confusion is obtained. By balancing with defocus, the standard deviation of astigmatism is reduced by a factor of 1.225. The point of observation with respect to which the aberration variance is minimum and, therefore, the irradiance at that point is maximum, is called the *diffraction focus.*

8.3.4 Comparison of Approximate and Exact Results

Figure 8-3 shows how the Strehl ratio of a primary aberration varies with its standard deviation. Approximate as well as exact results are shown in this figure.[8] The exact results are obtained by the use of Eq. 8-8. The curves for a given aberration and for the corresponding balanced aberration can be distinguished from each other by their behavior for large σ_w values (near 0.25λ). For example, coma is shown by the evenly dashed

curves; the higher dashed curve is for coma and the lower is for balanced coma. The same holds true for astigmatism. The curves for spherical and balanced spherical aberrations are identical since the Strehl ratio for a given value of σ_w is the same for the two aberrations. The following observations may be made from Figure 8-3:

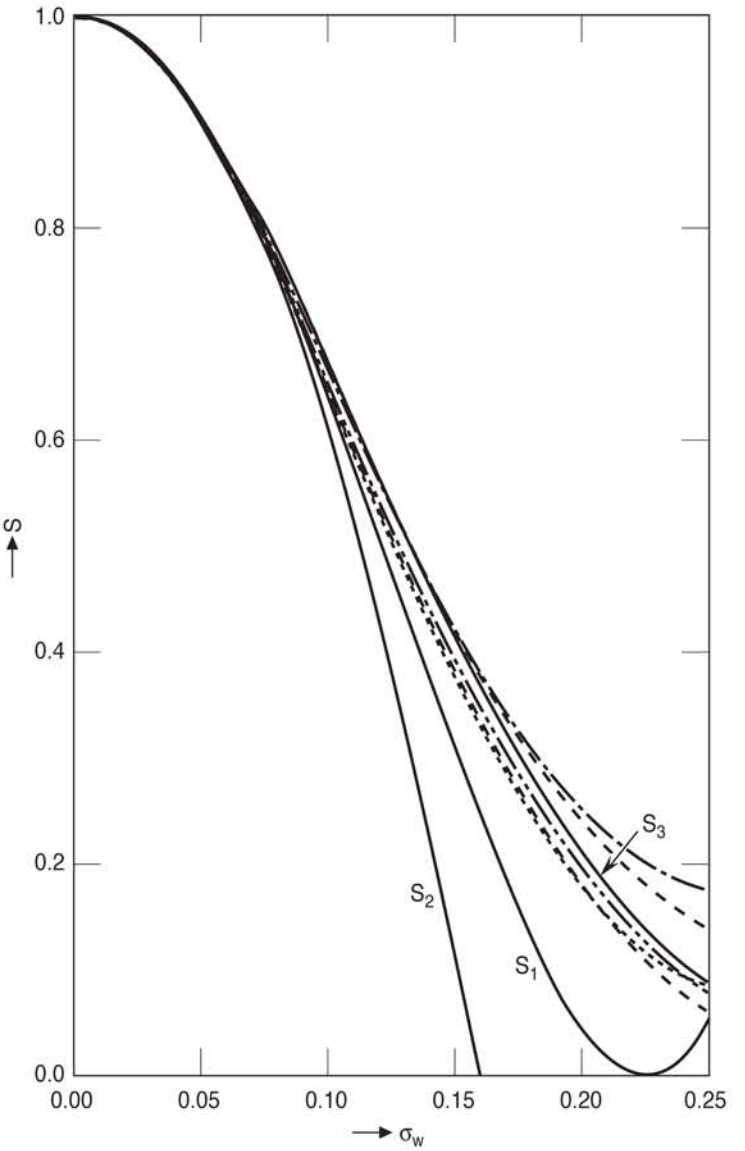


Figure 8-3. Strehl ratio for primary aberrations as a function of their standard deviation σ_w in units of optical wavelength λ. For large values of σ_w, coma and astigmatism give a higher Strehl ratio than the corresponding balanced aberration. The Strehl ratios for spherical and balanced spherical aberrations for the same value of σ_w are identical. $\sigma_\Phi = (2\pi/\lambda)\,\sigma_w$. Spherical or balanced spherical...., Coma– – – –, Astigmatism __.__.__. .

i. For small values of σ_w, the Strehl ratio is independent of the type of aberration. It depends only on its variance.

ii. The expressions for S_1 and S_2 underestimate the true Strehl ratio *S*.

iii. The expression for S_3 underestimates the true Strehl ratio only for coma and astigmatism; it overestimates for the other aberrations. Numerical analysis shows that the error, defined as $100(1 - S_3/S)$, is < 10% for $S > 0.3$.

iv. S_3 gives a better approximation for the true Strehl ratio than S_1 and S_2. The reason is that, for small values of σ_w, it is larger than S_1 by approximately $\sigma_\Phi^4/4$. Of course, S_1 is larger than S_2 by $\sigma_\Phi^4/4$.

v. The Strehl ratio depends strongly on the standard deviation of an aberration but weakly on its detailed distribution over a wide range of Strehl ratio values.

8.3.5 Strehl Ratio for Nonoptimally Balanced Aberrations

When a certain aberration is balanced with other aberrations to minimize its variance, the balanced aberration does not necessarily yield a higher or the highest possible Strehl ratio. For small aberrations, a maximum Strehl ratio should be obtained according to any of the Eqs. (8-13)–(8-15), when the variance is minimum. For large aberrations, however, there is no simple relationship between the Strehl ratio and the aberration variance. For example,[9] when $A_s = 3\lambda$, the optimum amount of defocus is $B_d = -3\lambda$ but the Strehl ratio is a minimum and equal to 0.12. The Strehl ratio is maximum and equal to 0.26 for $B_d \simeq -4\lambda$ or -2λ. For $A_s \lesssim 2.3\lambda$, the axial irradiance is maximum at a point with respect to which the aberration variance is minimum. Similarly, in the case of coma, the maximum irradiance in the image plane occurs at the point with respect to which the aberration variance is minimum only if $A_c \lesssim 0.7\lambda$, which in turn corresponds to $S \gtrsim 0.76$. For larger values of A_c, the distance of the point of maximum irradiance does not increase linearly with its value and even fluctuates in some regions.[10] Moreover, it is found that for $A_c > 2.3\lambda$, the Seidel coma gives a larger Strehl ratio than the balanced coma; i.e., the irradiance in the image plane at the origin is larger than at the point with respect to which the aberration variance is minimum. Thus, only for large Strehl ratios, the irradiance is maximum at the point associated with minimum aberration variance.

When *secondary spherical aberration* (varying as ρ^6) and *secondary coma* (varying as $\rho^5 \cos\theta$) are balanced with lower-order aberrations to minimize their variance, it is found[11] that a maximum of Strehl ratio is obtained only if its value comes out to be greater than about 0.5. Otherwise, a mixture of aberrations yielding a larger-than-minimum possible variance gives a higher Strehl ratio than the one provided by a minimum variance mixture.

8.3.6 Rayleigh's $\lambda/4$ Rule

Rayleigh[12] showed that a quarter-wave of primary spherical aberration reduces the irradiance at the Gaussian image point by 20%; i.e., the Strehl ratio for this aberration is

0.8. This result has brought forth the *Rayleigh's* $\lambda/4$ *rule*, namely, that a Strehl ratio of approximately 0.8 is obtained if the maximum absolute value of the aberration at any point in the pupil is equal to $\lambda/4$. A variant of this definition is that an aberrated wavefront that lies between two concentric spheres spaced a quarter-wave apart will give a Strehl ratio of approximately 0.8. Thus, instead of $|W_p| = \lambda/4$, we require $W_{p\ v} = \lambda/4$, where $|W_p|$ is the *peak absolute value* and $W_{p\ v}$ is the *peak-to-valley* value of the aberration. From Table 8-2, we note that a Strehl ratio of 0.8 is obtained for $|W_p| = \lambda/4 = W_{p\ v}$ for spherical aberration only. For other primary aberrations, distinctly different values of $|W_p|$ and $W_{p\ v}$ give a Strehl ratio of 0.8. In Table 8-2, $|W_p|$ and $W_{p\ v}$ are also given in terms of the aberration coefficient A_i. Thus, we see that it is advantageous to use σ_w for estimating the Strehl ratio. A Strehl ratio of $S \gtrsim 0.8$ is obtained for $\sigma_w \lesssim \lambda/14$.

8.3.7 Balanced Aberrations and Zernike Circle Polynomials

The phase aberration function of a system with a circular exit pupil for a certain point object can be expanded in terms of a complete set of *Zernike circle polynomials*[1,2] $Z_n^m(\rho,\theta)$ that are orthonormal over a unit circle in the form

$$\Phi(\rho,\theta) = \sum_{n=0}^{\infty}\sum_{m=0}^{n} c_{nm} Z_n^m(\rho,\theta) \quad , \quad 0 \le \rho \le 1 \quad , \quad 0 \le \theta \le 2\pi \quad , \tag{8-18}$$

where c_{nm} are the orthonormal expansion coefficients that depend on the location of the object, n and m are positive integers including zero, $n - m \ge 0$ and even, and

$$Z_n^m(\rho,\theta) = \left[2(n+1)/(1+\delta_{m0})\right]^{1/2} R_n^m(\rho)\cos m\theta \quad . \tag{8-19}$$

Here, δ_{ij} is a Kronecker delta, and

$$R_n^m(\rho) = \sum_{s=0}^{(n-m)/2} \frac{(-1)^s (n-s)!}{s!\left(\frac{n+m}{2}-s\right)!\left(\frac{n-m}{2}-s\right)!}\rho^{n-2s} \tag{8-20}$$

is a radial polynomial of degree n in ρ containing terms in ρ^n, ρ^{n-2}, …, and ρ^m. The radial circle polynomials $R_n^m(\rho)$ are even or odd in ρ, depending on whether n (or m) is even or odd. Also, $R_n^n(\rho) = \rho^n$, $R_n^m(1) = 1$, and $R_n^m(0) = \delta_{m0}$ for even $n/2$ and $-\delta_{m0}$ for odd $n/2$. The polynomials $R_n^m(\rho)$ obey the orthogonality relation

$$\int_0^1 R_n^m(\rho)\, R_{n'}^m(\rho)\,\rho\, d\rho = \frac{1}{2(n+1)}\delta_{nn'} \quad . \tag{8-21}$$

The orthogonality of the angular functions yields

$$\int_0^{2\pi} \cos m\theta \cos m'\theta\, d\theta = \pi(1+\delta_{m0})\delta_{mm'} \quad . \tag{8-22}$$

Therefore, the polynomials $Z_n^m(\rho,\theta)$ are orthonormal according to

$$\frac{1}{\pi}\int_0^1 \int_0^{2\pi} Z_n^m(\rho,\theta) Z_{n'}^{m'}(\rho,\theta)\rho\, d\rho\, d\theta = \delta_{nn'}\delta_{mm'} \quad . \tag{8-23}$$

The orthonormal Zernike expansion coefficients are given by

$$c_{nm} = \frac{1}{\pi}\int_0^1 \int_0^{2\pi} \Phi(\rho,\theta) Z_n^m(\rho,\theta)\rho\, d\rho\, d\theta \quad , \tag{8-24}$$

as may be seen by substituting Eq. (8-18) and utilizing the orthonormality of the polynomials.

The orthonormal Zernike polynomials and the names associated with some of them when identified with aberrations are listed in Table 8-3 for $n \le 8$. The polynomials independent of θ are the spherical aberrations, those varying as $\cos\theta$ are the coma aberrations, and those varying as $\cos 2\theta$ are the astigmatism aberrations. The number of Zernike (or orthogonal) aberration terms in the expansion of an aberration function through a certain order n is given by

$$N_n = \left(\frac{n}{2}+1\right)^2 \quad \text{for even } n \quad , \tag{8-25a}$$

$$= (n+1)(n+3)/4 \quad \text{for odd } n \quad . \tag{8-25b}$$

Each orthonormal expansion coefficient, with the exception of c_{00}, represents the standard deviation of the corresponding aberration term. The variance of the aberration function is accordingly given by

$$\sigma_\Phi^2 = \left\langle \Phi^2(\rho,\theta)\right\rangle - \left\langle \Phi(\rho,\theta)\right\rangle^2$$

$$= \sum_{n=0}^{\infty}\sum_{m=0}^{n} c_{nm}^2 - c_{00}^2$$

$$= \sum_{n=1}^{\infty}\sum_{m=0}^{n} c_{nm}^2 \quad . \tag{8-26}$$

Unless the mean value of the aberration $\langle\Phi\rangle = c_{00} = 0$, $\sigma_\Phi \neq \Phi_{rms}$, where $\Phi_{rms} = \left\langle \Phi^2\right\rangle^{1/2}$ is the rms value of the aberration.

As indicated in Table 8-3, the balanced aberrations can be identified with the Zernike circle polynomials. For example, Z_2^2, Z_3^1, and Z_4^0 represent balanced astigmatism, coma, and spherical aberration. For obvious reasons, a balanced aberration in this form may be referred to as a Zernike or an *orthogonal aberration.* The constant term in Z_4^0 makes its mean value to be zero. It does not change the standard deviation of the balanced aberration or the Strehl ratio corresponding to it. The circle polynomials are unique in the sense that they are the only polynomials that are orthogonal across a unit circle *and* represent balanced aberrations.

In a system without an axis of rotational symmetry, as for example in the case of fabrication errors, the aberration function will generally consist of terms not only in

Table 8-3. Orthonormal Zernike circle polynomials and balanced aberrations.

n	m	**Orthonormal Zernike Polynomial** $\left[\frac{2(n+1)}{1+\delta_{m0}}\right]^{1/2} R_n^m(\rho)\cos m\theta$	**Aberration Name***
0	0	1	Piston
1	1	$2\rho\cos\theta$	Distortion (tilt)
2	0	$\sqrt{3}\left(2\rho^2-1\right)$	Field curvature (defocus)
2	2	$\sqrt{6}\,\rho^2\cos 2\theta$	Primary astigmatism
3	1	$\sqrt{8}\left(3\rho^3-2\rho\right)\cos\theta$	Primary coma
3	3	$\sqrt{8}\,\rho^3\cos 3\theta$	
4	0	$\sqrt{5}\left(6\rho^4-6\rho^2+1\right)$	Primary spherical
4	2	$\sqrt{10}\left(4\rho^4-3\rho^2\right)\cos 2\theta$	Secondary astigmatism
4	4	$\sqrt{10}\,\rho^4\cos 4\theta$	
5	1	$\sqrt{12}\left(10\rho^5-12\rho^3+3\rho\right)\cos\theta$	Secondary coma
5	3	$\sqrt{12}\left(5\rho^5-4\rho^3\right)\cos 3\theta$	
5	3	$\sqrt{12}\,\rho^5\cos 5\theta$	
6	0	$\sqrt{7}\left(20\rho^6-30\rho^4+12\rho^2-1\right)$	Secondary spherical
6	2	$\sqrt{14}\left(15\rho^6-20\rho^4+6\rho^2\right)\cos 2\theta$	Tertiary astigmatism
6	4	$\sqrt{14}\left(6\rho^6-5\rho^4\right)\cos 4\theta$	
6	6	$\sqrt{14}\,\rho^6\cos 6\theta$	
7	1	$4\left(35\rho^7-60\rho^5+30\rho^3-4\rho\right)\cos\theta$	Tertiary coma
7	3	$4\left(21\rho^7-30\rho^5+10\rho^3\right)\cos 3\theta$	
7	5	$4\left(7\rho^7-6\rho^5\right)\cos 5\theta$	
7	7	$4\rho^7\cos 7\theta$	
8	0	$3\left(70\rho^8-140\rho^6+90\rho^4-20\rho^2+1\right)$	Tertiary spherical

*The words “orthonormal Zernike” are to be associated with these names, e.g., *orthonormal Zernike primary astigmatism*.

cos$m\theta$ but also in sin$m\theta$. The phase aberration function for such cases can be written in terms of orthonormal Zernike circle polynomials $Z_j(\rho,\theta)$ in the form

$$\Phi(\rho,\theta) = \sum_{j=1} a_j Z_j(\rho,\theta) \quad , \quad 0 \le \rho \le 1 \quad , \quad 0 \le \theta \le 2\pi \quad , \tag{8-27}$$

$$Z_{even\ j}(\rho,\theta) = \sqrt{2(n+1)}\, R_n^m(\rho)\cos m\theta, \ m \ne 0 \quad , \tag{8-28a}$$

$$Z_{odd\ j}(\rho,\theta) = \sqrt{2(n+1)}\, R_n^m(\rho)\sin m\theta, \ m \ne 0 \quad , \tag{8-28b}$$

$$Z_j(\rho,\theta) = \sqrt{n+1}\, R_n^0(\rho), \ m = 0 \quad . \tag{8-28c}$$

The index *j* is a polynomial-ordering number that is a function of both the radial degree *n* and the azimuthal frequency *m*. The polynomials are ordered such that an even *j* corresponds to a symmetric polynomial varying as $\cos m\theta$, and an odd *j* corresponds to an antisymmetric polynomial varying as $\sin m\theta$. A polynomial with a lower value of *n* is ordered first, and for a given value of *n*, a polynomial with a lower value of *m* is ordered first.

The polynomials are orthonormal according to

$$\int_0^1 \int_0^{2\pi} Z_j(\rho,\theta) Z_{j'}(\rho,\theta)\rho\, d\rho\, d\theta \Big/ \int_0^1 \int_0^{2\pi} \rho\, d\rho\, d\theta = \delta_{jj'} \quad . \tag{8-29}$$

The expansion coefficients are given by

$$a_j = \pi^{-1} \int_0^1 \int_0^{2\pi} \Phi(\rho,\theta) Z_{j'}(\rho,\theta)\rho\, d\rho\, d\theta \quad . \tag{8-30}$$

The variance of the aberration function is given by

$$\sigma_\Phi^2 = \sum_{j=1} a_j^2 - a_1^2$$

$$= \sum_{j=2} a_j^2 \tag{8-31}$$

The number of polynomials through a certain order *n* is given by

$$N_n = (n+1)(n+2)/2 \quad . \tag{8-32}$$

The first eleven such polynomials are listed in Table 8-4.

8.4 2D PSFs

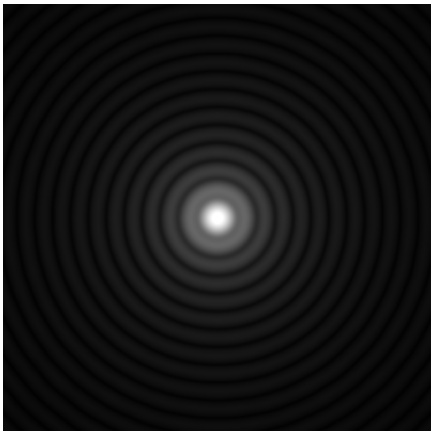
Now we show how the Airy pattern shown in Figure 8b, representing the 2D aberration-free PSF, is affected by a primary aberration. Our emphasis is to illustrate the structure of a PSF, i.e., on the distribution of its bright and dark regions, and not on the

Table 8-4. Orthonormal Zernike circle polynomials $Z_j(\rho, \theta)$.

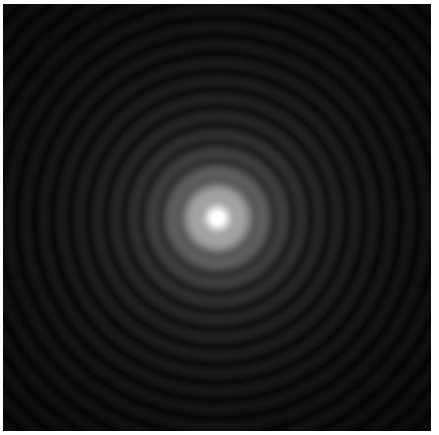
j	n	m	$Z_j(\rho, \theta)$	**Aberration***
1	0	0	1	Piston
2	1	1	$2\rho\cos\theta$	x tilt
3	1	1	$2\rho\sin\theta$	y tilt
4	2	0	$\sqrt{3}\left(2\rho^2 - 1\right)$	Defocus
5	2	2	$\sqrt{6}\,\rho^2\sin 2\theta$	45° Primary astigmatism
6	2	2	$\sqrt{6}\,\rho^2\cos 2\theta$	0° Primary astigmatism
7	3	1	$\sqrt{8}\left(3\rho^3 - 2\rho\right)\sin\theta$	Primary y coma
8	3	1	$\sqrt{8}\left(3\rho^3 - 2\rho\right)\cos\theta$	Primary x coma
9	3	3	$\sqrt{8}\,\rho^3\sin 3\theta$	
10	3	3	$\sqrt{8}\,\rho^3\cos 3\theta$	
11	4	0	$\sqrt{5}\left(6\rho^4 - 6\rho^2 + 1\right)$	Primary spherical

*The words “orthonormal Zernike circle” are to be associated with these names, e.g., *orthonormal Zernike circle* $0°$ *primary astigmatism*.

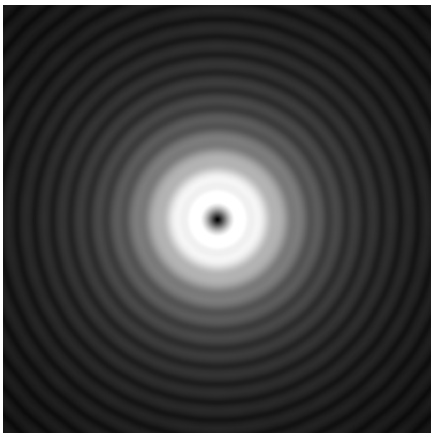
quantitative irradiance distribution. For example, we have accentuated some regions of very low irradiance to make their appearance visible. Some of the symmetry properties of the PSFs are clearly evident in these figures.[13] The PSFs for defocus (Figure 8-4) and spherical aberration (Figures 8-5 and 8-6) are radially symmetric like the Airy pattern. The central value of a PSF for an integral number of waves of defocus is zero, as may be seen from Eq. (8-7), yielding a dark center. The size of the central bright spot for spherical aberration does not change with increasing amount of spherical aberration.[14,15] The PSFs for astigmatism, shown in Figure 8-7, are symmetric about two orthogonal axes, one of them lying in the tangential plane. As the aberration increases, the diffraction PSFs begin to resemble the ray spots; elliptical spot in general, and line spots in particular. The PSFs in the minimum aberration-variance (circle of least confusion) plane are four-fold symmetric, as shown in Figure 8-8. The PSFs for coma are symmetric about the tangential plane, as shown in Figure 8-9. Thus, they have a line symmetry in any observation plane, the line lying in the tangential plane. It should be evident that a random mixture of various aberrations will yield a complex PSF.



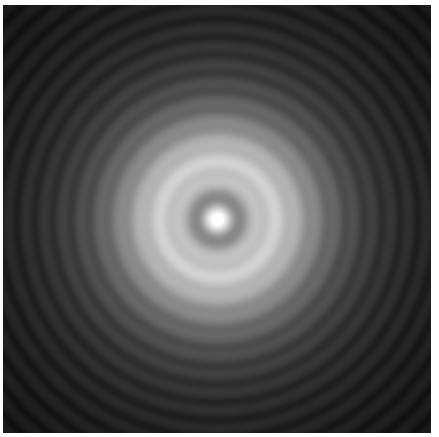
$B_d = 0$



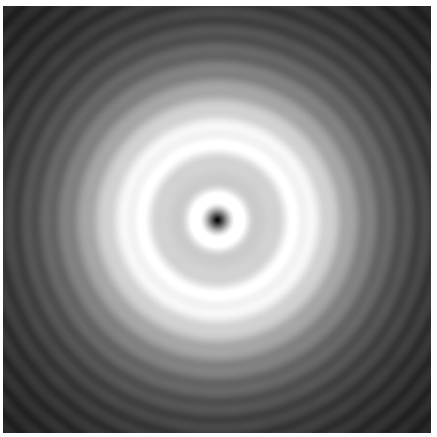
$B_d = 0.5$



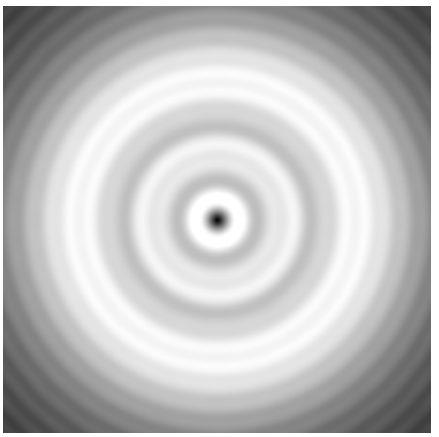
$B_d = 1$



$B_d = 1.5$



$B_d = 2$



$B_d = 3$

Figure 8-4. Defocused PSFs. B_d represents the peak value of defocus wave aberration in units of λ. The central value of the PSF is zero when B_d is equal to an integral number of wavelengths.

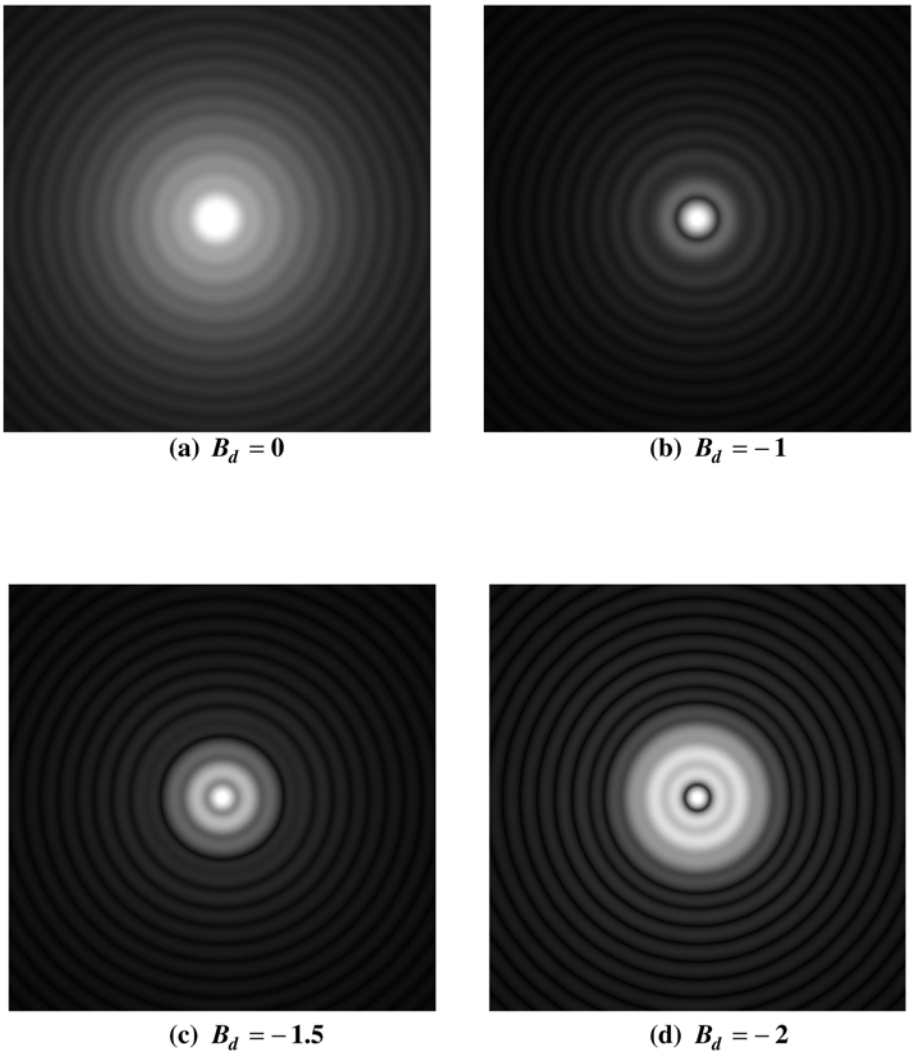
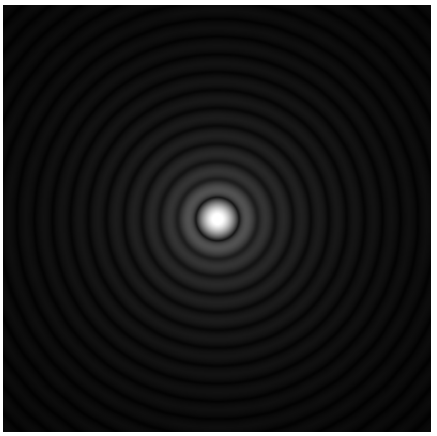
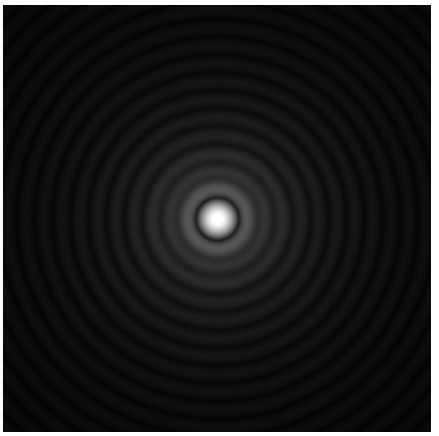


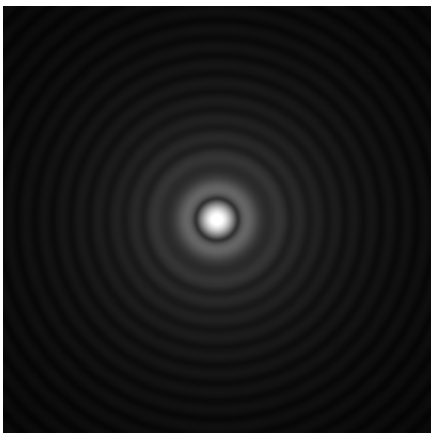
Figure 8-5. PSFs aberrated by one wave of spherical aberration ($A_s\rho^4$ with $A_s = 1\lambda$) observed in various defocused image planes indicated by B_d in units of λ. (a) Gaussian. (b) Minimum variance. (c) Least confusion. (d) Marginal.



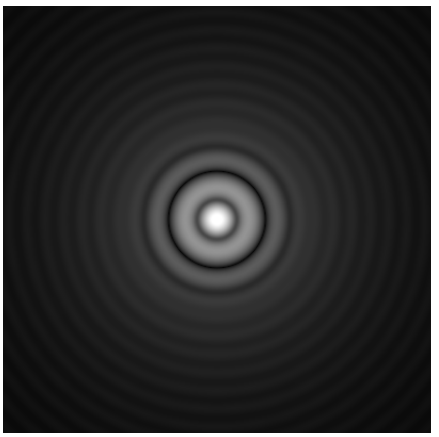
$A_s = 0.25$



$A_s = 0.5$



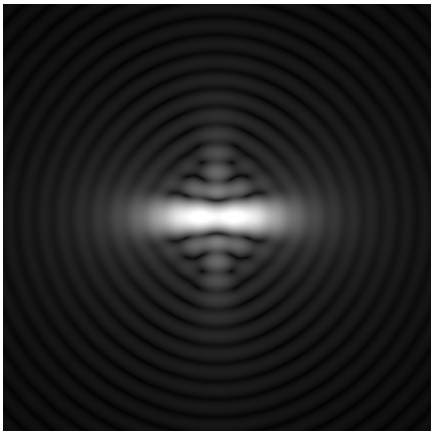
$A_s = 1$



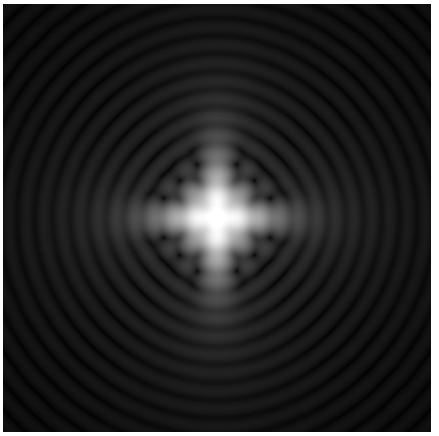
$A_s = 2$

$A_s = 3$

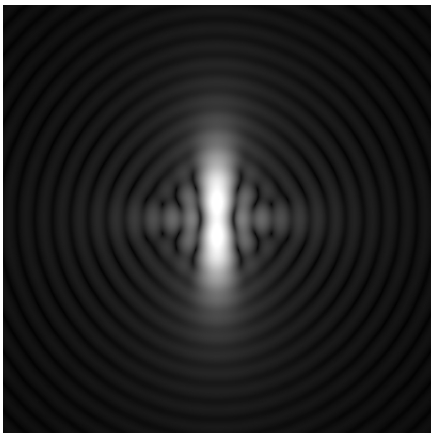
Figure 8-6. PSFs for balanced spherical aberration $\left[A_s\left(\rho^4 - \rho^2\right)\right]$. Thus, a PSF is observed in a defocused image plane corresponding to $B_d = -A_s$. The aberration coefficient A_s is in units of λ.



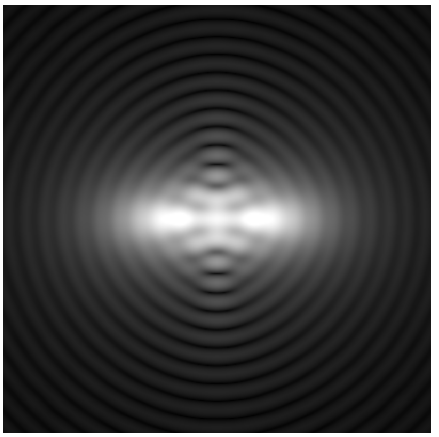
$A_a = 1,\ B_d = 0$



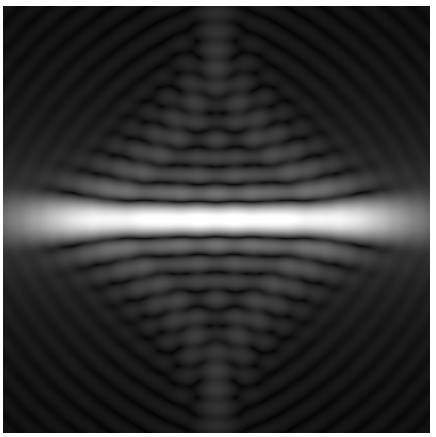
$A_a = 1,\ B_d = -0.5$



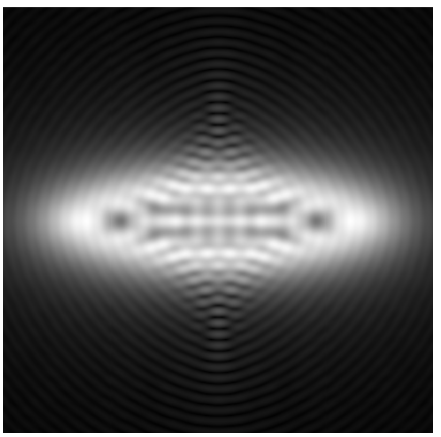
$A_a = 1,\ B_d = -1$



$A_a = 1,\ B_d = 0.5$



$A_a = 3,\ B_d = 0$



$A_a = 3,\ B_d = 1.5$

Figure 8-7. PSFs aberrated by astigmatism observed in various image planes. $B_d = 0$, $-A_a/2$, and $-A_a$ represent the Gaussian or sagittal, minimum-variance or circle-of-least (astigmatic)-confusion, and tangential image planes. The aberration coefficient A_a is in units of λ.

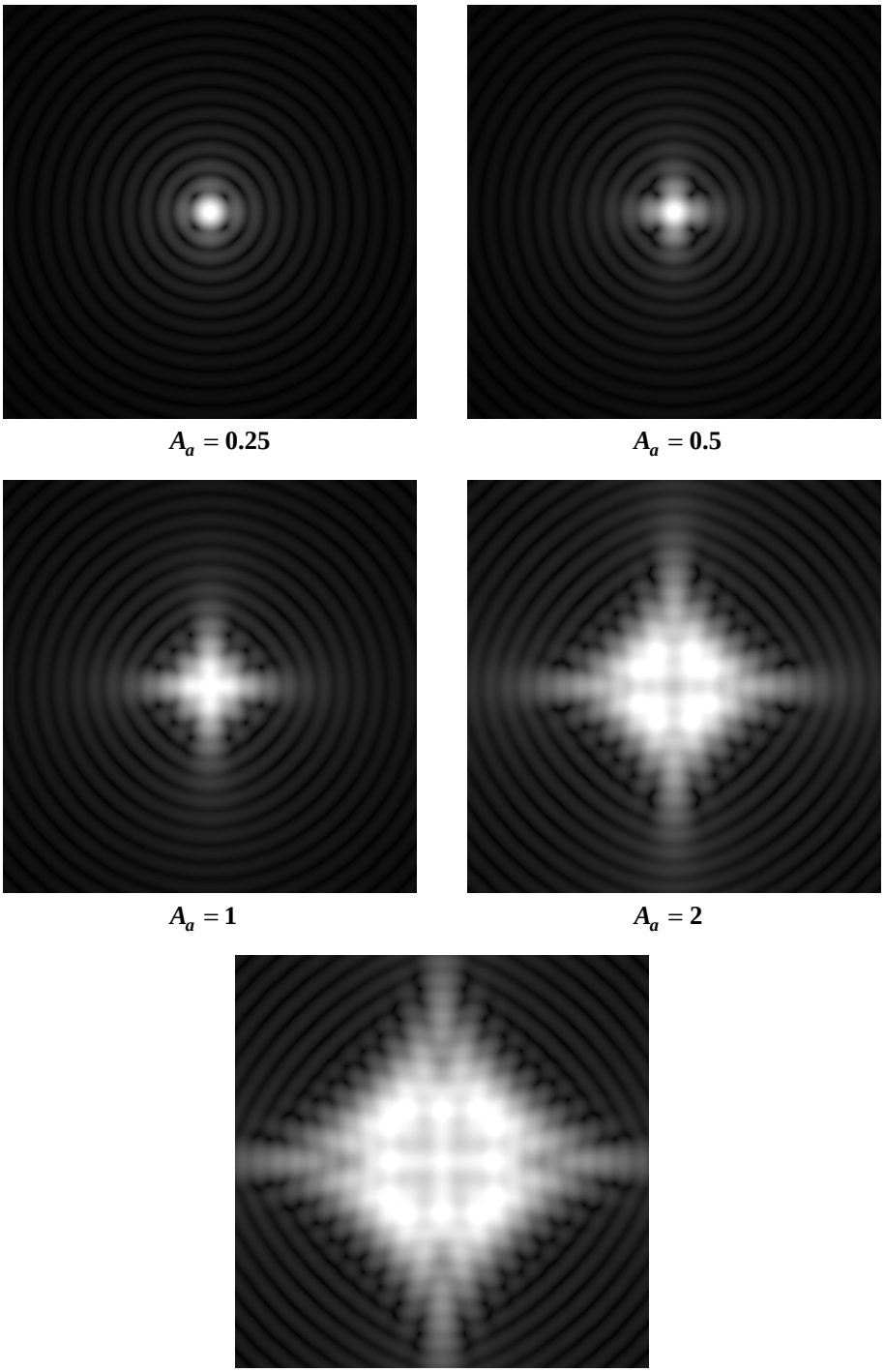


Figure 8-8. PSFs for balanced astigmatism $A_a\left(\rho^2\cos^2\theta - \rho^2/2\right)$. Thus, $B_d = -A_a/2$, and the PSFs are 4-fold symmetric. The aberration coefficient A_a is in units of λ.

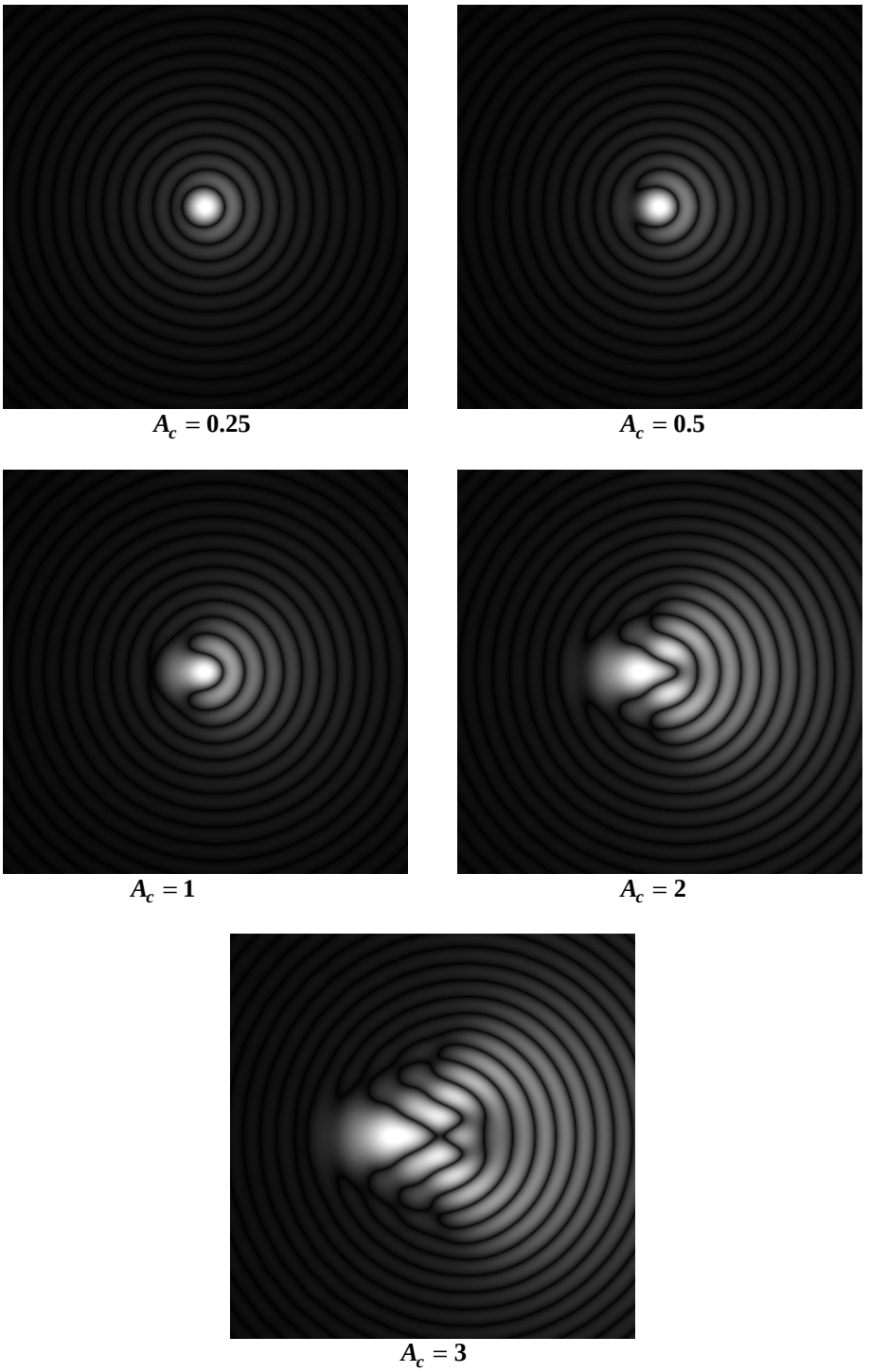


Figure 8-9. PSFs aberrated by increasing amount of coma $\left(A_c\rho^3\cos\theta\right)$. They are symmetric about the horizontal (x_i) axis. The peak and the centroid of the PSFs do not lie at the Gaussian image point. The aberration coefficient A_c is in units of λ.

8.5 OPTICAL TRANSFER FUNCTION (OTF)

Since the diffraction image of an incoherent object is given by the convolution of its Gaussian image and the system PSF, a Fourier transform of this relationship shows that the spatial frequency spectrum of the diffraction image is given by the product of the spectrum of the Gaussian image and the optical transfer function (OTF) of the system, where the OTF is equal to the Fourier transform of the PSF.[1-3] Because of the relationship of Eq. (8-1) between the PSF and the pupil function of the system, the OTF is also given by the autocorrelation of the pupil function. Thus, the OTF of a system can be obtained from its pupil function without having to calculate its PSF. In this section, we introduce the concept of OTF and discuss its physical significance. We also discuss how it is affected by aberrations and how it relates to the Strehl ratio. Also given is an expression for the aberration-free OTF of a system with a circular pupil. Contrast reversal is also illustrated, in which bright regions of certain bands of spatial frequencies in the object are imaged as dark, and dark regions are imaged as bright.

8.5.1 OTF and Its Physical Significance

The OTF of an incoherent imaging system is given by the Fourier transform of its PSF according to

$$\tau(\vec{v}_i) = \int PSF(\vec{r}_i) \exp(2\pi i \vec{v}_i \cdot \vec{r}_i)\, d\vec{r}_i \tag{8-33}$$

where $\vec{v}_i = (v_i, \phi)$ is a 2D *spatial frequency* vector in the image plane, $\vec{r}_i = (\lambda F r, \theta_i)$ is the position vector of a point in this plane, and the PSF is given by Eq. (8-1) with $P = 1$. In what follows, we assume that the Fresnel number of the system is large so that the defocus tolerance dictates that $z \simeq R$. However, if this is not the case, then we simply replace R by z in the following discussion. As mentioned above, because of Eq. (8-1) relating the PSF and the pupil function, the OTF may also be written as the autocorrelation of the pupil function, i.e.,

$$\tau(\vec{v}_i) = S_p^{-1} \int P(\vec{r}_p) P^*(\vec{r}_p - \lambda R \vec{v}_i)\, d\vec{r}_p \quad , \tag{8-34}$$

where

$$P(\vec{r}_p) = \exp[i\Phi(\vec{r}_p)] \quad , \quad 0 \le |\vec{r}_p| \le a \tag{8-35}$$

$$= 0 \quad , \quad \text{otherwise}$$

is the pupil function. Here, $\vec{r}_p = (a\rho, \theta)$ is the position vector of a point in the plane of the exit pupil. The integration in Eq. (8-34) is carried out across the region of overlap of two pupils centered at $\vec{r}_p = 0$ and the other at $\vec{r}_p = \lambda R \vec{v}_i$. The asterisk in Eq. (8-34) indicates a complex conjugate.

The OTF depends on the wavelength in two ways. First, the dependence of the phase aberration on it is evident. Second, it enters in the displacement of the pupil. It has the implication that for a longer wavelength the displacement approaches the diameter of the

pupil for a smaller frequency, thereby reducing the region of overlap of two pupils displaced with respect to each other to zero. Consequently, the *cutoff frequency* is smaller for a longer wavelength.

The physical significance of the OTF may be understood with the help of Figure 8-10. If we consider a sinusoidal object of a spatial frequency $\vec{v}_o$, modulation or contrast m, and phase δ, its Gaussian image is also sinusoidal with a spatial frequency $\vec{v}_i = \vec{v}_o / M$, where M is the magnification of the image. The modulation and phase of the Gaussian image are the same as those of the object. Its diffraction image is also sinusoidal with a spatial frequency $\vec{v}_i$. However, the modulation of the diffraction image is $m\left|\tau(\vec{v}_i)\right|$ and its phase is $\delta - \Psi(\vec{v}_i)$ where $\left|\tau(\vec{v}_i)\right|$ is the modulus of the OTF and $\Psi(\vec{v}_i)$ is its phase, i.e.,

$$\tau(\vec{v}_i) = \left|\tau(\vec{v}_i)\right| \exp\left[i\Psi(\vec{v}_i)\right] \quad , \tag{8-36}$$

The functions $\left|\tau(\vec{v}_i)\right|$ and $\Psi(\vec{v}_i)$ are also called the *modulation transfer function* (MTF) and the *phase transfer function* (PTF), respectively.

8.5.2 Aberration-Free OTF

It is evident from Eq. (8-34) that the aberration-free OTF represents the fractional overlap area of two circles whose centers are separated by a distance $\lambda R v_i$, as illustrated in Figure 8-11. The overlap area is equal to four times the difference between the area of

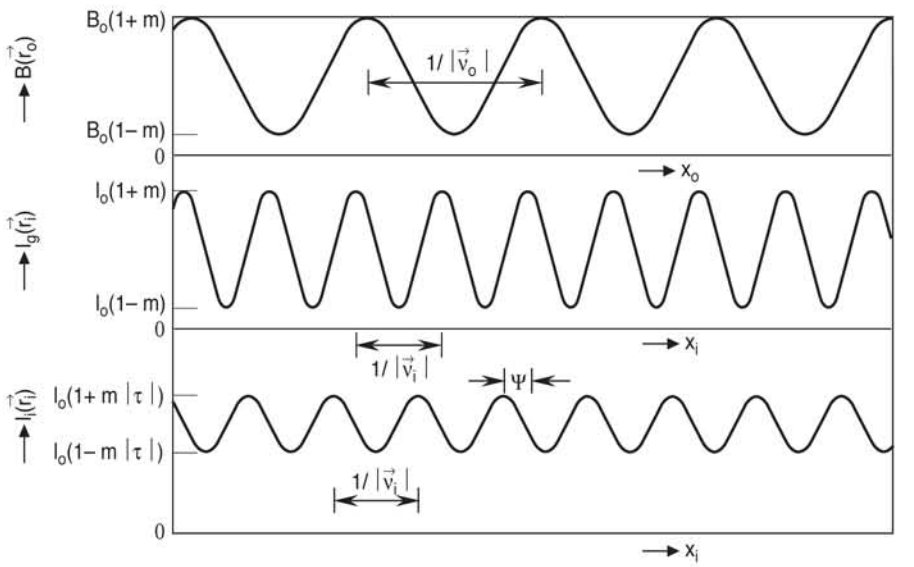


Figure 8-10. Image of a sinusoidal object of radiance B shown along the x axis. (a) Object, (b) Gaussian image of irradiance I_g, and (c) Diffraction image of irradiance I_i. B_0 and I_0 represent the average radiance and irradiance of the sinusoidal object and images, respectively.

a sector of radius a and cone angle β and the area of the triangle OAB. Thus the OTF is given by

$$\begin{aligned}\tau(v) &= (1/\pi)[2\beta - \sin 2\beta] \\ &= (2/\pi)\left[\cos^{-1} v - v(1-v^2)^{1/2}\right] \quad , \quad 0 \le v \le 1 \quad , \\ &= 0 \text{ , otherwise} \quad ,\end{aligned} \tag{8-37}$$

where

$$v = \cos\beta = v_i/(1/\lambda F) \tag{8-38}$$

is a normalized radial spatial frequency. The spatial frequency $v = 1$ or $v_i = 1/\lambda F$ is called the *cutoff frequency* of an incoherent imaging system, since the OTF for $v \ge 1$ is zero.

Figure 8-12 shows how the OTF given by Eq. (8-37) varies with v. We note that the OTF is radially symmetric; i.e.; its value depends on the magnitude of a spatial frequency but not on its direction. A system with a focal ratio F = 10 imaging an object radiating at a wavelength $\lambda = 0.5\,\mu$m corresponds to a cutoff frequency of 200 cycles/mm. The cutoff frequency decreases linearly with wavelength. The sinusoidal components of an object with spatial frequencies $v_o \ge M/\lambda F$ are not resolved by the system at all; i.e., their images are of uniform irradiance. From Eq. (8-37), we find that

$$\tau'(0) = \left[\frac{\partial \tau(v)}{\partial v}\right]_{v=0} = -4/\pi \tag{8-39}$$

and

$$\int_0^1 \tau(v)\, v\, dv = 1/8 \quad . \tag{8-40}$$

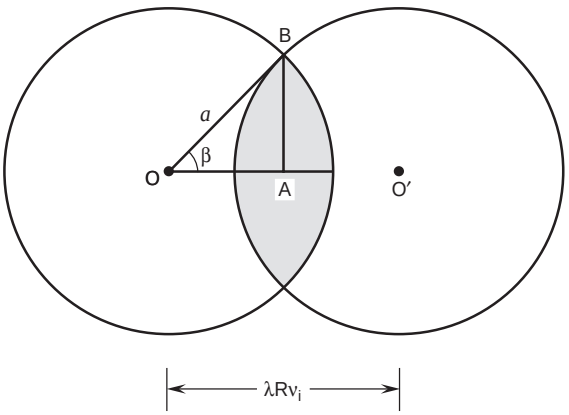


Figure 8-11. Aberration-free OTF as the fractional area of overlap of two circles whose centers are separated by a distance $\lambda R v_i$.

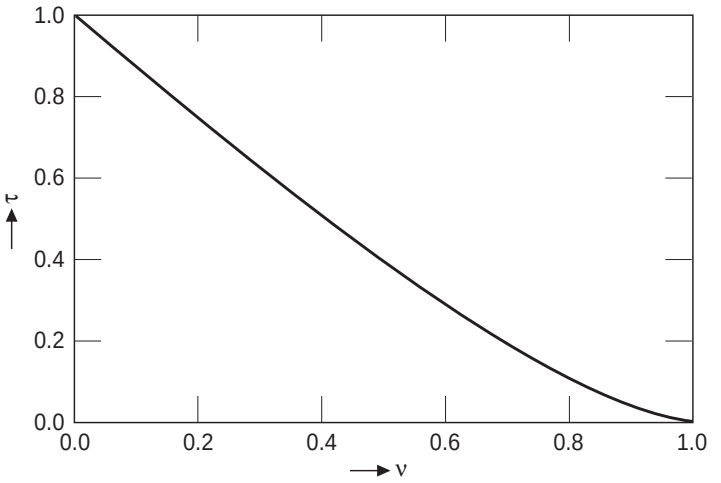


Figure 8-12. Aberration-free OTF, where the spatial frequency v is normalized by the cutoff frequency $1/\lambda F$.

Since the slope of the MTF or the real part of the OTF of a system evaluated at the origin is independent of its aberration, it is equal to $-4/\pi$ in the case of a circular pupil regardless of its aberration.

The Strehl ratio of an optical imaging system, discussed in Section 8.3, represents the ratio of its PSF (or the corresponding irradiance) values at the center $r = 0$ with and without aberration. From Eq. (8-33), we note that its PSF can be written as the inverse Fourier transform of its OTF, i.e.,

$$PSF\left(\vec{r}_i\right) = \int \tau\left(\vec{v}_i\right) \exp\left(-2\pi i\, \vec{v}_i \cdot \vec{r}_i\right) d\, \vec{v}_i \quad . \tag{8-41}$$

Accordingly, the Strehl ratio may be written

$$S = \left(4/\pi\right) \int \tau\left(\vec{v}_i\right) d\, \vec{v}_i \quad , \tag{8-42}$$

where we have used Eq. (8-40) for the integral involving the aberration-free OTF. Since S is a real quantity, the integral of the imaginary part of $\tau\left(\vec{v}_i\right)$ on the right-hand side of Eq. (8-42) must be zero. Hence, we may write Eq. (8-42) as

$$S = \left(4/\pi\right) \int \mathrm{Re}\, \tau\left(\vec{v}_i\right) d\, \vec{v}_i \quad , \tag{8-43}$$

where Re indicates a real part. Thus, the Strehl ratio of a system gives a measure of the mean value of the real part of its OTF, averaged over all spatial frequencies.

8.5.3 Hopkins Ratio and Aberration Tolerance

In Section 8.3, we calculated aberration tolerances for a system with a Strehl ratio of 0.8. Such a system forms the image of an object with a quality that is only slightly inferior to the corresponding quality for an aberration-free system, regardless of the spatial frequencies (or the size of detail) of interest in the object. A Strehl ratio of 0.8 is

obtained when the standard deviation of the aberration of the system across its exit pupil is approximately $\lambda/14$, regardless of the type of aberration. However, systems that have much larger aberrations form good-quality images of objects in which the size of the detail is much coarser than the limiting resolution $1/\lambda F$ of the system.

We now consider aberration tolerances based on a certain amount of reduction in the MTF of the system corresponding to a certain spatial frequency. Following Hopkins,[16] we define a modulation ratio $H(\vec{v}_i)$ as the ratio of the MTFs of a system at a spatial frequency $\vec{v}_i$ with and without aberration, i.e.,

$$H(\vec{v}_i) = \left|\tau(\vec{v}_i)\right| / \tau_u(\vec{v}_i) \quad , \tag{8-44}$$

where $\tau(\vec{v}_i)$ is the aberrated OTF, and $\tau_u(\vec{v}_i)$ is the aberration-free or unaberrated OTF given by Eq. (8-37) with v replaced by v_i and $0 \le v_i \le 1/\lambda F$. For obvious reasons, we call $H(\vec{v}_i)$ the *Hopkins modulation* or *contrast ratio.* Its value is ≤ 1, since the aberrated MTF is always less than its corresponding aberration-free value.

Based on numerical analysis for primary aberrations, Hopkins[16] has shown that $H(v) \gtrsim 0.8$ for $v \lesssim 0.1$, provided their coefficients obey the following conditions:

$$B_d \lesssim \pm \lambda/20v \quad , \tag{8-45}$$

$$A_a \lesssim \pm \lambda/10v \text{ in the plane } B_d = -A_a/2 \quad , \tag{8-46}$$

$$A_c \lesssim \pm \lambda\left(\frac{0.071}{v} + 0.16\right) \text{with } \Psi(v) = \mp 0.89 + 0.48v \text{ when } \phi = 0 \quad , \tag{8-47a}$$

$$\lesssim \pm \lambda\left(\frac{0.123}{v} + 0.19\right) \text{with } \Psi(v) = 0 \text{ when } \phi = \pi/2 \quad , \tag{8-47b}$$

and

$$A_s \lesssim \pm \lambda\left(\frac{0.106}{v} + 0.33\right) \text{in the plane } B_d = -\left(1.33 - 2.2v + 2.8v^2\right)A_s \quad . \tag{8-48}$$

As in Eq. (1-7), B_d, A_a, A_c, and A_s represent the peak coefficients of defocus, astigmatism, coma, and spherical aberration. We note that the amount of balancing defocus in the case of spherical aberration is different from its corresponding value given in Table 8-2 for optimizing the Strehl ratio. Moreover, its value depends on the magnitude of the spatial frequency at which the MTF is optimized. For spatial frequencies $v > 0.1$, it is more appropriate to use the Strehl ratio as the criterion of image quality and aberration balancing.

8.5.4 Contrast Reversal

Figure 8-13 shows how the OTF of a defocused system varies with the spatial frequency. We note that it is real and radially symmetric; i.e., its value depends on the

value of v but not on the value of ϕ. For $B_d \leq 0.64\,\lambda$, the OTF is positive for all spatial frequencies. However, for larger values of B_d, it becomes negative, corresponding to a PTF of π, for certain bands of spatial frequencies. It becomes negative for smaller and smaller spatial frequencies as the amount of defocus B_d increases. The OTF is independent of the sign of B_d (when the Fresnel number of the system is large).

To illustrate the significance of the OTF and, in particular, the *contrast reversal*, we consider, as shown in Figure 8-14a, a 2D object that is sinusoidal along the vertical axis with a spatial frequency that increases linearly in the horizontal direction. The maximum frequency in the object is chosen to equal the cutoff frequency of the aberration-free system. This frequency is normalized to unity. The aberration-free or the diffraction-limited image of the object is shown in Figure 8-14b. The monotonic reduction in contrast with increasing spatial frequency is quite evident from this figure. A defocused image corresponding to $B_d = 2\,\lambda$ is shown in Figure 8-14c. It is clear that the contrast in the image reduces with frequency rapidly to zero, reverses its sign back and forth as the frequency increases, with practically zero values for frequencies $v \gtrsim 0.3$. As a convenience, the aberration-free and defocused OTFs are shown in Figure 8-14d to illustrate the regions of zero and near zero contrast as well as the regions of contrast reversal.

The OTF for a symmetric aberration, such as spherical aberration or astigmatism, is real. However, the OTF for coma is a complex function with real and imaginary parts, or an MTF and a PTF. As stated in Section 8.4, the Strehl ratio is obtained by integrating the real part of the OTF, and the integral of its imaginary part is zero. The spatial frequency bands for which the contrast reversal takes place depends on the type and the magnitude of an aberration.

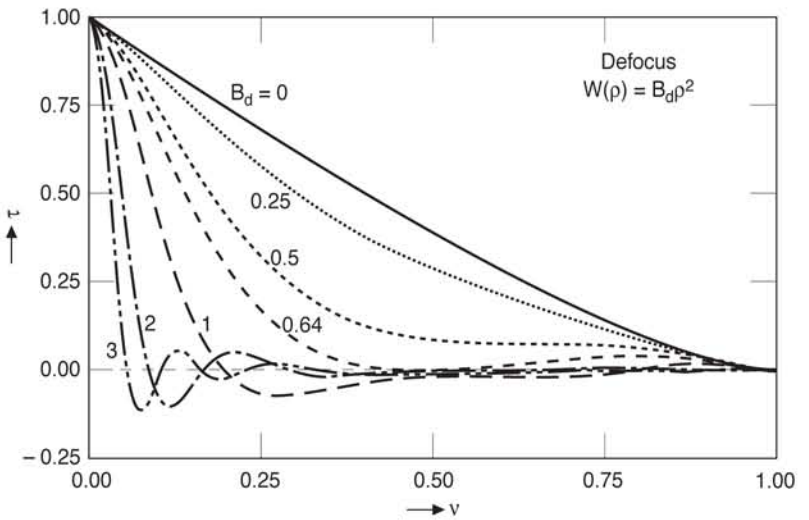
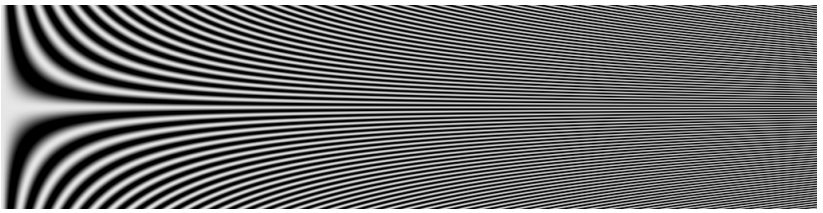
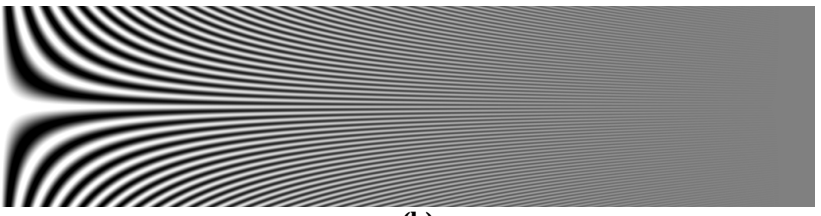


Figure 8-13. OTFs of a defocused system. B_d represents the peak defocus aberration in units of λ.

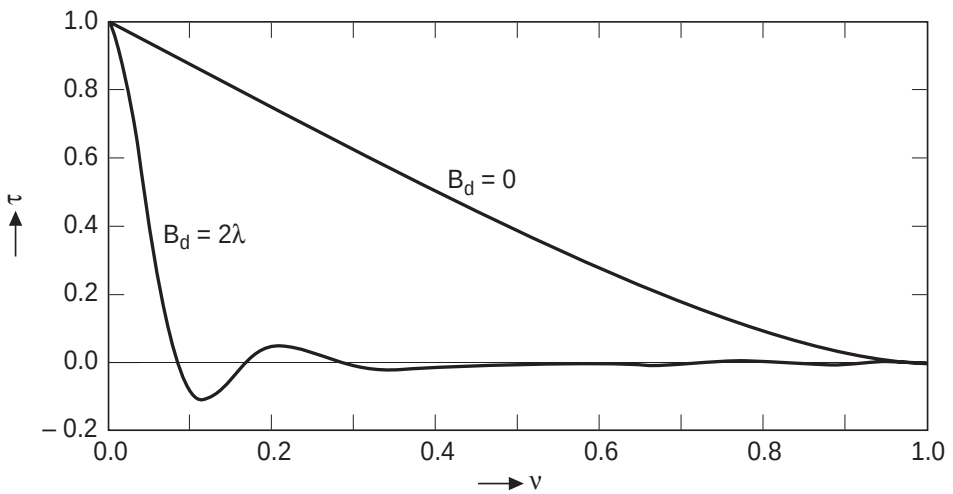


(a)



(b)

(c)



(d)

Figure 8-14. Aberration-free and defocused images of an object illustrating contrast reversal. (a) Sinusoidal object along the vertical axis with a spatial frequency that increases linearly in the horizontal direction. The Gaussian image is identical, except for any magnification. (b) Aberration-free image. (c) Defocused image with $B_d = 2\lambda$. (d) Aberration-free and defocused OTFs.

8.6 SUMMARY

While the amount of light in the image of a point object is determined by the entrance pupil of the system, its distribution is determined by the diffraction of that light at the exit pupil. The aberration-free image of a point object formed by a system with a circular exit pupil is called the Airy pattern, as described in Section 8.2.2. It consists of a bright central disc surrounded by alternating dark and bright diffraction rings. The central point of the disc is the brightest, coinciding with the center of curvature of the spherical wavefront, because that is where the Huygens' secondary wavelets interfere constructively.

When the wavefront is not spherical and, therefore, aberrations are present, the Huygens' secondary wavelets are not in phase and interfere (partially) destructively to yield a smaller irradiance at the center. The ratio of the irradiances at the center with and without aberration is called the Strehl ratio. As should be evident, the Strehl ratio is always less than or equal to one. For small aberrations, its value can be estimated from the aberration variance according to Eq. (8-15). Since the smaller the variance, the higher the Strehl ratio, we combine a given aberration with one or more lower-order aberrations to minimize its variance and thereby maximize the Strehl ratio. Combining aberrations in this manner is called aberration balancing for improving image quality. Thus, for example, spherical aberration or astigmatism is combined with defocus to improve the Strehl ratio or increase the aberration tolerance.

The higher-order aberrations, e.g., secondary aberrations, can be balanced in a similar manner to reduce their variance. The balanced aberrations thus obtained can be identified with the corresponding Zernike circle polynomials. These polynomials are unique in the sense that they are not only orthogonal across a unit circle, but also represent balanced aberrations for a circular pupil. These polynomials are given in their orthonormal form in this chapter so that when an aberration function is expanded in terms of them, the expansion coefficients represent the standard deviation of the corresponding aberration term.

The image of an incoherent object can be obtained by adding the irradiance images of its object elements. This imaging can also be considered in the spatial frequency domain. Thus the object is broken into its sinusoidal spatial frequency components, the image of each component is determined, and they are all summed or combined to determine the overall image of the object. The image of each sinusoidal object component is also sinusoidal, except that its contrast is lower and its phase may be shifted. In aberration-free imaging, the phase shifts are zero. Every system has a cutoff spatial frequency (due to the finite size of its exit pupil) above which the contrast is zero. Another way of saying this is that the spatial frequencies above a certain value are not transmitted by the system, or are missing from the image, thus limiting the resolution of the system. The aberrations reduce the contrast at every spatial frequency and introduce phase shifts depending on the value of the frequency. Thus, while the image generally resembles the object, it is not identical. Depending on the type and the magnitude of the

aberration, the phase shift for certain spatial frequency bands can be π, resulting in contrast reversal of those frequency bands. The dark regions of an object are imaged as bright and bright regions are imaged as dark, as illustrated in Figure 8-14. The ratio of contrasts at a certain spatial frequency with and without aberration is called the Hopkins ratio at that spatial frequency. The balanced aberrations obtained for maximizing the Hopkins ratio are different from those for maximizing the Strehl ratio.

References

1. V. N. Mahajan, *Optical Imaging and Aberrations*, Part II: *Wave Diffraction Optics* (SPIE Press, Bellingham, WA, Second Edition 2011).

2. M. Born and E. Wolf, *Principles of Optics*, 7th ed. (Cambridge University Press, New York, 1999).

3. J. W. Goodman, *Introduction to Fourier Optics*, 2nd ed. (McGraw-Hill, New York, 1996).

4. G. B. Airy, "On the diffraction of an object-glass with circular aperture," *Trans. Camb. Phil. Soc*. **5**, 283–291 (1835).

5. V. N. Mahajan, "Axial irradiance and optimum focusing of laser beams," *Appl. Opt.* **22**, 3042–3053 (1983).

6. A. Maréchal, "Etude des effets combines de la diffraction et des aberrations geometriques sur 1'image d'un point lumineux,"*Rev. d'Opt*. **26**, 257–277 (1947).

7. B. R. A. Nijboer, "The diffraction theory of aberrations," Ph.D. thesis (University of Groningen, Groningen, The Netherlands, 1942), p. 17.

8. V N. Mahajan, "Strehl ratio for primary aberrations in terms of their aberration variance," *J. Opt. Soc. Am*. **73**, 860–861 (1983).

9. V N. Mahajan, "Strehl ratio for primary aberrations: some analytical results for circular and annular pupils," *J. Opt. Soc. Am*. **72**, 1258–1266 (1982).

10. V N. Mahajan, "Line of sight of an aberrated optical system," *J. Opt. Soc. Am*. *A* **2**, 833–846 (1985).

11. W. B. King, "Dependence of the Strehl ratio on the magnitude of the variance of the wave aberration," *J. Opt. Soc. Am*. **58**, 655–661 (1968).

12. Lord Rayleigh, "Investigations in optics, with special reference to the spectroscope," *Phil. Mag.* (5) **8**, 403–411 (1879); also his *Scientific Papers*, Vol. 1 (Dover, New York, 1964), p. 432.

13. V. N. Mahajan, "Symmetry properties of aberrated point-spread functions," *J. Opt. Soc. Am. A* **11**, 1993–2003 (1994).

14. V N. Mahajan, "Aberrated point spread functions for rotationally symmetric aberrations," *Appl. Opt.* **22**, 3035–3041 (1983).

15. S. Szapiel, "Aberration-variance-based formula for calculating point-spread functions: rotationally symmetric aberrations," *Appl. Opt*. **25**, 244–251 (1986).

16. H. H. Hopkins, "The aberration permissible in optical systems," *Proc. Phys. Soc*. (London) B**52,** 449–470 (1957).

CHAPTER 9

Systems with Annular and Gaussian Pupils

9.1 INTRODUCTION

In Chapter 8 we have considered optical systems with circular exit pupils. Now we consider systems with *annular pupils*, for example, a Cassegrain telescope in which its secondary mirror *obscures* the central portion of its primary mirror. As in the case of a system with a circular pupil, we discuss the aberration-free PSF, axial irradiance, and the Strehl ratio of a system with an annular pupil. We show that the radius of the central bright spot of the PSF decreases, its principal or central maximum decreases in its value, and the secondary maxima increase in their values as the obscuration increases.

However, the tolerance for a given Strehl ratio increases or decreases depending on the type of the aberration. The Zernike annular polynomials representing balanced aberrations are also discussed, but the aberrated PSFs and OTFs are not. Optical systems with circular pupils and *Gaussian illumination* across them are also considered along similar lines. For these systems, it is shown that the tolerance for an aberration increases compared to the corresponding tolerance for a system with a uniformly illuminated circular pupil. Finally, systems with *weakly truncated Gaussian pupils*, i.e., those having a very wide pupil compared to the width or the radius of the Gaussian illumination, are considered. In this case, a Gaussian beam propagates as a Gaussian, and the tolerance for a primary aberration is obtained in terms of its peak value at the *Gaussian radius* rather than at the edge of the pupil.

9.2 ANNULAR PUPILS

In this section, we discuss the imaging characteristics of systems with annular pupils. The aberration-free PSF, encircled power, axial irradiance, and Strehl ratio are discussed for increasing value of the obscuration of the pupil. The results obtained are compared with the corresponding results for systems with circular pupils.

9.2.1 Aberration-Free PSF

Consider a system with an annular exit pupil having inner and outer radii of ϵa and a, where ϵ is called its *obscuration ratio*, as illustrated in Figure 9-1. The PSF of the system, i.e., the irradiance distribution of the image of a point object formed by it, is given by Eq. (8-1) except that now the lower limit in the radial integration is ϵ instead of zero. The aberration-free PSF thus obtained is given by

$$I(r;\epsilon) = \frac{1}{\left(1-\epsilon^2\right)^2}\left[\frac{2J_1(\pi r)}{\pi r} - \epsilon^2\,\frac{2J_1(\pi\epsilon r)}{\pi\epsilon r}\right]^2 , \tag{9-1}$$

where $J_1(\cdot)$ is the first-order Bessel function of the first kind. It is normalized to unity at the center $r = 0$ by the central irradiance $PS_p/\lambda^2 R^2$, where P is the total power

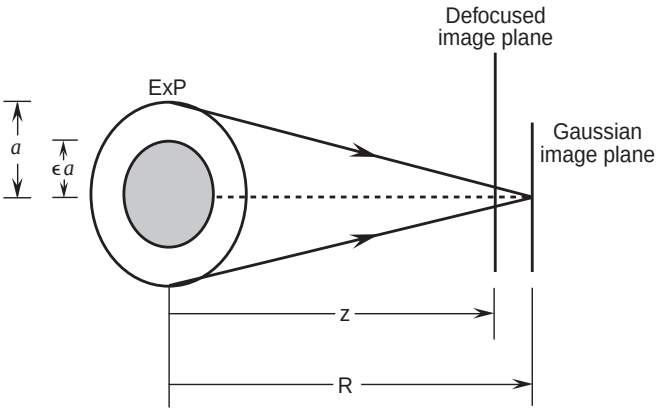


Figure 9-1. Imaging by a system with an annular exit pupil of inner and outer radii ϵa and a, respectively.

transmitted by the annular pupil, $S_p = \pi a^2\left(1-\epsilon^2\right)$ is its clear area, λ is the wavelength of object radiation, and R is the distance between the pupil plane and the Gaussian image plane.

Note that r is in units of λF, as in the case of a circular pupil, where $F = R/2a$ is the *focal ratio* or the *f-number* of the image-forming light cone. For a given total power P, the value of the *central maximum* decreases as $1-\epsilon^2$ as ϵ increases due to the decrease in the clear pupil area. However, if the irradiance of the pupil is held constant, then the total power P also decreases as $1-\epsilon^2$ and, therefore, the central irradiance decreases as $\left(1-\epsilon^2\right)^2$ as ϵ increases.

The *minima* of the distribution have a value of zero at r values given by

$$J_1(\pi r) = \epsilon J_1(\pi\epsilon r),\ r \neq 0 \quad . \tag{9-2a}$$

Its *maxima* occur at r values given by

$$J_2(\pi r) = \epsilon^2 J_2(\pi\epsilon r),\ r \neq 0 \quad , \tag{9-2b}$$

where $J_2(\cdot)$ is a second-order Bessel function of the first kind. By integrating the irradiance distribution across a circle of radius r_c, we obtain the encircled power $P(r_c)$. Both the irradiance and encircled-power distributions are shown in Figure 9-2 for several values of ϵ. We note that the radius of the central bright disc (first dark ring corresponding to the first minimum) decreases as ϵ increases. It can be shown that as $\epsilon \to 1$, the irradiance distribution approaches $J_0^2(\pi r)$. Its first zero occurs at 0.76 compared to a value of 1.22 [first zero of $J_1(\pi r)$ when ϵ = 0]. The values of the secondary maxima of a distribution relative to the value of its principal maximum at r = 0 become higher as ϵ increases. For example, when ϵ = 0.5, the first secondary maximum

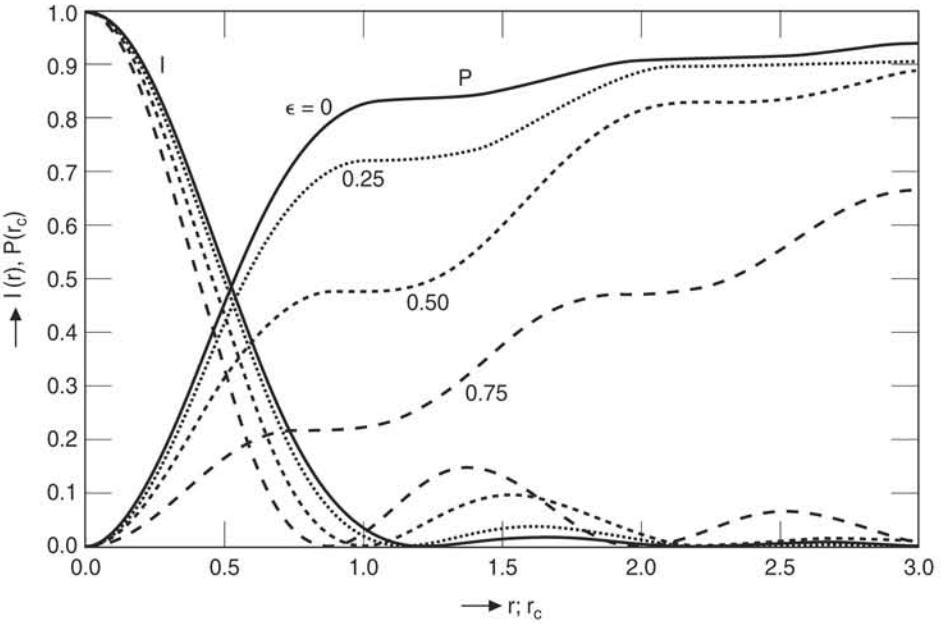


Figure 9-2. Irradiance and encircled-power distributions for an annular pupil. ϵ is the obscuration ratio of the pupil. The example of a circular pupil is shown for comparison.

has a value of 9.63% of the principal maximum compared to a value of 1.75% for a circular pupil. The minima, maxima, and the corresponding irradiances and encircled powers are given in Table 9-1 for $\epsilon = 0(0.1)0.9$.

An interesting observation[1] comes about when the irradiance distribution is considered for large values of r and large values of ϵ. Figure 9-3a shows the distributions for $\epsilon = 0$, 0.5, 0.8, and $\epsilon \to 1$. A corresponding picture of these distributions is shown in Figure 9-3b. We note that, for a circular pupil, the distribution consists of maxima and minima indicating a bright central disc surrounded by dark and bright rings. The successive maxima decrease in value monotonically. However, for an annular pupil, the distribution consists of not only the dark and bright rings but also of a *periodic ring group structure.* The number of maxima in a period is given by $n = 2/(1-\epsilon)$, which is equal to the ratio of the outer diameter and the width of the annulus, provided that n is an integer. The distribution is divided into ring groups. The group minima are the lowest ring maxima and correspond to ring numbers that are multiples of n, e.g., 10, 20, 30, etc., for $\epsilon = 0.8$. The radius of a ring group is also a multiple of n (in units of λF) since the spacing between two successive maxima or minima is approximately unity. The central bright spot or the first dark ring of radius 1.22 contains 83.8% of the total power in the image when $\epsilon = 0$. For $\epsilon = 0.8$, as may be seen from Table 9-1, the first dark ring has a radius of 0.85 and contains only 17.2% of the total power. However, the central ring group in this case has a radius of 10.10 and contains 90.3% of the total power.

Table 9-1. Positions r of PSF maxima and minima for an annular pupil in units of λF, and the corresponding irradiance and the encircled power.

ϵ		0.0			0.1			0.2			0.3			0.4	
Max/ Min	r, r_C	$I(r)$	$P(r_C)$	r, r_C	$I(r)$	$P(r_C)$	r, r_C	$I(r)$	$P(r_C)$	r, r_C	$I(r)$	$P(r_C)$	r, r_C	$I(r)$	$P(r_C)$
Max	0	1	0	0	1	0	0	1	0	0	1	0	0	1	0
Min	1.22	0	0.838	1.21	0	0.818	1.17	0	0.764	1.11	0	06.82	1.06	0	0.584
Max	1.63	0.0175	0.867	1.63	0.0206	0.853	1.63	0.0304	0.818	1.61	0.0475	0.766	1.58	0.0707	0.702
Min	2.23	0.	0.910	2.27	0	0.906	2.36	0	0.900	2.42	0	0.899	2.39	0	0.885
Max	2.68	0.0042	0.922	2.68	0.0031	0.914	2.69	0.0015	0.904	2.73	0.0011	0.902	2.77	0.0033	0.893
Min	3.24	0	0.938	3.18	0	0.925	3.09	0	0.908	3.10	0	0.904	3.30	0	0.903
Max	3.70	0.0016	0.944	3.70	0.0024	0.936	3.68	0.0037	0.926	3.64	0.0028	0.916	3.66	0.0007	0.905
Min	4.24	0	0.952	4.32	0	0.949	4.37	0	0.947	4.22	0	0.929	4.04	0	0.907
Max	4.71	0.0008	0.957	4.71	0.0004	0.951	4.74	0.0004	0.949	4.75	0.0016	0.938	4.66	0.0028	0.922
Min	5.24	0	0.961	5.15	0	0.953	5.16	0	0.951	5.42	0	0.949	5.31	0	0.939
Max	5.72	0.0004	0.964	5.71	0.0008	0.959	5.69	0.0006	0.955	5.73	0.0001	0.950	5.79	0.0008	0.944
Min	6.24	0	0.968	6.35	0	0.965	6.23	0	0.959	6.07	0	0.950	6.43	0	0.950
Max	6.72	0.0003	0.970	6.73	0.0001	0.966	6.74	0.0004	0.962	6.67	0.0006	0.955	6.72	0.0001	0.950
Min	7.25	0	0.972	7.14	0	0.967	7.35	0	0.966	7.27	0	0.961	7.03	0	0.950
Max	7.73	0.0002	0.974	7.72	0.0003	0.970	7.72	0.0001	0.967	7.77	0.0003	0.963	7.65	0.0004	0.954
Min	8.25	0	0.975	8.34	0	0.974	8.11	0	0.967	8.38	0	0.966	8.22	0	0.958
Max	8.73	0.0001	0.977	8.74	0.0001	0.975	8.72	0.0003	0.971	8.72	0.0000	0.966	8.77	0.0004	0.962
Min	9.25	0	0.978	9.16	0	0.975	9.38	0	0.974	9.06	0	0.967	9.46	0	0.966
Max	9.73	0.0001	0.979	9.72	0.0001	0.977	9.75	0.0000	0.975	9.70	0.0002	0.970	9.78	0.0000	0.966
Min	10.25	0	0.980	10.30	0	0.979	10.16	0	0.975	10.32	0	0.973	10.13	0	0.966

ϵ		0.5			0.6			0.7			0.8			0.9	
Max/ Min	r, r_C	$I(r)$	$P(r_C)$	r, r_C	$I(r)$	$P(r_C)$	r, r_C	$I(r)$	$P(r_C)$	r, r_C	$I(r)$	$P(r_C)$	r, r_C	$I(r)$	$P(r_C)$
Max	0	1	0	0	1	0	0	1	0	0	1	0	0	1	0
Min	1.000	0	0.479	0.95	0	0.372	0.90	0	0.269	0.85	0	0.172	0.81	0	0.082
Max	1.54	0.0963	0.618	1.48	0.1203	0.512	1.41	0.1395	0.389	1.35	0.1527	0.256	1.28	0.1600	0.124
Min	2.29	0	0.829	2.17	0	0.717	2.06	0	0.560	1.95	0	0.376	1.85	0	0.184
Max	2.76	0.0124	0.859	2.69	0.0306	0.784	2.58	0.0533	0.649	2.47	0.0734	0.456	2.35	0.0861	0.229
Min	3.49	0	0.901	3.39	0	0.873	3.22	0	0.761	3.05	0	0.554	2.90	0	0.284
Max	3.78	0.0004	0.902	3.84	0.0045	0.886	3.74	0.0192	0.808	3.57	0.0401	0.619	3.40	0.0566	0.328
Min	4.12	0	0.903	4.52	0	0.902	4.38	0	0.865	4.16	0	0.695	3.95	0	0.379
Max	4.50	0.0009	4.80	0.0001		0.903	4.86	0.0050	0.880	4.68	0.0218	0.741	4.46	0.0404	0.421
Min	5.05	0	0.910	5.11	0	0.903	5.52	0	0.899	5.27	0	0.795	5.00	0	0.468
Max	5.66	0.0022	0.923	5.58	0.0004	0.905	5.91	0.0005	0.901	5.78	0.0110	0.824	5.51	0.0299	0.507
Min	6.30	0	0.938	6.00	0	0.906	6.47	0	0.903	6.37	0	0.857	6.05	0	0.549
Max	6.81	0.0008	0.943	6.61	0.0016	0.916	6.72	0.000	0.903	6.87	0.0048	0.872	6.56	0.0224	0.584
Min	7.50	0	0.950	7.19	0	0.925	6.97	0	0.903	7.47	0	0.889	7.10	0	0.622
Max	7.79	0.0000	0.950	87.75	0.0013	0.943	7.53	0.0004	0.905	7.95	0.0016	0.894	6.61	0.0169	0.652
Min	8.12	0	0.950	8.40	0	0.944	7.98	0	0.906	8.57	0	0.901	8.16	0	0.685
Max	8.62	0.0001	0.951	8.87	0.0004	0.947	8.58	0.0010	0.913	8.98	0.0003	0.902	8.67	0.0127	0.711
Min	9.05	0	0.952	9.53	0	0.950	9.13	0	0.919	9.58	0	0.903	9.21	0	0.739
Max	9.68	0.0004	0.957	9.80	0.0000	0.950	9.69	0.0011	0.927	9.83	0.0000	0.903	9.72	0.0094	0.761
Min	10.31	0	0.962	10.11	0	0.950	10.28	0	0.935	10.10	0	0.903	10.26	0	0.784

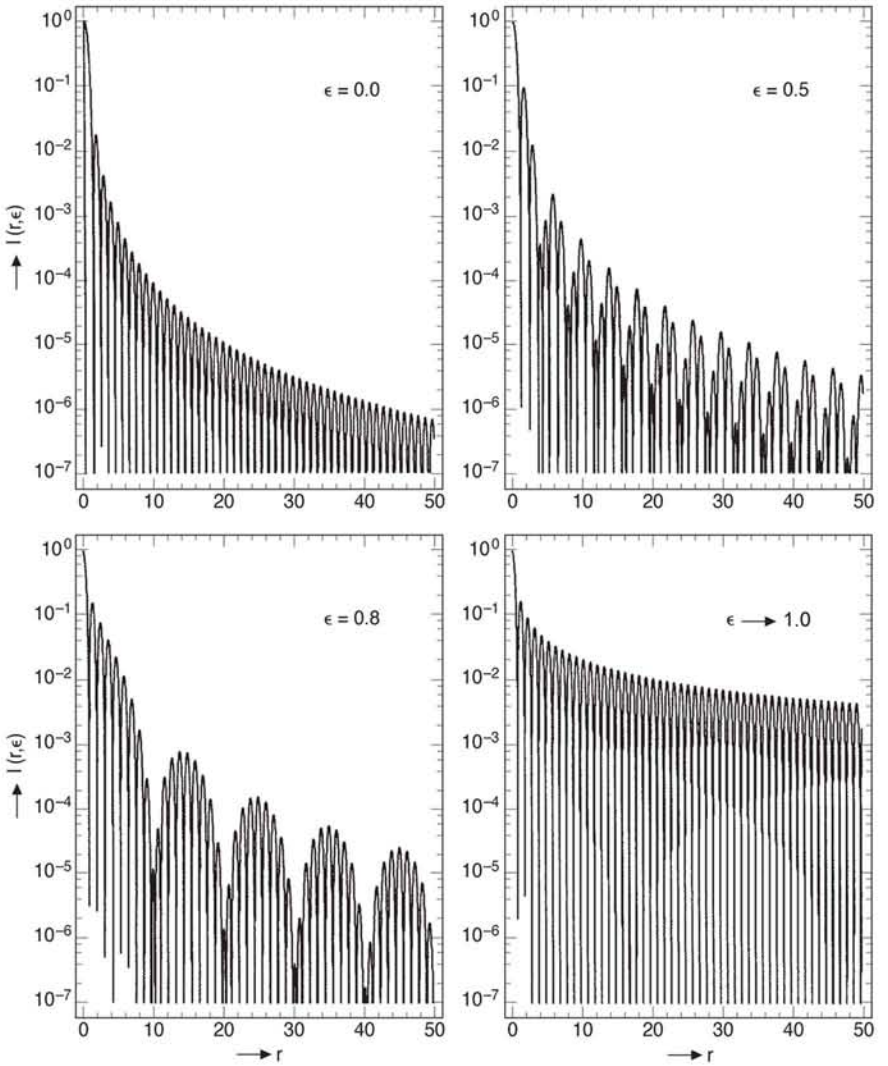


Figure 9-3a. Irradiance distribution for systems with *circular* $(\epsilon = 0)$ and *annular* $(\epsilon \neq 0)$ pupils. The case $\epsilon \to 1$ represents the limiting case of a totally obscured pupil. In practice, it approximates the PSF for a system with a very thin annular or a ring pupil.

When n is not an integer, then the distribution becomes complex. For example, for $\epsilon = 0.7$, $n = 6.67$, and the distribution has a double periodicity with the number of maxima in the two periods equal to 6 and 7 (two integers closest to n).

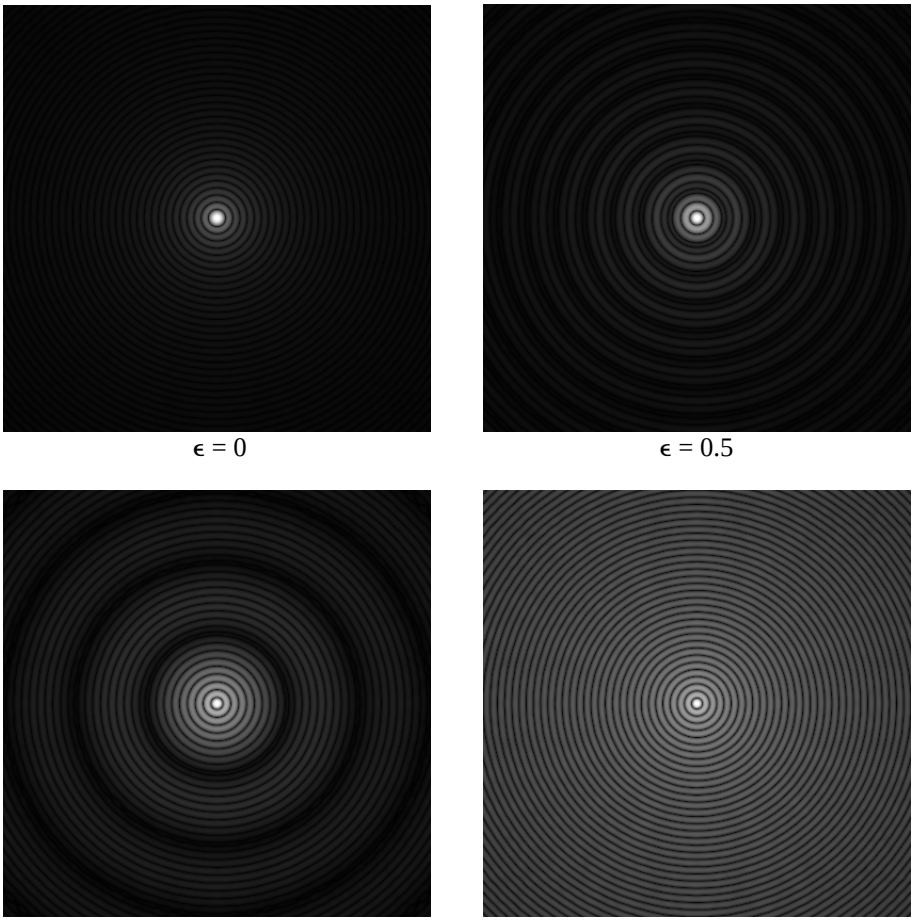


Figure 9-3b. 2D PSFs for systems with circular $(\epsilon = 0)$ and annular $(\epsilon \neq 0)$ pupils.

9.2.2 Aberration-Free OTF

As in the case of a system with a circular pupil, the aberration-free OTF of a system with an annular pupil for a spatial frequency v_i is also given by the fractional overlap area of two annuli whose centers are separated by a distance $\lambda R v_i$. Thus, it can be shown that the OTF is given by

$$\tau(v;\epsilon) = \frac{1}{1-\epsilon^2}\left[\tau(v) + \epsilon^2\tau(v/\epsilon) - \tau_{12}(v;\epsilon)\right], \quad 0 \le v \le 1 \quad , \tag{9-3}$$

where $\tau(v)$ given by Eq. (8-37) represents the OTF of the system if there were no obscuration, $v = v_i/(1/\lambda F)$ is a normalized radial spatial frequency, as in Eq. (8-38), and

$$\tau_{12}(v;\epsilon) = 2\epsilon^2, \quad \le v \le (1-\epsilon)/2 \tag{9-4a}$$

$$= (2/\pi)\left(\theta_1 + \epsilon^2\theta_2 - 2v\sin\theta_1\right), \quad (1-\epsilon)/2 \le v \le (1+\epsilon)/2 \tag{9-4b}$$

$$= 0, \text{ otherwise} \quad . \tag{9-4c}$$

In Eq. (9-4b), θ_1 and θ_2 are given by

$$\cos\theta_1 = \frac{4v^2 + 1 - \epsilon^2}{4v} \tag{9-4d}$$

and

$$\cos\theta_2 = \frac{4v^2 - 1 + \epsilon^2}{4\epsilon v} \quad , \tag{9-4e}$$

respectively. It should be evident that the *cutoff frequency* $v = 1$ or $v_i = 1/\lambda F$, which depends on the outer diameter of the annular pupil, is the same as that for a circular pupil. Moreover, we note from Eq. (9-3) that at least for spatial frequencies,

$$\frac{1+\epsilon}{2} < v < 1 \ , \ \tau(v;\epsilon) > \tau(v) \tag{9-5}$$

by a factor of $\left(1-\epsilon^2\right)^{-1}$. The overlap area in this frequency range is independent of ϵ, but the fractional area is larger owing to the smaller area of the obscured exit pupil. For a thin annular pupil, as $\epsilon \to 1$, a sharp peak near the cutoff frequency is obtained. The peak frequency represents the spatial frequency of fringes obtained in a 2D analog of a Young's double-slit aperture.

How $\tau(v;\epsilon)$ varies with v is shown in Figure 9-4 for various values of ϵ, including zero. We note that an annular pupil gives a higher OTF at high frequencies but a lower OTF at low frequencies, compared to the OTF for a corresponding circular $(\epsilon = 0)$ pupil. This is the frequency domain analog of smaller radius of the central bright ring and higher secondary maxima of the PSF for an annular pupil compared to those for a circular pupil.

As pointed out in Section 8.5.2, the slope of $\tau(v)$ at the origin is equal to $-4/\pi$. From Eq. (9-3) we find that the slope of the OTF for an annular pupil at the origin is given by

$$\tau'(0;\epsilon) = -4/\pi\left(1-\epsilon\right) \quad . \tag{9-6}$$

This slope does not change when aberrations are introduced into the system. We note that

$$\int_0^1 \tau(v;\epsilon)\, v dv = \left(1-\epsilon^2\right)/8 \quad . \tag{9-7}$$

9.2.3 Axial Irradiance

The axial irradiance of the image-forming beam for an aberration-free system with an annular pupil may be obtained in the same manner as for a system with a circular pupil. Thus, we let $r = 0$, $\Phi(\rho; z) = B_d(z)\rho^2$ [see Eq. (8-4)], and replace the lower limit of radial integration from 0 to ϵ in Eq. (8-1), thereby obtaining the result

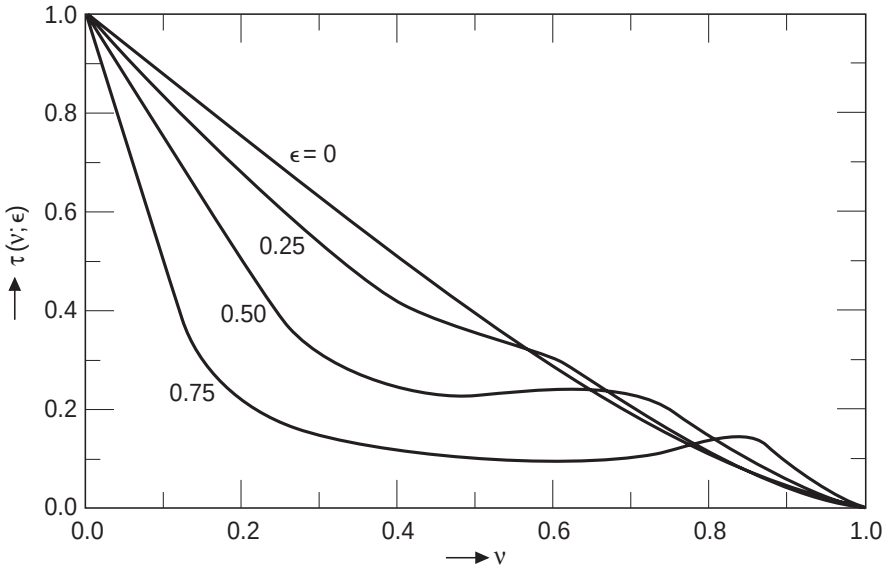


Figure 9-4. OTF of an aberration-free system with an annular pupil of obscuration ratio ϵ.

$$I(0; z; \epsilon) = \frac{PS_p}{\lambda^2 z^2}\left\{\frac{\sin\left[B_d\left(1-\epsilon^2\right)/2\right]}{B_d\left(1-\epsilon^2\right)/2}\right\}^2 \quad . \tag{9-8}$$

Equation (9-8) differs from the corresponding Eq. (8-7) for systems with circular pupils in that the quantity B_d in the latter has been replaced by $B_d\left(1-\epsilon^2\right)$. It represents the peak defocus phase aberration at the outer edge of the annular pupil relative to its value at the inner edge. Accordingly, the defocus tolerance or *depth of focus* for a given Strehl ratio for a system with an annular pupil is *larger* by a factor of $\left(1-\epsilon^2\right)^{-1}$ compared to its corresponding value if ϵ were zero. The axial irradiance is minimum and equal to zero at z values given by

$$R/z = 1 + 2n/N\left(1-\epsilon^2\right) \quad , \quad n = \pm 1, \pm 2, \ldots \quad , \tag{9-9}$$

where $N = a^2/\lambda R$ is the Fresnel number of the pupil if ϵ were zero. The maxima of axial irradiance, obtained by equating the derivative of Eq. (9-8) with respect to z equal to zero, are given by the solutions of

$$\tan\left[B_d\left(1-\epsilon^2\right)/2\right] = \left(R/z\right) B_d\left(1-\epsilon^2\right)/2 \quad , \quad z \neq R \quad . \tag{9-10}$$

Figure 9-5 shows how the axial irradiance of an annular beam with $\epsilon = 0.5$ varies for N = 1, 10, and 100. Comparing it with Figure 8-2, we note that the effect of the obscuration is to reduce the irradiance at the principal maximum but to increase it at the secondary maxima. Also, the maxima and minima occur at smaller z values for an annular pupil. As in the case of circular beams, the axial irradiance of annular beams also becomes symmetric about the focal point $z = R$ as N increases.

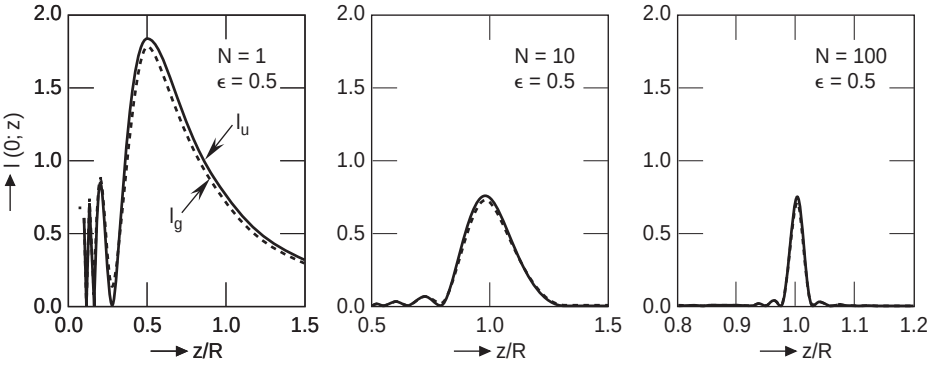


Figure 9-5. Axial irradiance of an annular beam with $\epsilon = 0.5$ focused at a distance R. The minima of irradiance occur at $z/R = 3/11, 3/19, 3/27$, when $N = 1$. The irradiance is in units of the focal-point irradiance of a corresponding circular beam with the same total power. Accordingly, the focal-point irradiance in this figure is $1 - \epsilon^2 = 0.75$. The axial irradiance becomes symmetric about the focal point as N increases. The dashed curves are for a Gaussian beam with $\gamma = 1$, as discussed in Section 9.3.3.

9.2.4 Strehl Ratio

The Strehl ratio is given by

$$S = \frac{1}{\pi\left(1-\epsilon^2\right)} \int_{\epsilon}^{1} \int_{0}^{2\pi} \exp\left[i\Phi(\rho, \theta; \epsilon)\right] \rho \, d\rho \, d\theta \quad . \tag{9-11}$$

For small aberrations, the Strehl ratio of an aberrated image is still given by Eqs. (8-13)–(8-15), except that the variance σ_Φ^2 of the aberration $\Phi(\rho, \theta; \epsilon)$ is across the annular region of the pupil. This in turn implies that the mean and the mean square values of the aberration are given by

$$< \Phi^n > = \frac{1}{\pi\left(1-\epsilon^2\right)} \int_{\epsilon}^{1} \int_{0}^{2\pi} \Phi^n(\rho, \theta; \epsilon) \rho \, d\rho \, d\theta \quad , \tag{9-12}$$

with n = 1 and 2, respectively.

The form of a primary aberration and its standard deviation are listed in Table 9-2. The balanced aberrations listed in the table represent balancing of an aberration with another to minimize its variance across the annular pupil. It should be evident that the diffraction focus for spherical aberration or coma lies at a larger distance from the Gaussian image point than that for a circular pupil. However, it is independent of ϵ for astigmatism.

Figure 9-6 shows how the standard deviation of an aberration varies with the obscuration ratio of the pupil. In Figures 9-6a and 9-6b, the amounts of defocus and tilt

Table 9-2. Primary aberrations and their standard deviations for a system with a uniformly illuminated annular pupil.

Aberration	$\Phi(\rho,\theta)$	σ_Φ
Spherical	$A_s\rho^4$	$\left(4-\epsilon^2-6\epsilon^4-\epsilon^6+4\epsilon^8\right)^{1/2}A_s/3\sqrt{5}$
Balanced spherical	$A_s\left[\rho^4-\left(1+\epsilon^2\right)\rho^2\right]$	$\frac{1}{6\sqrt{5}}\left(1-\epsilon^2\right)^2A_s$
Coma	$A_c\rho^3\cos\theta$	$\left(1+\epsilon^2+\epsilon^4+\epsilon^6\right)^{1/2}A_c/2\sqrt{2}$
Balanced coma	$A_c\left(\rho^3-\frac{2}{3}\frac{1+\epsilon^2+\epsilon^4}{1+\epsilon^2}\rho\right)\cos\theta$	$\frac{\left(1-\epsilon^2\right)\left(1+4\,\epsilon^2+\epsilon^4\right)^{1/2}}{6\sqrt{2}\left(1+\epsilon^2\right)^{1/2}}A_c$
Astigmatism	$A_a\rho^2\cos^2\theta$	$\left(1+\epsilon^2\right)^{1/2}A_a/4$
Balanced astigmatism	$A_a\rho^2\left(\cos^2\theta-1/2\right)$	$\frac{1}{2\sqrt{6}}\left(1+\epsilon^2+\epsilon^4\right)^{1/2}A_a$
Field curvature (defocus)	$A_d\rho^2$	$\left(1-\epsilon^2\right)A_d/2\sqrt{3}$
Distortion (tilt)	$A_t\rho\cos\theta$	$\left(1+\epsilon^2\right)^{1/2}A_t/2$

required to minimize the variance of spherical aberration and coma, respectively, are also shown. We observe from these figures that the standard deviation of spherical and balanced spherical aberrations and defocus decreases as ϵ increases. Correspondingly, the tolerance in terms of their aberration coefficients A_s and A_d, for a given Strehl ratio, increases. The standard deviation of coma, astigmatism, balanced astigmatism, and tilt increases as ϵ increases. The standard deviation of balanced coma first slightly increases, achieves its maximum value at ϵ = 0.29, and then decreases rapidly as ϵ increases. The factor by which the standard deviation of an aberration is reduced by balancing it with another aberration decreases in the case of spherical aberration and coma, but increases in the case of astigmatism, as ϵ increases.

Figures 9-7a and 9-7b show how the Strehl ratio of a primary aberration varies with its standard deviation for ϵ = 0.5 and 0.75. Approximate as well as exact results are shown in these figures.[2] The curves for a given aberration and for the corresponding balanced aberration can be distinguished from each other by their behavior for large σ_W values (near 0.25λ). For example, coma is shown by the evenly dashed curves; the higher dashed curve is for coma and the lower is for balanced coma.

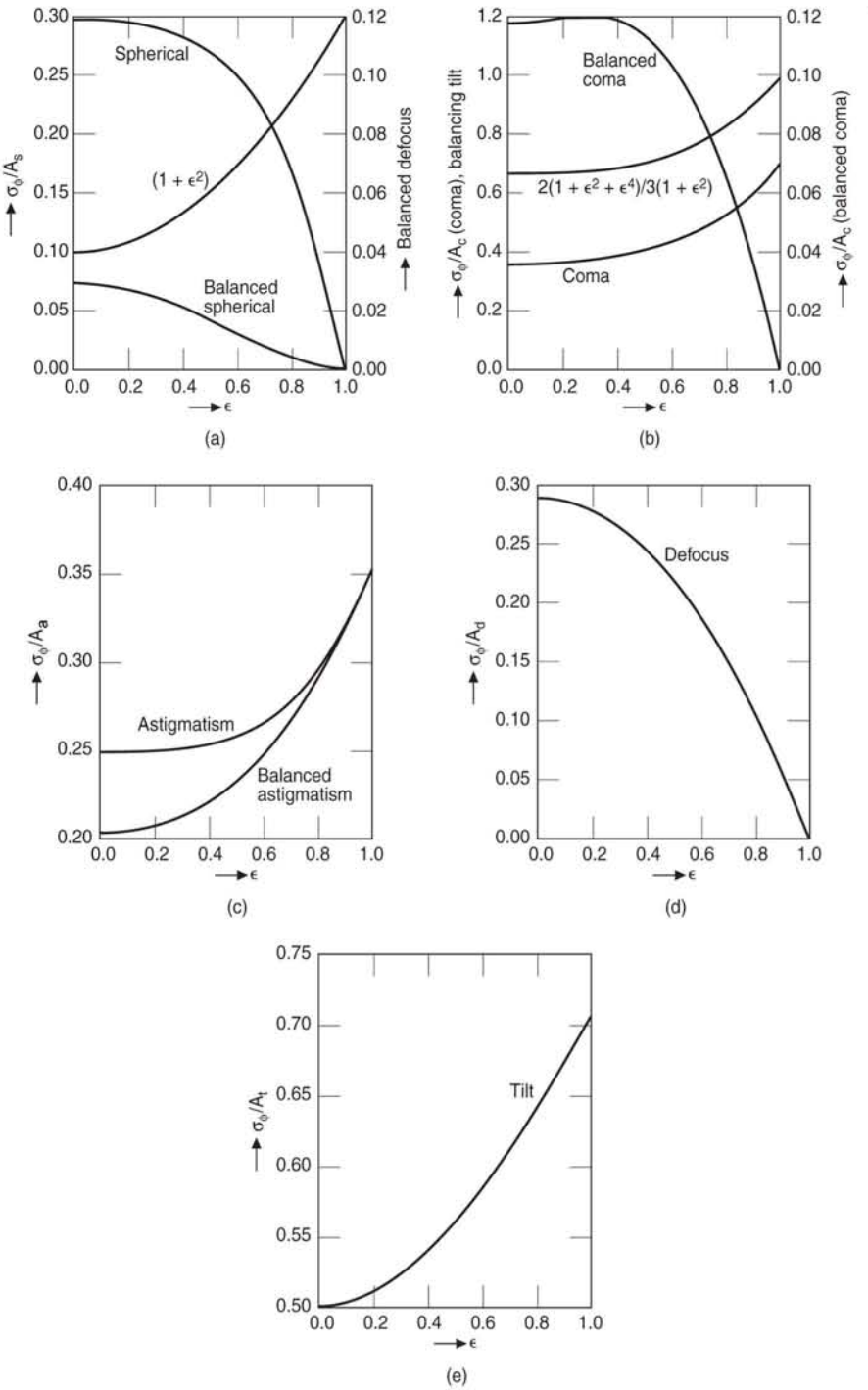


Figure 9-6. Variation of standard deviation of a primary and a balanced primary aberration with obscuration ratio ϵ. Variation of balancing defocus in the case of spherical aberration and tilt in the case of coma are also shown. (a) Spherical aberration, (b) coma, (c) astigmatism, (d) defocus, and (e) tilt.

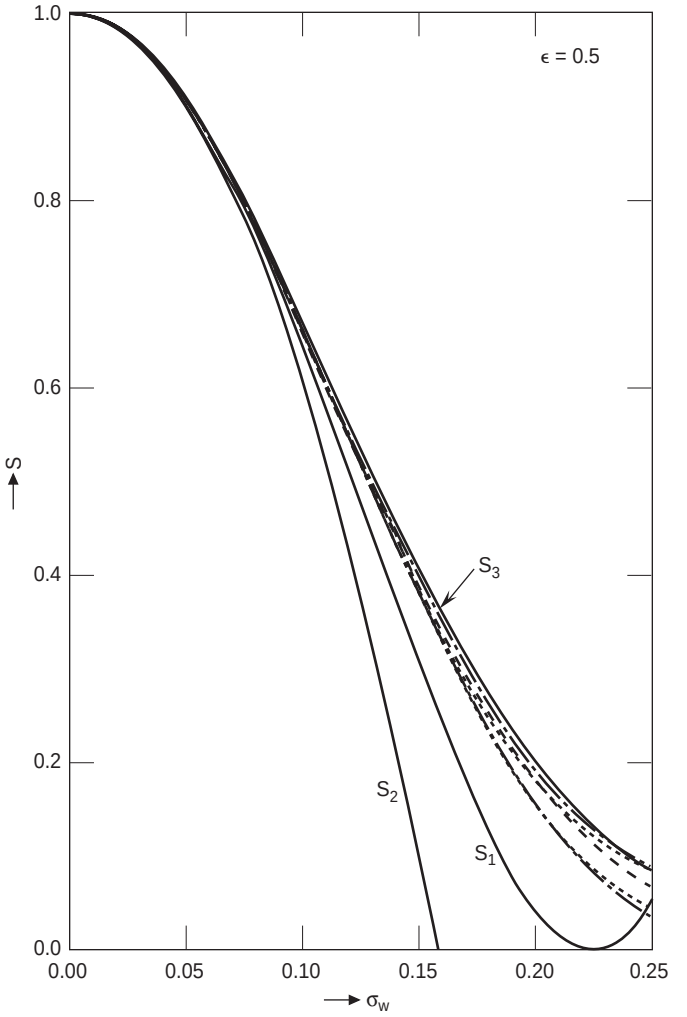


Figure 9-7a. Strehl ratio for annular pupils with $\epsilon = 0.5$ as a function of the standard deviation σ_w of an aberration in units of λ. The Strehl ratio for a given value of the standard deviation for classical coma is practically the same as that for balanced coma. For large values of σ_w, the Strehl ratio for classical astigmatism is larger than that for balanced astigmatism. Spherical...., Coma----, Astigmatism–.–.

As in the case of circular pupils, the expressions for S_1 and S_2 underestimate the true Strehl ratio. The expression for S_3 overestimates the true Strehl ratio for $\epsilon \geq 0.5$. It gives the Strehl ratio with an error of less than 10% for $S \geq 0.4$. For smaller obscurations, the error is less than 10% for $S \geq 0.3$. The percent error is defined as $100(1 - S_3/S)$.

Using S_1 to estimate the Strehl ratio, Figure 9-8 shows how the aberration coefficient A_i of a primary aberration for 10% error varies with the obscuration ratio.[3] It is evident that this coefficient increases with obscuration in the case of spherical, balanced spherical, and balanced coma, but decreases in the case of astigmatism, balanced

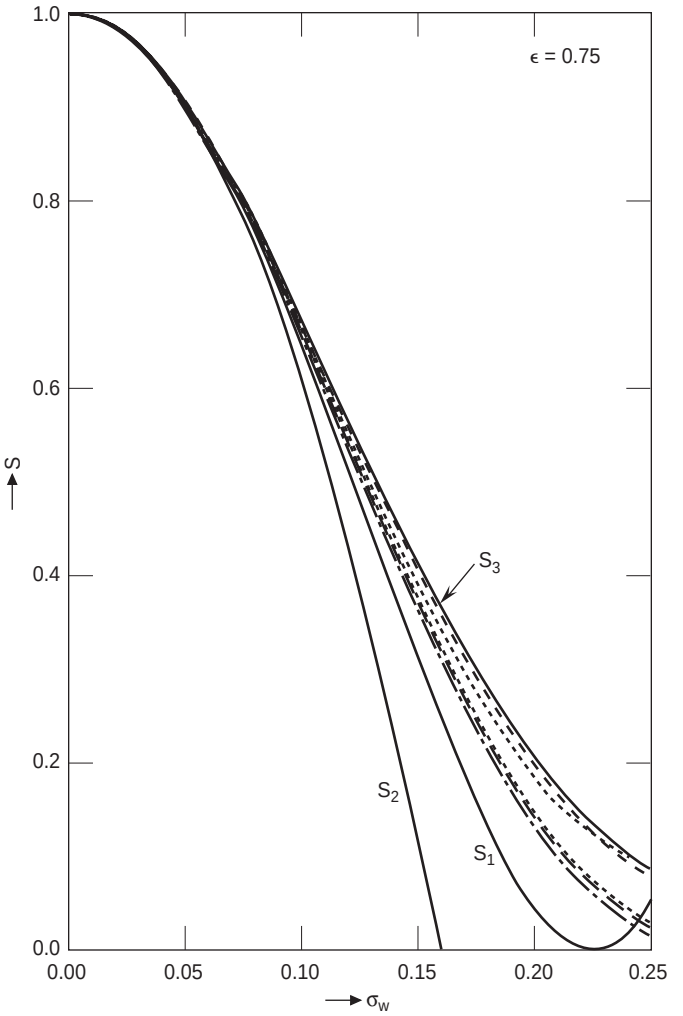


Figure 9-7b. Strehl ratio for annular pupils with $\epsilon = 0.75$ as a function of the standard deviation σ_w in units of λ. For large values of σ_w, the Strehl ratio for balanced coma is higher than that for coma. The opposite is true for astigmatism. Note that the curves for coma and astigmatism are practically identical. Spherical...., Coma----, Astigmatism–.–.

astigmatism, and coma. When the aberration coefficient A_i of an aberration is equal to a quarter wave, the variation of the corresponding Strehl ratio with ϵ is shown in Figure 9-9. It is evident that a Strehl ratio of 0.8 is obtained in very few cases. Comparing this figure with Figures 9-7a and 9-7b, we again conclude, as in the case of circular pupils, that it is advantageous to use the standard deviation of an aberration instead of the aberration coefficient to estimate the Strehl ratio. For example, a Strehl ratio of 0.8 is obtained for any aberration with a standard deviation of $\sigma_W = \lambda/14$. On the other hand, this value of Strehl ratio is obtained for different values of the aberration coefficient for different aberrations.

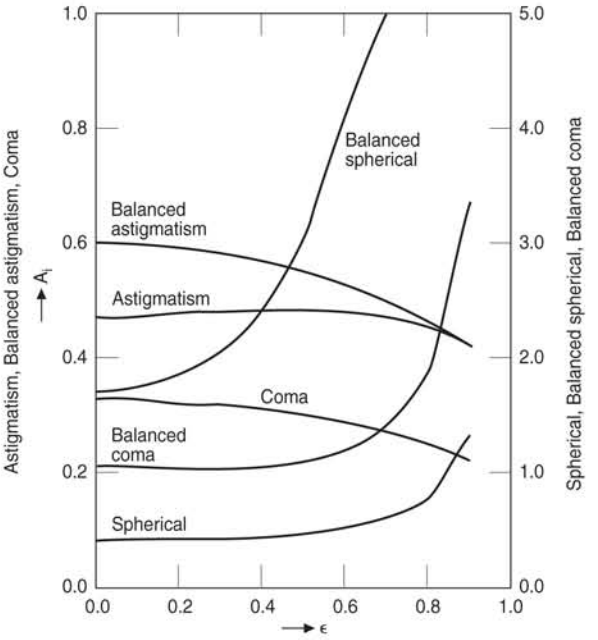


Figure 9-8. Variation of a primary aberration coefficient A_i (in units of λ) with ϵ for 10% error when S_1 is used to estimate the Strehl ratio.

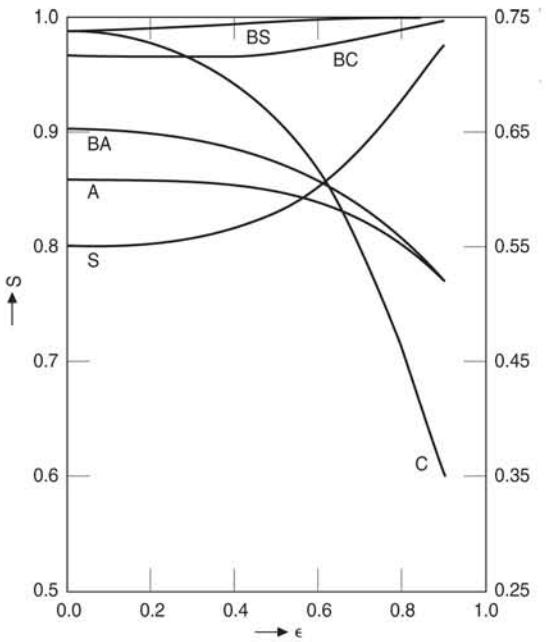


Figure 9-9. Strehl ratio for $A_i = \lambda/4$ as a function of ϵ. S, Spherical; BS, balanced spherical; C, coma; BC, balanced coma; A, astigmatism; and BA, balanced astigmatism. The right-hand side vertical scale is only for coma.

9.2.5 Balanced Aberrations and Zernike Annular Polynomials

The phase aberration function of a system with an annular pupil for a point object at a certain angle from its axis can be expanded in terms of a complete set of *Zernike annular polynomials* $Z_n^m(\rho,\theta;\epsilon)$ that are orthonormal over a unit annulus in the form[4]

$$\Phi(\rho,\theta;\epsilon) = \sum_{n=0}^{\infty}\sum_{m=0}^{n} c_{nm} Z_n^m(\rho,\theta;\epsilon) \quad , \quad \epsilon \le \rho \le 1 \quad , \quad 0 \le \theta \le 2\pi \quad , \tag{9-13}$$

where c_{nm} are the orthonormal expansion coefficients that depend on the field angle of the object, n and m are positive integers, $n - m \ge 0$ and even, and

$$Z_n^m(\rho,\theta;\epsilon) = \left[2(n+1)/(1+\delta_{m0})\right]^{1/2} R_n^m(\rho;\epsilon)\cos m\theta \quad . \tag{9-14}$$

The polynomials $Z_n^m(\rho,\theta;\epsilon)$ are orthonormal according to

$$\frac{1}{\pi(1-\epsilon^2)}\int_{\epsilon}^{1}\int_{0}^{2\pi} Z_n^m(\rho,\theta;\epsilon) Z_{n'}^{m'}(\rho,\theta;\epsilon)\,\rho\,d\rho\,d\theta = \delta_{nn'}\delta_{mm'} \quad . \tag{9-15}$$

The annular polynomials $Z_n^m(\rho,\theta;\epsilon)$ are similar to the circle polynomials $Z_n^m(\rho,\theta)$ discussed in Section 8.3.7, except that they are orthogonal over an annular pupil. They can be obtained from the corresponding circle polynomials by the Gram-Schmidt orthogonalization process.

The radial polynomials $R_n^m(\rho;\epsilon)$ obey the orthogonality relation

$$\int_{\epsilon}^{1} R_n^m(\rho;\epsilon) R_{n'}^m(\rho;\epsilon)\rho\,d\rho = \frac{1-\epsilon^2}{2(n+1)}\delta_{nn'} \quad . \tag{9-16}$$

A radial polynomial $R_n^m(\rho;\epsilon)$ is of degree n in ρ containing terms in ρ^n, ρ^{n-2},..., and ρ^m with coefficients that depend on ϵ. It is even or odd in ρ depending on whether n (or m) is even or odd. Moreover,

$$\begin{aligned} R_n^m(1;\epsilon) &= 1, \quad m = 0 \\ &\ne 1, \quad m \ne 0 \quad . \end{aligned} \tag{9-17}$$

It should be evident that the angular part $\cos m\theta$ of an annular polynomial $Z_n^m(\rho,\theta;\epsilon)$ is identically the same as that of a corresponding circle polynomial $Z_n^m(\rho,\theta)$.

The orthonormal annular coefficients are given by

$$c_{nm} = \frac{1}{\pi(1-\epsilon^2)}\int_{\epsilon}^{1}\int_{0}^{2\pi} \Phi(\rho,\theta;\epsilon)\, Z_n^m(\rho,\theta;\epsilon)\,\rho\,d\rho\,d\theta \quad , \tag{9-18}$$

as may be seen by substituting Eq. (9-13) and using the orthonormality Eq. (9-15) of the polynomials. Each annular coefficient, with the exception of c_{00}, represents the standard

deviation of the corresponding polynomial term. The variance of the aberration function is given by

$$\sigma_\Phi^2 = \left\langle \Phi^2(\rho, \theta; \epsilon) \right\rangle - \left\langle \Phi(\rho, \theta; \epsilon) \right\rangle^2$$

$$= \sum_{n=0}^{\infty} \sum_{m=0}^{n} c_{nm}^2 - c_{00}^2$$

$$= \sum_{n=1}^{\infty} \sum_{m=0}^{n} c_{nm}^2 \quad . \tag{9-19}$$

The Zernike annular radial polynomials for $n \leq 6$ are listed in Table 9-3. The number of Zernike (or orthogonal) aberration terms in the expansion of an aberration function through a certain order n is the same as in the case of circle polynomials. The balanced aberrations given in Table 9-2 can be identified with the annular polynomials. Thus the polynomials Z_2^2, Z_3^1, and Z_4^0 represent balanced astigmatism, coma, and spherical aberration. From the form of the annular polynomial $R_2^2(\rho; \epsilon)\cos 2\theta$, it is evident that the balancing defocus in the case of astigmatism is independent of the value of ϵ. The annular polynomials are unique in that they are the only polynomials that are orthogonal across an annular pupil and represent balanced aberrations for such a pupil, just as the circle polynomials discussed in Section 8.3.7 are unique for the circular pupils. Whereas the aberration function for a rotationally symmetric system consists of polynomials varying as $\cos m\theta$, an aberration function representing fabrication errors will generally consist of polynomials varying as $\sin m\theta$ as well. The single-index annular polynomials $Z_j(\rho, \theta; \epsilon)$ can be constructed in the same manner as the corresponding single index circle polynomials $Z_j(\rho, \theta)$ discussed in Section 8.3.7.

9.3 GAUSSIAN PUPILS

So far we have considered optical systems that have *uniform* amplitude across their exit pupils. Now we consider systems with exit pupils having nonuniform amplitude across them in the form of a *Gaussian.*[5,6] Such pupils are often referred to as *Gaussian pupils.* The Gaussian amplitude may, for example, be obtained by placing a filter with Gaussian transmission at the pupil. A system with a nonuniform amplitude across its pupil is called an *apodized system.* The motivation for apodizing a system is to reduce the values of the secondary maxima of its PSF relative to the value of the principal maximum. The discussion given here applies equally well to the propagation of *Gaussian laser beams.* For a Gaussian pupil transmitting the same total power as a circular pupil with uniform transmission, the central value of the PSF is smaller and the tolerance for an aberration is higher.

9.3.1 Aberration-Free PSF

The Gaussian amplitude may be written

$$A(\rho) = A_0 \exp\left(-\gamma \rho^2\right) \quad , \tag{9-20}$$

Table 9-3. Zernike annular radial polynomials.

n	m	$R_n^m(\rho;\epsilon)$
0	0	1
1	1	$\rho\Big/\left(1+\epsilon^2\right)^{1/2}$
2	0	$\left(2\rho^2-1-\epsilon^2\right)\Big/\left(1-\epsilon^2\right)$
2	2	$\rho^2\Big/\left(1+\epsilon^2+\epsilon^4\right)^{1/2}$
3	1	$\dfrac{3\left(1+\epsilon^2\right)\rho^3-2\left(1+\epsilon^2+\epsilon^4\right)\rho}{\left(1-\epsilon^2\right)\left[\left(1+\epsilon^2\right)\left(1+4\epsilon^2+\epsilon^4\right)\right]^{1/2}}$
3	3	$\rho^3\Big/\left(1+\epsilon^2+\epsilon^4+\epsilon^6\right)^{1/2}$
4	0	$\left[6\rho^4-6\left(1+\epsilon^2\right)\rho^2+1+4\epsilon^2+\epsilon^4\right]\Big/\left(1-\epsilon^2\right)^2$
4	2	$\dfrac{4\rho^4-3\left[\left(1-\epsilon^8\right)\Big/\left(1-\epsilon^6\right)\right]\rho^2}{\left\{\left(1-\epsilon^2\right)^{1}\left[16\left(1-\epsilon^{10}\right)-15\left(1-\epsilon^8\right)^2\Big/\left(1-\epsilon^6\right)\right]^{1/2}\right\}}$
4	4	$\rho^4\Big/\left(1+\epsilon^2+\epsilon^4+\epsilon^6+\epsilon^8\right)^{1/2}$
5	1	$\dfrac{10\left(1+4\epsilon^2+\epsilon^4\right)\rho^5-12\left(1+4\epsilon^2+4\epsilon^4+\epsilon^6\right)\rho^3+3\left(1+4\epsilon^2+10\epsilon^4+4\epsilon^6+\epsilon^8\right)\rho}{\left(1-\epsilon^2\right)^2\left[\left(1+4\epsilon^2+\epsilon^4\right)\left(1+9\epsilon^2+9\epsilon^4+\epsilon^6\right)\right]^{1/2}}$
5	3	$\dfrac{5\rho^5-4\left[\left(1-\epsilon^{10}\right)\Big/\left(1-\epsilon^8\right)\right]\rho^3}{\left\{\left(1-\epsilon^2\right)^{1}\left[25\left(1-\epsilon^{12}\right)-24\left(1-\epsilon^{10}\right)^2\Big/\left(1-\epsilon^8\right)\right]\right\}^{1/2}}$
5	5	$\rho^5\Big/\left(1+\epsilon^2+\epsilon^4+\epsilon^6+\epsilon^8+\epsilon^{10}\right)^{1/2}$
6	0	$\left[20\rho^6-30\left(1+\epsilon^2\right)\rho^4+12\left(1+3\epsilon^2+\epsilon^4\right)\rho^2-\left(1+9\epsilon^2+9\epsilon^4+\epsilon^6\right)\right]\Big/\left(1-\epsilon^2\right)^3$
6	2	$\dfrac{15\left(1+4\epsilon^2+10\epsilon^4+4\epsilon^6+\epsilon^8\right)\rho^6-20\left(1+4\epsilon^2+10\epsilon^4+10\epsilon^6+4\epsilon^8+\epsilon^{10}\right)\rho^4+6\left(1+4\epsilon^2+10\epsilon^4+20\epsilon^6+10\epsilon^8+4\epsilon^{10}+\epsilon^{12}\right)\rho^2}{\left(1+\epsilon^2\right)^2\left[\left(1+4\epsilon^2+10\epsilon^4+4\epsilon^6+\epsilon^8\right)\left(1+9\epsilon^2+45\epsilon^4+65\epsilon^6+45\epsilon^8+9\epsilon^{10}+\epsilon^{12}\right)\right]^{1/2}}$
6	4	$\dfrac{6\rho^6-5\left[\left(1-\epsilon^{12}\right)\Big/\left(1-\epsilon^{10}\right)\right]\rho^4}{\left\{\left(1-\epsilon^2\right)^{1}\left[36\left(1-\epsilon^{14}\right)-35\left(1-\epsilon^{12}\right)^2\Big/\left(1-\epsilon^{10}\right)\right]\right\}^{1/2}}$
6	6	$\rho^6\Big/\left(1+\epsilon^2+\epsilon^4+\epsilon^6+\epsilon^8+\epsilon^{10}+\epsilon^{12}\right)^{1/2}$

where A_0 is a constant and γ is a parameter that defines truncation of the Gaussian by the pupil. If we let ω be the radial distance at which the amplitude drops to $1/e$ of its value at the center, then $\gamma = (a/\omega)^2$ where a is the radius of the exit pupil. We refer to ω as the *Gaussian radius.* In the limit $\gamma \to 0$, we obtain a uniformly illuminated pupil. The total power transmitted by the pupil is obtained by integrating $A^2(\rho)$ across the pupil.

The PSF for a Gaussian pupil or the irradiance distribution of a focused Gaussian beam may be obtained from Eq. (8-1) provided the Gaussian amplitude is inserted under the integral in this equation. The irradiance and encircled power distributions thus obtained for an aberration-free system are given by[6]

$$I(r;\gamma) = 4\left[\int_0^1 \sqrt{I(\rho)}\, J_0(\pi r\rho)\, \rho\, d\rho\right]^2 . \tag{9-21}$$

and

$$P(r_c;\gamma) = (\pi^2/2)\int_0^{r_c} I(r;\gamma)\, r dr , \tag{9-22}$$

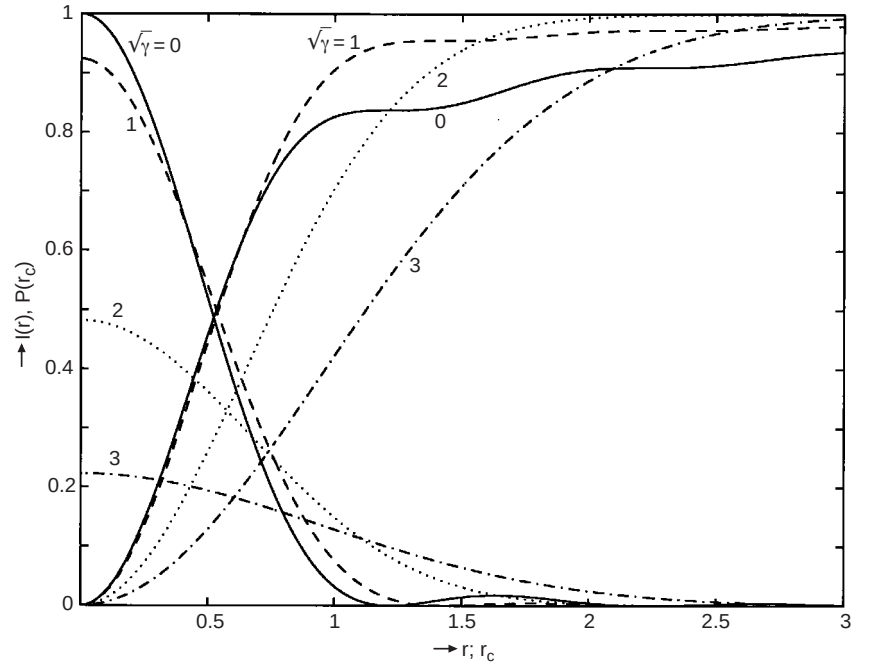
respectively, where

$$I(\rho) = 2\gamma \exp(-2\gamma\rho^2)/\left[1 - \exp(-2\gamma)\right] \tag{9-23}$$

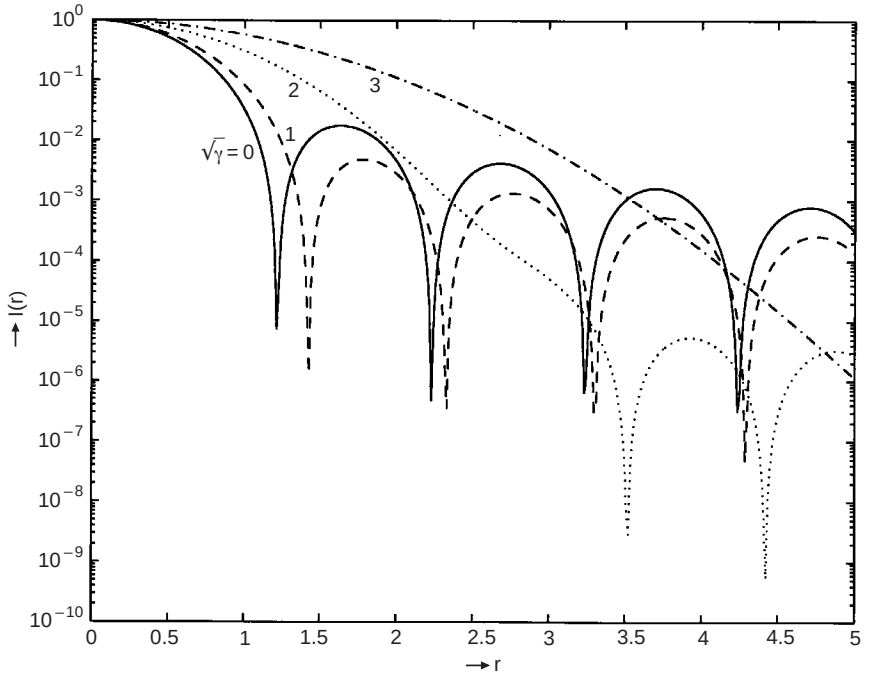
is the irradiance in units of power in the pupil per unit area.

Figure 9-10a shows the irradiance and encircled-power distributions for various values of γ, including $\gamma = 0$. For clarity, the irradiance distributions are also plotted on a logarithmic scale in Figure 9-10b to highlight the differences between the secondary maxima of uniform and Gaussian beams. It is evident that the *Gaussian illumination broadens the central disc but reduces the power in the secondary rings.* As γ increases, the value at the center [obtained from Eq. (9-21) by letting $r = 0$] and the values of secondary maxima decrease. For large values of γ, the diffracted beam is also Gaussian, as discussed in Section 9.3.6.

The positions of maxima and minima and the corresponding irradiance and encircled-power values are given in Table 9-4 for $\gamma = 1$. Comparing them with those in Table 8-1 for a uniform pupil, it is evident that the corresponding maxima and minima for a Gaussian beam are located at larger values of r than those for a uniform beam. Moreover, whereas the principal maximum for a Gaussian beam is only slightly lower (0.924 compared with 1), the secondary maxima are lower by a factor > 3 compared with the corresponding maxima for a uniform beam. Note that $I_u > I_g$ for $r < 0.42$. For larger values of r, $I_g > I_u$, except in the secondary rings, where again $I_u > I_g$. The encircled power $P_u \gtrless P_g$ for $r_c \lesseqgtr 0.63$. Of course, as $r_c \to \infty$, $P_u \to P_g \to 1$.



(a)



(b)

Figure 9-10. (a) PSF and encircled power for a Gaussian pupil with $\sqrt{\gamma}$ = 0, 1, 2 and 3, where the irradiance is in units of PS_p/λ^2R^2, encircled power is in units of *P*, and *r* and r_c are in units of λF. (b) Irradiance distribution normalized to unity at the center shown on a log scale to highlight the differences between the secondary maxima of uniform and Gaussian beams.

Table 9-4. Maxima and minima of image-plane irradiance distribution and corresponding encircled powers for a Gaussian pupil with $\gamma = 1$ compared with those for a uniform $(\gamma = 0)$ pupil, which are given in parentheses.

Max/Min	r, r_c	$I(r)$	$P(r_c)$
Max	0	0.924	0
	(0)	(1)	(0)
Min	1.43	0	0.955
	(1.22)	(0)	(0.838)
Max	1.79	0.0044	0.962
	(1.64)	(0.0175)	(0.867)
Min	2.33	0	0.973
	(2.23)	(0)	(0.910)
Max	2.76	0.0012	0.976
	(2.68)	(0.0042)	(0.922)
Min	3.30	0	0.981
	(3.24)	(0)	(0.938)
Max	3.76	0.0005	0.983
	(3.70)	(0.0016)	(0.944)
Min	4.29	0	0.985
	(4.24)	(0)	(0.952)
Max	4.75	0.0002	0.986
	(4.71)	(0.0008)	(0.957)

9.3.2 Aberration-Free OTF

The OTF for an aberration-free Gaussian pupil is given by

$$\tau(v;\gamma) = \frac{8\gamma \exp(-2\gamma v^2)}{\pi[1-\exp(-2\gamma)]} \int_0^{\sqrt{1-v^2}} dq \int_0^{\sqrt{1-q^2}-v} \exp\left[-2\gamma\left(p^2+q^2\right)\right] dp \,, \; 0 \le v \le 1 \quad , \quad (9\text{-}24)$$

where the coordinates of a pupil point are normalized by the pupil radius a and the integration is over a quadrant of the overlap region of two pupils whose centers are separated by a distance v along the p axis.

For large values of γ (e.g., $\gamma \ge 4$), the contribution to the integral in Eq. (9-24) is negligible unless $v = 0$, in which case it represents the Gaussian-weighted area of a quadrant of the pupil, and the equation reduces to

$$\tau(v;\gamma) = \exp(-2\gamma v^2) \quad , \quad 0 \le v \le 1 \quad . \quad (9\text{-}25)$$

Figure 9-11 shows how the OTF varies with v for several values of γ. We note that

compared to a uniform pupil (i.e., for $\gamma = 0$), the OTF of a Gaussian pupil is higher for low spatial frequencies, and lower for the high. Moreover, as γ increases, the bandwidth of low frequencies for which the OTF is higher decreases, and the OTF at high frequencies becomes increasingly smaller. This is due to the fact that the Gaussian weighting across the overlap region of two pupils whose centers are separated by small values of v is higher than that for large values of v. If we consider an apodization such that the amplitude increases from the center toward the edge of the pupil, then the OTF is lower for low frequencies and higher for high frequencies. Thus, unlike aberrations, which reduce the MTF of a system at all frequencies within its passband, the amplitude variations can increase or decrease the MTF at any of those frequencies.

9.3.3 Axial Irradiance

The irradiance distribution in a defocused image plane at a distance z from the plane of the exit pupil is given by[7]

$$I(r; z; \gamma) = \left(\frac{2R}{z}\right)^2 \left| \int_0^1 \sqrt{I(\rho)}\, \exp\left(iA_d \rho^2\right) J_0(\pi \rho r) \rho \, d\rho \right|^2 . \tag{9-26}$$

If we let $r = 0$ in Eq. (9-26), we obtain the axial irradiance of the beam

$$I(0; z; \gamma) = \left(\frac{R}{z}\right)^2 \left(\frac{2\gamma}{B_d^2 + \gamma^2}\right) \frac{1}{\sinh \gamma} (\cosh \gamma - \cos B_d) . \tag{9-27}$$

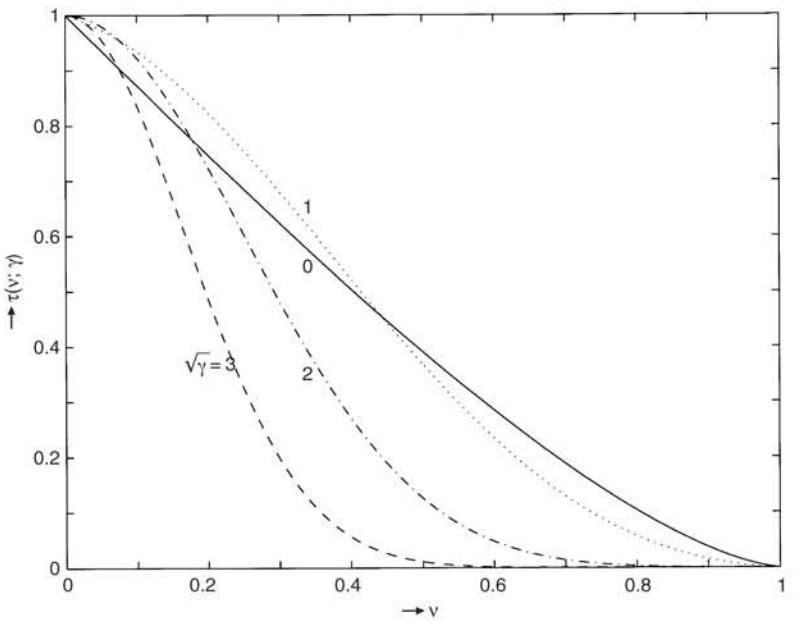


Figure 9-11. The OTF of a Gaussian pupil. A uniform pupil corresponds to $\gamma = 0$, and a large value of γ represents a weakly truncated pupil.

It goes through a series of maxima and minima as a function of z because of the $\cos B_d$ term.

Figure 8-2 shows how the axial irradiance of a focused Gaussian beam with $\gamma = 1$ differs from that of a focused uniform beam when the Fresnel number $N = 1$, 10, and 100. We note that the principal maximum is higher for the uniform beam compared with that for the Gaussian beam. However, the secondary maxima are higher for the Gaussian beam. Moreover, whereas the axial minima for the uniform beam have a value of zero, the minima for the Gaussian beam have non-zero values. We note that the curves become symmetric about the focal point $z = R$ as N increases. It should be noted that even though the principal maximum of axial irradiance does not lie at the focus, unless N is very large, the maximum central irradiance on a target at a given distance from the pupil is obtained when the beam is focused on it. (Similarly, for a weakly truncated Gaussian pupil discussed later, minimum Gaussian radius is obtained on a target when the beam is focused on it, even though a smaller radius occurs at a distance $z < R$ when N is small.)

9.3.4 Strehl Ratio

The Strehl ratio (representing the ratio of the central irradiances with and without aberration) is given by[8]

$$S = \left| \int_0^1 \int_0^{2\pi} A(\rho) \exp\left[i\Phi(\rho,\theta)\right] \rho \, d\rho \, d\theta \right|^2 \Bigg/ \left[\int_0^1 \int_0^{2\pi} A(\rho) \, \rho \, d\rho \, d\theta \right]^2$$

$$= \left\{ \frac{\gamma}{\pi\left[1 - \exp(-\gamma)\right]} \right\}^2 \left| \int_0^1 \int_0^{2\pi} \exp\left(-\gamma\rho^2\right) \exp\left[i\Phi(\rho,\theta)\right] \rho \, d\rho \, d\theta \right|^2 . \qquad (9\text{-}28)$$

Its approximate value can be obtained from Eqs. (8-13)–(8-15), where the variance of the aberration is now across the amplitude-weighted pupil. Thus, for a circular pupil, the mean and the mean square values of the aberration are given by

$$\langle \Phi^n \rangle = \int_0^1 \int_0^{2\pi} A(\rho) \left[\Phi(\rho,\theta)\right]^n \rho \, d\rho \, d\theta \Bigg/ \int_0^1 \int_0^{2\pi} A(\rho) \, \rho \, d\rho \, d\theta \qquad (9\text{-}29)$$

with $n = 1$ and 2, respectively. Following the same procedure as for a uniformly illuminated circular pupil, we can obtain the balanced primary aberrations and their standard deviations. Table 9-5 gives the aberrations and their standard deviations for $\gamma = 1$, i.e., when $a = \omega$. Comparing these results with those given in Table 9-2 for $\epsilon = 0$, it is evident that the standard deviation of an aberration for a Gaussian pupil is somewhat smaller than the corresponding value for a uniform pupil. Accordingly, for a given small amount of aberration A_i, the Strehl ratio for a Gaussian pupil is somewhat higher than that for a uniform pupil. Thus the depth of focus increases as γ increases, or the beam

Table 9-5. Primary aberrations and their standard deviations for optical systems with Gaussian circular pupils with $\gamma = 1$.

Aberration	$\Phi(\rho,\theta)$	σ_Φ
Spherical	$A_s\rho^4$	$\frac{A_s}{3.67}$
Balanced spherical	$A_s(\rho^4 - 0.933\rho^2)$	$\frac{A_s}{13.71}$
Coma	$A_c\rho^3\cos\theta$	$\frac{A_c}{3.33}$
Balanced coma	$A_c(\rho^3 - 0.608\rho)\cos\theta$	$\frac{A_c}{8.80}$
Astigmatism	$A_a\rho^2\cos^2\theta$	$\frac{A_a}{4.40}$
Balanced astigmatism	$A_a\rho^2(\cos^2\theta - 1/2)$	$\frac{A_a}{5.61}$
Defocus	$B_d\rho^2$	$\frac{B_d}{3.55}$
Tilt	$B_t\rho\cos\theta$	$\frac{B_t}{2.19}$

becomes narrower. Similarly, for a given Strehl ratio, the aberration tolerance for a Gaussian pupil is somewhat higher than that for a uniform pupil. Moreover, the balancing defocus in the case of spherical aberration and the balancing tilt in the case of coma are somewhat smaller for a Gaussian pupil, compared to their corresponding values for a uniform pupil; i.e., the *diffraction focus* for these aberrations in the case of a Gaussian pupil is slightly different from the corresponding focus for a uniform pupil. We also note that, although aberration balancing in the case of a uniform pupil reduces the standard deviation of spherical aberration and coma by factors of 4 and 3, respectively, the reduction in the case of astigmatism is only a factor of 1.22. For a Gaussian pupil, the trend is similar but the reduction factors are smaller for spherical aberration and coma, and larger for astigmatism. They are 3.74, 2.64, and 1.27, corresponding to spherical aberration, coma, and astigmatism, respectively. (In Reference 5, the factor for astigmatism is incorrectly stated as 1.16 in the text and 1.66 in Table 5.)

9.3.5 Balanced Aberrations and Zernike-Gauss Circle Polynomials

The *Zernike-Gauss polynomials* $Z_n^m(\rho,\theta;\gamma)$ orthonormal over a circular Gaussian pupil and representing balanced aberrations for such pupils can be obtained from the polynomials $Z_n^m(\rho,\theta)$ for uniform illumination by the Gram-Schmidt orthogonalization process. The phase aberration function of a system with a circular exit pupil can be expanded in terms of these polynomials in the form[5,6]

$$\Phi(\rho,\theta;\gamma) = \sum_{n=0}^{\infty}\sum_{m=0}^{n} c_{nm}Z_n^m(\rho,\theta;\gamma) \quad , \quad 0 \le \rho \le 1 \quad , \quad 0 \le \theta \le 2\pi \quad , \tag{9-30}$$

where c_{nm} are the orthonormal expansion coefficients, n and m are positive integers including zero, $n-m \geq 0$ and even, and

$$Z_n^m(\rho,\theta;\gamma) \;=\; \left[2(n+1)/(1+\delta_{m0})\right]^{1/2} R_n^m(\rho;\gamma)\cos m\theta \quad . \tag{9-31}$$

The polynomials are orthonormal according to

$$\int_0^1 \int_0^{2\pi} Z_n^m(\rho,\theta;\gamma) Z_{n'}^{m'}(\rho,\theta;\gamma) A(\rho)\rho\, d\rho\, d\theta \Big/ 2\pi \int_0^1 A(\rho)\rho\, d\rho \;=\; \delta_{nn'}\delta_{mm'} \quad . \tag{9-32}$$

The radial polynomials obey the orthogonality relation

$$\int_0^1 R_n^m(\rho;\gamma)\, R_{n'}^m(\rho;\gamma)\, A(\rho)\,\rho\, d\rho \Big/ \int_0^1 A(\rho)\,\rho\, d\rho \;=\; \frac{1}{n+1}\delta_{nn'} \quad . \tag{9-33}$$

The radial polynomial $R_n^m(\rho;\gamma)$ is a polynomial of degree n in ρ containing terms in ρ^n, ρ^{n-2}, ..., and ρ^m, whose coefficients depend on the Gaussian amplitude through γ; i.e., it has the form

$$R_n^m(\rho;\gamma) \;=\; a_n^m \rho^n + b_n^m \rho^{n-2} + \ldots + d_n^m \rho^m \quad , \tag{9-34}$$

where the coefficients a_n^m, etc., depend on γ.

The radial polynomials corresponding to balanced primary aberrations are listed in Table 9-6. As in the case of annular polynomials, the angular part $\cos m\theta$ of a Zernike-Gauss polynomial $Z_n^m(\rho,\theta;\gamma)$ is identically the same as that of a corresponding circle polynomial $Z_n^m(\rho,\theta)$. From the form of the Zernike-Gauss polynomial $R_2^2(\rho;\gamma)\cos 2\theta$, it is evident that the balancing defocus in the case of astigmatism is independent of the value of γ. The Zernike-Gauss polynomials are unique in the sense that they are the only polynomials that are orthogonal across a Gaussian amplitude-weighted pupil and represent balanced aberrations for such a pupil.

The number of Zernike (or orthogonal) aberration terms in the expansion of an aberration function through a certain order n is the same as in the case of circle or annular polynomials. The Zernike-Gauss expansion coefficients are given by

$$c_{nm} = \int_0^1 \int_0^{2\pi} \Phi(\rho,\theta;\gamma) Z_n^m(\rho,\theta;\gamma) A(\rho)\rho\, d\rho\, d\theta \Big/ 2\pi \int_0^1 A(\rho)\rho\, d\rho \quad , \tag{9-35}$$

as may be seen by substituting Eq. (9-30) and utilizing the orthonormality Eq. (9-32) of the polynomials. Each expansion coefficient, with the exception of c_{00}, represents the standard deviation of the corresponding polynomial term. The variance of the aberration function is given by

$$\sigma_\Phi^2 \;=\; \left\langle \Phi^2(\rho,\theta;\gamma)\right\rangle - \left\langle \Phi(\rho,\theta;\gamma)\right\rangle^2 \;=\; \sum_{n=1}^{\infty} \sum_{m=0}^{n} c_{nm}^2 \quad , \tag{9-36}$$

where the coefficients a_n^m, etc., depend on γ.

Table 9-6. Zernike-Gauss radial polynomials $R_n^m(\rho;\gamma)$ representing balanced primary aberrations for uniform ($\gamma = 0$), Gaussian ($\gamma = 1$), and weakly truncated Gaussian beams.

Aberration	Radial Polynomial	Gaussian*	Gaussian γ 1	Uniform γ 0	Weakly Truncated Gaussian
Piston	R_0^0	1	1	1	1
Distortion (tilt)	R_1^1	$a_1^1\rho$	1.09367ρ	ρ	$\sqrt{\gamma/2}\rho$
Field curvature (defocus)	R_2^0	$a_2^0\rho^2 + b_2^0$	$2.04989\rho^2 - 0.85690$	$2\rho^2 - 1$	$(\gamma\rho^2 - 1)/\sqrt{3}$
Astigmatism	R_2^2	$a_2^2\rho^2$	$1.14541\rho^2$	ρ^2	$(\gamma/\sqrt{6})\rho^2$
Coma	R_3^1	$a_3^1\rho^3 + b_3^1\rho$	$3.11213\rho^3 - 1.89152\rho$	$3\rho^3 - 2\rho$	$\sqrt{\gamma/2}\left(\frac{\gamma}{2}\rho^3 - \rho\right)$
Spherical aberration	R_4^0	$a_4^0\rho^4 + b_4^0\rho^2 + c_4^0$	$6.12902\rho^4 - 5.71948\rho^2 + 0.83368$	$6\rho^4 - 6\rho^2 + 1$	$(\gamma^2\rho^4 - 4\gamma\rho^2 + 2)/2\sqrt{5}$

*$a_1^1 = (2p_2)^{-1/2}$, $a_2^0 = [3(p_4 - p_2^2)]^{-1/2}$, $b_2^0 = -p_2 a_2^0$, $a_2^2 = (3p_4)^{-1/2}$, $a_3^1 = \frac{1}{2}(p_6 - p_4^2/p_2)$, $b_3^1 = -(p_4/p_2)a_3^1$,

$a_4^0 = \left\{5[p_8 - 2K_1 p_6 + (K_1^2 + 2K_2)p_4 - 2K_1 K_2 p_2 + K_2^2]\right\}^{-1/2}$, $b_4^0 = -K_1 a_4^0, c_4^0 = K_2 a_4^0$,

$p_s = <\rho^s> = (1 - \exp\gamma)^{-1} + (s/2\gamma)p_{s-2}$, s is an even integer,

$p_0 = 1, K_1 = (p_6 - p_2 p_4)/(p_4 - p_2^2), K_2 = (p_2 p_6 - p_4^2)/(p_4 - p_2^2)$.

9.3.6 Weakly Truncated Pupils

For a weakly truncated Gaussian pupil, i.e., for large values of γ, the upper limit on the radial variable in Eq. (9-20) and any associated equations may be changed from 1 to ∞ with negligible error. Numerical calculations show that for $\gamma \geq 9$ (or $a \geq 3\omega$), the difference between the exact PSF and the approximate result thus obtained may be neglected.[5] Moreover, in the limit of an untruncated beam, the ring structure of the diffraction pattern disappears and an aberration-free Gaussian beam propagates as a Gaussian. The beam radius and the irradiance distribution in a plane at a distance z from a plane where its beam radius is ω_z are given by

$$\omega_z^2 = (\lambda z/\pi\omega)^2 + \omega^2(1 - z/R)^2 \tag{9-37}$$

and

$$I(r;z) = \left(2P/\pi\omega_z^2\right)\exp\left(-2r^2/\omega_z^2\right), \tag{9-38}$$

respectively. In Eq. (9-38), r is the radial distance of a point in the observation plane from the axis of the beam without any normalization. Since the PSF is Gaussian, its Fourier transform representing the OTF is also Gaussian, as indicated earlier in Eq. (9-25). For a weakly truncated beam, since the power in the pupil is concentrated in a small region near its center, the effect of the aberration in its outer region is negligible. Accordingly, the aberration tolerance in terms of the peak value of a primary aberration at the edge ($\rho = 1$) of the pupil is not very meaningful. It is more appropriate, for example, to consider the tolerance in terms of the peak value at the Gaussian radius. If we define

$$\rho' = \sqrt{\gamma}\,\rho \quad , \tag{9-39}$$

then, $\rho' = 1$ corresponds to the Gaussian radius. Correspondingly, we define the aberration coefficients

$$A_s' = A_s/\gamma^2,\ A_c' = A_c/\gamma^{3/2},\ A_a' = A_a/\gamma,\ B_d' = B_d/\gamma,\ B_t' = B_t/\sqrt{\gamma} \quad , \tag{9-40}$$

which represent the peak values of aberrations at the Gaussian radius.

Table 9-7 lists the aberrations in terms of the radial variable ρ' and the aberration coefficients A_i'. The standard deviations of these aberrations across the Gaussian pupil are also given in this table. We note that the balancing of an aberration reduces its standard deviation by a factor of $\sqrt{5}$, $\sqrt{3}$, and $\sqrt{2}$, in the case of spherical aberration, coma, and astigmatism, respectively. The amount of a balancing aberration decreases as γ increases in the case of spherical aberration and coma but does not change in the case of astigmatism. For example, in the case of spherical aberration, the amount of balancing defocus for a weakly truncated Gaussian pupil is $4/\gamma$ times the corresponding amount for a uniform pupil. Similarly, in the case of coma, the balancing tilt for a weakly truncated Gaussian pupil is $3/\gamma$ times the corresponding amount for a uniform pupil. Aberration tolerances in terms of the aberration coefficients A_i' for a Strehl ratio of 0.8 are given in Table 9-7. The tolerances in terms of the coefficients A_i may be obtained by use of Eq. (9-14).

Table 9-7. Primary aberrations and their standard deviations for optical systems with weakly $\left(\sqrt{\boldsymbol{\gamma}} \geq \mathbf{3}\right)$ truncated Gaussian circular pupils.

Aberration	$\Phi(\rho', \theta)$	σ_Φ	A_i' **for** $S = 0.8$
Spherical	$A_s' \rho'^4$	$2\sqrt{5}A_s'$	$\lambda/63$
Balanced spherical	$A_s' \left(\rho'^4 - 4\rho'^2\right)$	$2\,A_s'$	$\lambda/28$
Coma	$A_c' \rho'^3 \cos\theta$	$\sqrt{3}A_c'$	$\lambda/24$
Balanced coma	$A_c' \left(\rho'^3 - 2\rho'\right)\cos\theta$	A_c'	$\lambda/14$
Astigmatism	$A_a' \rho'^2 \cos^2\theta$	$A_a'/\sqrt{2}$	$\lambda/10$
Balanced astigmatism	$A_a' \rho'^2 \left(\cos^2\theta - 1/2\right)$	$A_a'/2$	$\lambda/7$
Defocus	$B_d' \rho'^2$	$\sqrt{3}B_d'$	$\lambda/24$
Tilt	$B_t' \rho' \cos\theta$	$\sqrt{3}B_t'$	$\lambda/20$

9.4 SUMMARY

Imaging by systems with an annular pupil is quite common, e.g., the Hubble telescope. Compared to a circular pupil, a corresponding annular pupil reduces the amount of light passing thorough it, yields a smaller central bright disc with a smaller central irradiance but brighter diffraction rings. Correspondingly, the OTF of a system with an annular pupil is lower at small spatial frequencies but higher at the high frequencies. While the obscuration does not change the cutoff frequency, its effect on the aberration tolerances for a certain Srehl ratio depends on the type of the aberration. As may be seen from Figure 9-6, the depth of focus, for example, increases with obscuration, but tolerance for astigmatism decreases. The balanced aberrations for annular pupils can be identified with the corresponding Zernike annular polynomials. They are unique for the annular pupils, just as the circle polynomials are unique for the circular pupils.

Pupils with Gaussian illumination across them are referred as Gaussian pupils. The Gaussian illumination may be due to a Gaussian-transmittance filter placed at the pupil, as in an apodized system, or the beam incident on the pupil may itself be Gaussian, as in the case of a laser transmitter. Whereas pupil or beam obscuration reduces the size of the central bright spot of the diffraction image or the pattern, the Gaussian illumination of the pupil increases the size. For a given total power, the central value of the aberration-free diffraction image for a Gaussian pupil is lower than that for a corresponding uniform pupil.

As the Gaussian illumination becomes narrower, the diffraction pattern approaches a Gaussian distribution. The image distribution can be approximated by a Gaussian when the pupil radius is twice the Gaussian radius (at which the amplitude drops to $1/e$ of its value at the center). The Strehl ratio for a small aberration depends on the aberration variance calculated across the Gaussian amplitude-weighted pupil. The aberration tolerance for a Gaussian pupil is higher than that for a uniformly illuminated pupil, because the illumination decreases but the aberrations generally increase with the distance from the center. The Gaussian approximation of the aberrated image distribution is valid when the pupil radius is three times the Gaussian radius.

The balanced aberrations for a Gaussian pupil can be identified with the corresponding Zernike-Gauss polynomials. These polynomials are unique for the Gaussian pupils, just as the Zernike circle or annular polynomials are unique for the circular or annular pupils. We note that the form of balanced astigmatism for annular and Gaussian pupils is the same as that for the circular pupils.

References

1. H. F Tschunko, “Imaging performance of annular apertures,” *Appl. Opt.* **18**, 1820–1823 (1974).

2. V. N. Mahajan, “Strehl ratio for primary aberrations in terms of their aberration variance,” *J. Opt. Soc. Am.* **73**, 860–861 (1983).

3. V. N. Mahajan, “Strehl ratio for primary aberrations: some analytical results for circular and annular pupils,” *J. Opt. Soc. Am*. **72**, 1258–1266 (1982).

4. V. N. Mahajan, “Zernike annular polynomials for imaging systems with annular pupils,” *J. Opt. Soc. Am*. **71**, 75–85, 1408 (1981), and *J. Opt. Soc. Am A* **l**, 685 (1984).

5. V. N. Mahajan, “Uniform versus Gaussian beams: a comparison of the effects of diffraction, obscuration, and aberrations,” *J. Opt. Soc. Am. A* **3**, 470–485(1986).

6. V. N. Mahajan, “Gaussian apodization and beam propagation,” *Progress in Optics*, **49**, 1–96, (2006).

7. V. N. Mahajan, “Axial irradiance of a focused beam,” *J. Opt. Soc. Am. A* **22**, 1813–1823 (2005).

8. V. N. Mahajan, “Strehl ratio of a Gaussian beam,” *J. Opt. Soc. Am. A* **22**, 1824–1833 (2005).

CHAPTER 10
Line of Sight of an Aberrated System

10.1 INTRODUCTION

In this chapter we consider the *line of sight* (LOS) of an aberrated optical system. The LOS is assumed here to coincide with the *centroid* of its diffraction point-spread function (PSF). For an aberration-free system, it coincides with the center of the PSF. For an aberrated system, it depends on the various orders of its coma aberrations. Thus, a coma aberration not only reduces the central value of the PSF like any other aberration, but it also shifts its centroid. We consider here PSFs aberrated by primary coma and give numerical results on the location of their peaks and centroids.

10.2 THEORY

The LOS of an aberration-free optical system coincides with the center of its diffraction PSF. For an aberrated system, let us define its LOS as the centroid of its aberrated PSF. Thus, if $I(x, y)$ represents the irradiance distribution of the aberrated image of a point object, its centroid $\langle x, y \rangle$ representing the LOS error of the system is given by

$$\langle x, y \rangle = P^{-1} \iint (x, y) I(x, y)\, dx\, dy \quad , \tag{10-1}$$

where P is the total power in the image. It can be shown that the centroid thus obtained is identical to that obtained from the *geometrical* PSF.[1] Let the aberration function in terms of the Zernike circle polynomials (see Section 8.3.7) for a system with a *circular* exit pupil be given by

$$W(\rho, \theta) = \sum_{n=0}^{\infty} \sum_{m=0}^{n} \left[2(n+1)/(1+\delta_{m0})\right]^{1/2} R_n^m(\rho)(c_{nm} \cos m\theta + s_{nm} \sin m\theta) \quad , \tag{10-2}$$

where c_{nm} and s_{nm} are the Zernike aberration coefficients representing the standard deviations of the corresponding aberration terms across the pupil (with the exception of the piston term $n = 0 = m$, which has a standard deviation of zero). It can be shown that the centroid of its aberrated PSF for a uniformly illuminated pupil is given by

$$\langle x, y \rangle = 2F \sum_{n=1}^{\infty}{}' \sqrt{2(n+1)}(c_{n1}, s_{n1}) \quad , \tag{10-3}$$

where F is the focal ratio or the f-number of the image-forming light cone and a prime indicates a summation over odd integral values of n. We note that only those aberrations contribute to the LOS *errors* that vary with θ as $\cos\theta$ and $\sin\theta$. Aberrations varying as $\cos\theta$ contribute to $\langle x \rangle$ and those varying as $\sin\theta$ contribute to $\langle y \rangle$. For a given value of c_{n1} or s_{n1}, an aberration of a higher order gives a larger LOS error because of the $\sqrt{2(n+1)}$ factor. Thus, two Zernike aberrations with $m = 1$ but different values of n having the same standard deviation give different LOS errors, even though they give

(approximately) the same Strehl ratio. (See Section 8.3.1 for a relationship between the Strehl ratio and the standard deviation of an aberration.)

If we consider an aberration in the form

$$W(\rho, \theta) = W_k \rho^k \cos\theta \quad , \tag{10-4}$$

where k is an odd integer, we find that

$$\langle x, y \rangle = (2FW_k, 0) \quad . \tag{10-5}$$

Thus, the LOS error depends on the value of the peak aberration W_k but not on k. We note that for $k = 3$, the aberration is primary coma, and for $k = 3$ it is secondary coma, but they both give the same LOS error if $W_3 = W_5$, even though the corresponding PSFs are quite different. The reason for the same LOS error is that for a uniform circular pupil, the centroid depends only on the aberration along the perimeter of the pupil, which depends on W_k but not on k.[1]

10.3 NUMERICAL RESULTS

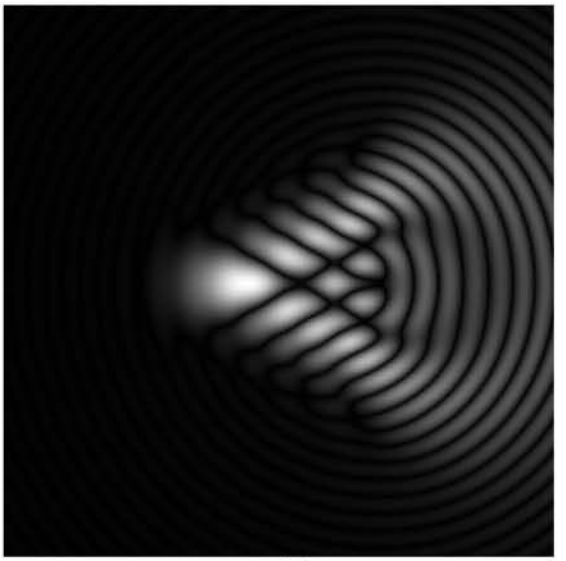
Figure 10-1 shows a 2D PSF for 5λ of primary coma, and the central profiles of the PSFs for coma varying from 0 to 2λ, normalized by the aberration-free central irradiance. Note that λ is the optical wavelength. The locations of the peak x_p and centroid <x> of the aberrated PSFs are given in Table 10-1. The irradiances I_p and I_c at these points and $I(0, 0)$ at the PSF center are also given in this table. For example, when $W_3 = 0.5\lambda$, the Strehl ratio of the PSF is approximately equal to 0.32, but its peak value of 0.87 lies at the point (0.66, 0) compared to a value of unity at the center (0, 0) of the corresponding aberration-free PSF. The centroid of the PSF lies at (1, 0). Thus, the centroid of the PSF shifts by an amount approximately equal to the radius 1.22 (in units of λF) of the Airy disc.

The point with respect to which the variance of coma aberration is minimized is indicated by x_m (which from Section 8.3.3 is equal to $4FW_3/3$), and the irradiance at this point is given by I_m. We note that x_m and x_p are approximately equal to each other only for small values of W_3 $(< 0.7\lambda)$, showing that coma balanced with wavefront tilt to give minimum aberration variance across the pupil (i.e., Zernike coma) yields a maximum of irradiance only for small aberrations.

Figure 10-2 and Table 10-2 give similar information for secondary coma. Comparing the figures and tables, we note that, although the PSFs for the same value of primary coma W_3 and secondary coma W_5 are different, their centroids are the same.

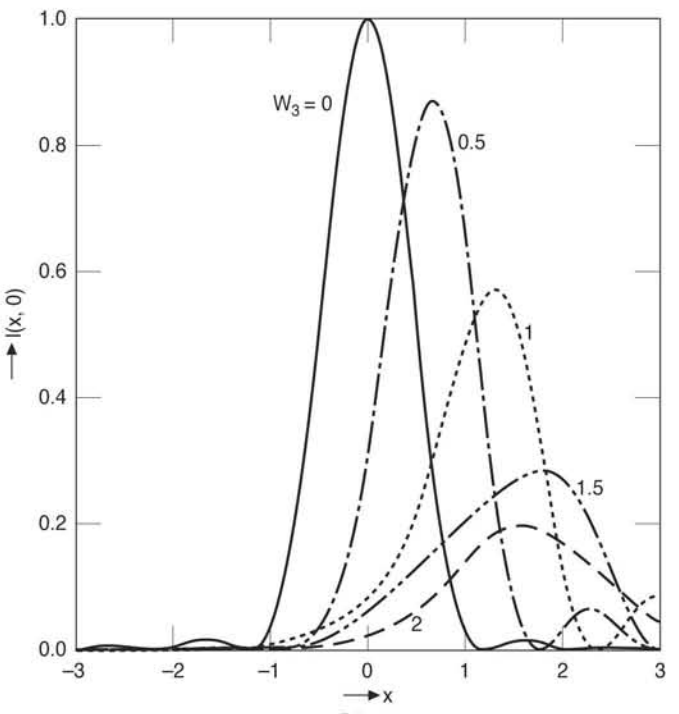
10.4 COMMENTS

The results given here are applicable to both imaging systems, e.g., those used for optical surveillance, as well as to laser transmitters used for active illumination of a target. In both cases, the LOS of the optical system is extremely important. An LOS error



(a)

$W_3 = 5\,\lambda$



(b)

Figure 10-1. (a) 2D PSF for $5\,\lambda$ of primary coma W_3 and (b) PSF profiles $I(x, 0)$ for several typical values of W_3 in units of λ.

Table 10-1. Typical values of x_n, x_p, and $<x>$ and corresponding irradiances I_m, I_p, and I_c for circular pupils aberrated by primary coma.

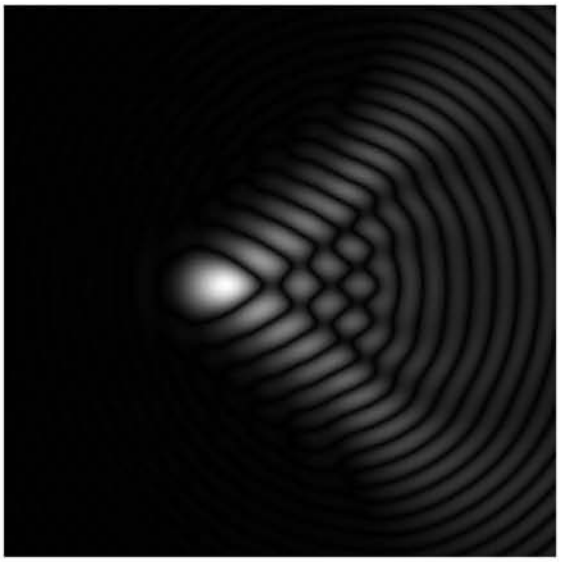
W_3	x_m	x_p	$<x>$	I_m	I_p	I_c	$I(0)$
0	0	0	0	1	1	1	1
0.5	0.67	0.66	1.00	0.8712	0.8712	0.6535	0.3175
1.0	1.33	1.30	2.00	0.5708	0.5717	0.1445	0.0791
1.5	2.00	1.80	3.00	0.2715	0.2844	0.0004	0.0618
2.0	2.67	1.57	4.00	0.0864	0.1978	0.0061	0.0341

Table 10-2. Typical values of x_m, x_p, and $<x>$ and corresponding irradiances I_m, I_p, and I_c for circular pupils aberrated by secondary coma.

W_5	x_m	x_p	$<x>$	I_m	I_p	I_c	$I(0)$
0	0	0	0	1	1	1	1
0.5	0.50	0.49	1.00	0.8150	0.8153	0.4114	0.4955
1.0	1.00	0.83	2.00	0.4464	0.4664	0.0025	0.2332
1.5	1.50	0.81	3.00	0.1685	0.3237	0.0098	0.1873
2.0	2.00	1.11	4.00	0.0420	0.2523	0.0073	0.1389

of a surveillance system will produce an error in the location of the target. In the case of a laser transmitter, a large LOS error may cause the laser beam to miss the target altogether. Whereas for static aberrations we may be able to calibrate the LOS, for dynamic aberrations it is the analysis given here that will determine the tolerances of aberrations of the type $\rho^k \cos\theta$ and $\rho^k \sin\theta$. Although we have defined the LOS of an optical system in terms of the centroid of its PSF, it could have been defined in terms of the peak of the PSF (assuming that the aberrations are small enough so that the PSF has a distinguishable peak). For an aberration-free PSF, its peak value and its centroid both lie at its origin, regardless of the amplitude variations across its pupil. The two are not coincident when $\cos\theta$ and/or $\sin\theta$ dependent aberrations are present. The precise definition of the LOS will perhaps depend on the nature of the application of the optical system. Moreover, in practice, only a finite central portion of the PSF will be sampled to measure its centroid, and the precision of this measurement will be limited by the noise characteristics of the photodetector array.

For simplicity, we have limited our discussion here to optical systems with uniform circular pupils. However, the analysis can be extended to obtain the LOS errors of aberrated systems with *annular* and/or *Gaussian pupils.*[1] For example, for an annular pupil with a central obscuration ϵ, the right-hand side of Eq. (10-5) is multiplied by $1+\epsilon^2$ for k = 3. Compared to a uniform pupil, the value of <x> for a Gaussian pupil is smaller; i.e., the centroid for a Gaussian pupil is closer to the true (aberration-free) LOS.



$W_5 = 5\,\lambda$

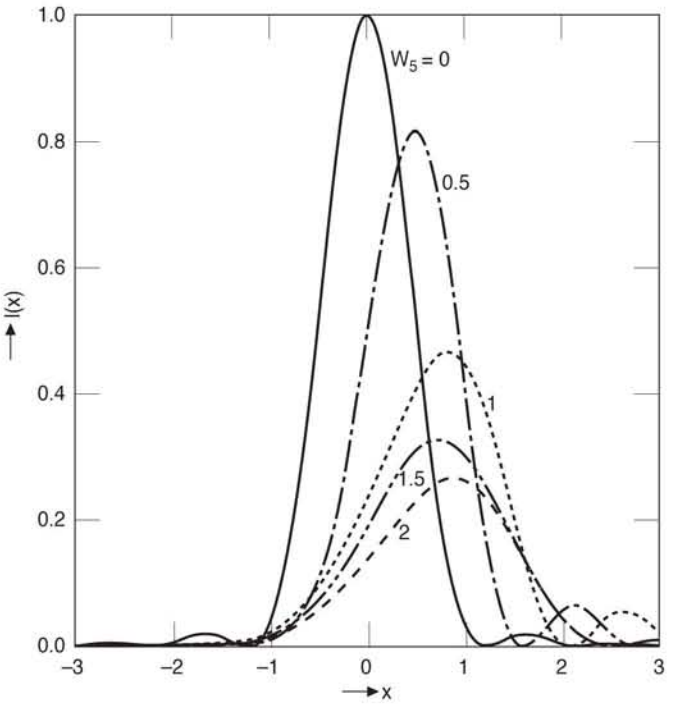


Figure 10-2. (a) 2D PSF for $5\,\lambda$ of secondary coma W_5 and (b) PSF profiles $I(x, 0)$ for several typical values of W_5 in units of λ.

10.5 SUMMARY

The line of sight of a system is determined by the centroid of the image of a point object. The centroid of an aberrated diffraction image is the same as that of the corresponding spot diagram. For an aberration-free system, the centroid lies at the center of the image due to its radial symmetry. Only the coma aberrations displace the centroid. The displacement depends on the magnitude of a coma aberration, but not on its order. Figures 10-1 and 10-2 illustrate this point for five waves of primary and secondary comas. The location of the peak value, centroid, and the irradiance at these locations are given in Tables 10-1 and 10-2 for the primary and secondary comas, respectively. The Strehl ratio $I(0)$ is also given in these tables.

References

1. V. N. Mahajan, "Line of sight of an aberrated optical system," *J. Opt. Soc. Am. A* **2**, 833–846 (1985).

CHAPTER 11
Random Aberrations

11.1 INTRODUCTION

So far we have considered *deterministic aberrations* such as those that are inherent in the design of an optical imaging system. These aberrations are deterministic in the sense that they are known or can be calculated, for example, by ray tracing the system. Now we consider the effects of aberrations that are *random* in nature on the quality of images. The aberration is random in the sense that it varies randomly with time for a given system, or it varies randomly from one sample of a system to another. An example of the first kind is the aberration introduced by *atmospheric turbulence* when an optical wave propagates through it, as in ground-based astronomical observations. An example of the second kind is the aberration introduced due to *polishing errors* of the optical elements of the system. The polishing errors of an element fabricated similarly in large quantities vary randomly from one sample to another. In either case, we cannot obtain the exact image unless the instantaneous aberration or the exact polishing errors are known. However, based on the statistics of the aberrations, we can obtain the time- or ensemble-averaged image.

We discuss the effects of two types of random aberrations: *random wavefront tilt* or *defocus* causing *random image motion*, and random aberrations introduced by atmospheric turbulence. The time-averaged Strehl ratio, PSF, OTF, and encircled power are discussed for the two types of aberrations. Although much of our discussion is on systems with circular pupils, systems with annular pupils are also considered. A brief discussion on the aberrations resulting from *fabrication errors* is also given.

11.2 RANDOM IMAGE MOTION

In many optical imaging systems, especially those used in space, there is always some image motion during an exposure interval. The source of image motion may, for example, be vibration of optical elements and servo dither in the pointing system. The image motion may be transverse or longitudinal, i.e., normal to or along the optical axis, respectively. In the case of beam transmitting systems, the beam itself may have some motion associated with it. We give expressions for the time-averaged PSF, Strehl ratio, OTF, and encircled power for an imaging system with a circular or an annular exit pupil undergoing Gaussian random motion. We show that the Strehl ratio is more sensitive to obscuration in the case of transverse motion, but less sensitive in the case of longitudinal motion.

11.2.1 Transverse Image Motion[1]

The time-averaged PSF for a system with a circular pupil in the case of transverse image motion characterized by Gaussian functions of zero mean and equal standard deviation σ in units of λF along the two orthogonal axes of the image plane is given by

$$\langle I(r;\sigma)\rangle = 8\int_0^1 \langle \tau(v;\sigma)\rangle\, J_0(2\pi r v)\, v\, dv \quad , \tag{11-1}$$

where

$$\langle \tau(v;\sigma)\rangle = \tau(v)\exp\left(-2\pi^2\sigma^2 v^2\right) \tag{11-2}$$

is the time-averaged OTF. The motion-free OTF $\tau(v)$ is given by Eq. (8-37). Letting $r = 0$ in Eq. (11-1), we obtain the time-averaged Strehl ratio:

$$\langle S(\sigma)\rangle = 8\int_0^1 \langle \tau(v;\sigma)\rangle\, v\, dv \quad . \tag{11-3}$$

The time-averaged encircled power in terms of the OTF is given by

$$\langle P(r_c;\sigma)\rangle = 2\pi r_c \int_0^1 \langle \tau(v;\sigma)\rangle\, J_1(2\pi r_c v)\, dv \quad . \tag{11-4}$$

The corresponding equations for a system with an annular pupil with an obscuration ratio ϵ are:

$$\langle I(r;\epsilon;\sigma)\rangle = \left[8/\left(1-\epsilon^2\right)\right]\int_0^1 \langle \tau(v;\epsilon;\sigma)\rangle\, J_0(2\pi r v)\, v\, dv \quad , \tag{11-5}$$

$$\langle S(\epsilon;\sigma)\rangle = \left[8/\left(1-\epsilon^2\right)\right]\int_0^1 \langle \tau(v;\epsilon;\sigma)\rangle\, v\, dv \quad , \tag{11-6}$$

$$\langle P(r_c;\epsilon,\sigma)\rangle = 2\pi r_c \int \tau(v;\epsilon,\sigma)\, J_1(2\pi v r_c)\, dv \quad , \tag{11-7}$$

and

$$\langle \tau(v;\epsilon;\sigma)\rangle = \tau(v;\epsilon)\exp\left(-2\pi^2\sigma^2 v^2\right) \quad . \tag{11-8}$$

The motion-free OTF $\tau(v;\epsilon)$ is given by Eq. (9-3).

Figure 11-1 shows how the Strehl ratio varies with σ for $\epsilon = 0(0.25)\,0.75$. It decreases monotonically as σ increases. We note that as ϵ increases, the drop in Strehl ratio due to image motion for a given value of σ increases. This occurs because the motion-free PSF (normalized to unity at the origin) for a larger value of ϵ is smaller for small values of r for $r \le 1$.

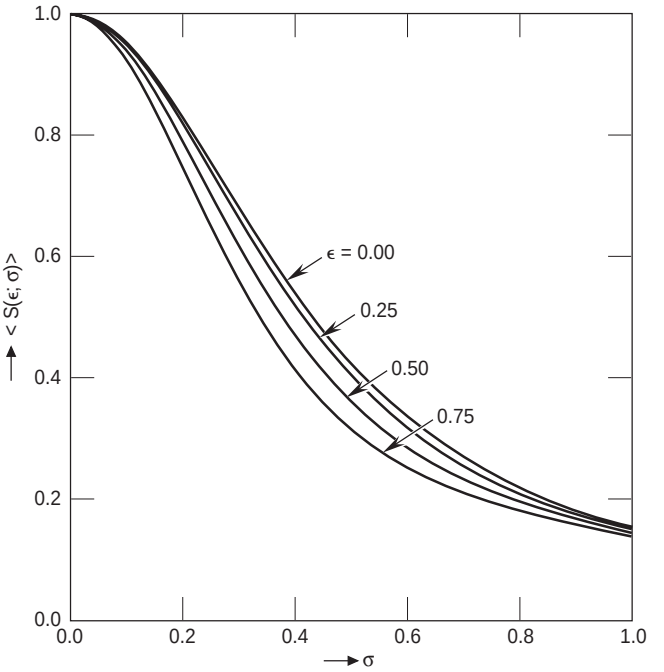


Figure 11-1. Time-averaged Strehl ratio as a function of the standard deviation σ of the transverse image motion for several typical values of ϵ.

11.2.2 Longitudinal Image Motion[2]

In the case of transverse image motion, the aberration-free image is randomly displaced in the image plane. However, in the case of longitudinal image motion, the image is randomly defocused. If the random defocus varies very slowly with time and the exposure time of an observation is small, then the defocused image at the time of an observation is given by the defocused PSF. However, if the exposure time is long enough that the image moves back and forth during that time, then we must average the defocused images.

The defocused PSF of a system with an annular pupil is given by

$$I(r;\Delta;\epsilon) = \left(\frac{2}{1-\epsilon^2}\right)^2 \left| \int_{\epsilon}^{1} \exp\left(-2\pi i \Delta \rho^2\right) J_0(\pi r \rho) \rho \, d\rho \right|^2 , \tag{11-9}$$

where Δ is the longitudinal defocus in units of $8\lambda F^2$. Thus, a value of $\Delta = 1$ in these units represents a defocus phase aberration of 2π or a wave aberration of one wave. Letting $r = 0$ in Eq. (11-9), the corresponding Strehl ratio is given by

$$S(\epsilon) = \left\{ \frac{\sin\left[\pi\Delta\left(1-\epsilon^2\right)\right]}{\pi\Delta\left(1-\epsilon^2\right)} \right\}^2 . \tag{11-10}$$

The time-averaged PSF, Strehl ratio, and encircled power for longitudinal Gaussian image motion characterized by zero mean and standard deviation σ in the same units as Δ is given by

$$\langle I(r;\sigma;\epsilon)\rangle = \frac{1}{\sqrt{2\pi}\,\sigma}\int_{\infty}^{\infty} I(r;\Delta;\epsilon)\exp\left(-\Delta^2/2\sigma^2\right)d\Delta \quad , \tag{11-11}$$

$$\langle S(\sigma;\epsilon)\rangle = \frac{1}{\sqrt{2\pi}\,\sigma}\int_{\infty}^{\infty}\left\{\frac{\sin\left[\pi\Delta\left(1-\epsilon^2\right)\right]}{\pi\Delta\left(1-\epsilon^2\right)}\right\}^2 \exp\left(-\Delta^2/2\sigma^2\right)d\Delta \quad , \tag{11-12}$$

and

$$\langle P(r_c;\sigma;\epsilon)\rangle = \frac{\pi^2}{2}\left(1-\epsilon^2\right)\int_0^{r_c}\langle I(r;\sigma;\epsilon)\rangle r\,dr \quad . \tag{11-13}$$

Figure 11-2 shows how the time-averaged Strehl ratio $\langle S(\sigma;\epsilon)\rangle$ varies with the standard deviation σ of the image motion. As expected, the Strehl ratio decreases as the image motion increases. However, the decrease is smaller for a large value of the obscuration ratio ϵ, or the Strehl ratio for a given value of σ is larger for a large value of ϵ. This is a consequence of the fact that the depth of focus is larger for a large value of ϵ, as discussed in Section 9.2.4. This effect is opposite to that of the transverse image motion, where the drop in Strehl ratio with σ increases as ϵ increases due to the narrower PSF for an obscured pupil.

We note from Eq. (11-10) that the static Strehl ratio is zero for integral values of $\Delta\left(1-\epsilon^2\right)$. Thus, for a circular pupil, for example, it is zero when the defocus wave aberration is one wave or the longitudinal defocus Δ (in units of $8\lambda F^2$) is unity. However, the time-averaged value of the dynamic Strehl ratio for $\sigma = 1$ is 0.3483. Similarly, for an annular pupil with $\epsilon = 0.5$, it is zero when $\Delta = 4/3$, but the time-averaged dynamic Strehl ratio for $\sigma = 4/3$ is approximately 0.35.

11.3 IMAGING THROUGH ATMOSPHERIC TURBULENCE[3]

11.3.1 Introduction

The resolution of a telescope forming an aberration-free image is determined by the diameter *D* of its pupil; the larger the diameter, the better the resolution. However, in ground-based astronomy, the resolution is degraded considerably because of the aberrations introduced by atmospheric turbulence. A plane wave of uniform amplitude and phase representing the light from a star propagating through the atmosphere undergoes both amplitude and phase variations due to the random inhomogeneities in its refractive index. The amplitude variations, called scintillations, result in the twinkling of

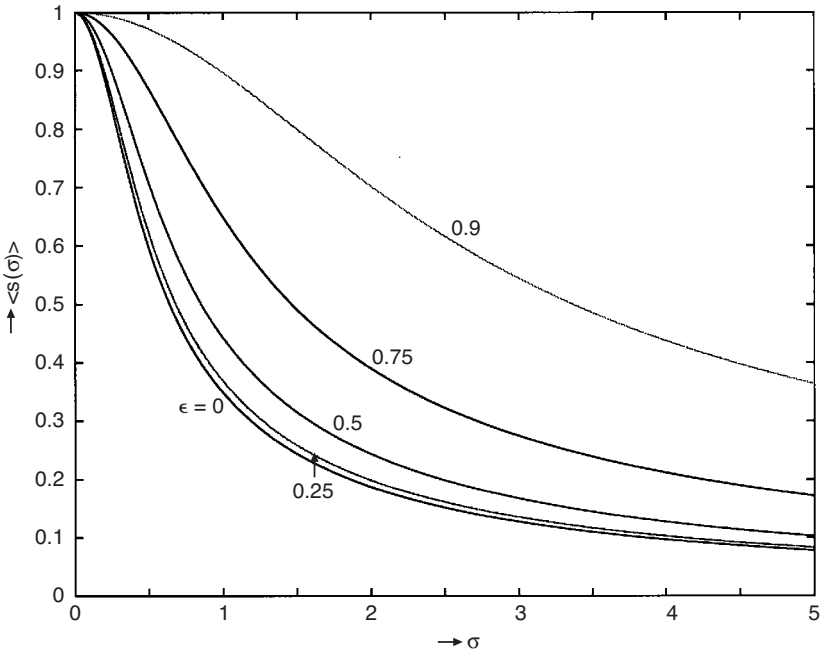


Figure 11-2. Time-averaged Strehl ratio $\langle S(\sigma;\epsilon)\rangle$ as a function of the standard deviation σ of the longitudinal Gaussian random image motion, where ϵ is the obscuration ratio of the annular pupil. σ is in units of $8\lambda F^2$, and its numerical value represents the peak defocus wave aberration in units of wavelength.

stars. The purpose of a large ground-based telescope, such as the 5-m telescope at Mt. Palomar, has generally not been better resolution but to collect more light so that dim objects may be observed. Of course, with the advent of *adaptive optics*,[4-6] the resolution can be improved by correcting the phase aberrations with a deformable mirror.

11.3.2 Long-Exposure Image

For Kolmogorov turbulence, the time-averaged OTF for a distorted wavefront representing a long-exposure (LE) image is given by[7]

$$\langle \tau(v; D/r_0)\rangle = \tau(v)\exp\left[-3.44\left(vD/r_0\right)^{5/3}\right] , \tag{11-14}$$

where D is the diameter of the telescope and r_0 is Fried's coherence length of turbulence.[8,9] The exponential factor in Eq. (11-14) represents the mutual coherence function of the wave at the telescope.

Since $\exp(-3.44) \simeq 0.03$, atmospheric turbulence reduces the overall system MTF corresponding to a spatial frequency $v = r_0/D$ by a factor of 0.03. Similarly, the degree of coherence of complex amplitudes at two points on a wave separated by r_0 is only 0.03, or that the visibility of the fringes formed by the secondary waves from these points is 0.03. The value of r_0 on a mountain site may vary from 5 to 10 cm in the visible region of the spectrum and increases with wavelength as $\lambda^{6/5}$.

11.3.2.1 *Imaging With Circular Pupils*

Substituting Eq. (11-14) into the imaging equations given in Section 11.2.1, we can obtain the Strehl ratio, and the irradiance and encircled power distributions. Figure 11-3 shows how the Strehl ratio decreases monotonically as D/r_0 increases. Thus, for example, for a given value of D, the Strehl ratio decreases rapidly as r_0 decreases. Even when r_0 is as large as D, the Strehl ratio is only 0.445.

The phase aberration variance for Kolmogorov turbulence is given by

$$\sigma_\Phi^2 = 1.03\left(D/r_0\right)^{5/3} \quad . \tag{11-15}$$

Substituting Eq. (11-15) into Eq. (8-15), we obtain the approximate Strehl ratio:

$$\left\langle S_1\left(D/r_0\right)\right\rangle \simeq \exp\left[-1.03\left(D/r_0\right)^{5/3}\right] \quad . \tag{11-16}$$

Its variation with D/r_0 is also shown in Figure 11-3. We note that it considerably underestimates the true Strehl ratio $\langle S\rangle$. A much better approximation is given by

$$\left\langle S_2\left(D/r_0\right)\right\rangle \simeq \left[1+\left(D/r_0\right)^{5/3}\right]^{6/5} \quad , \tag{11-17}$$

as is evident from its plot in Figure 11-3.

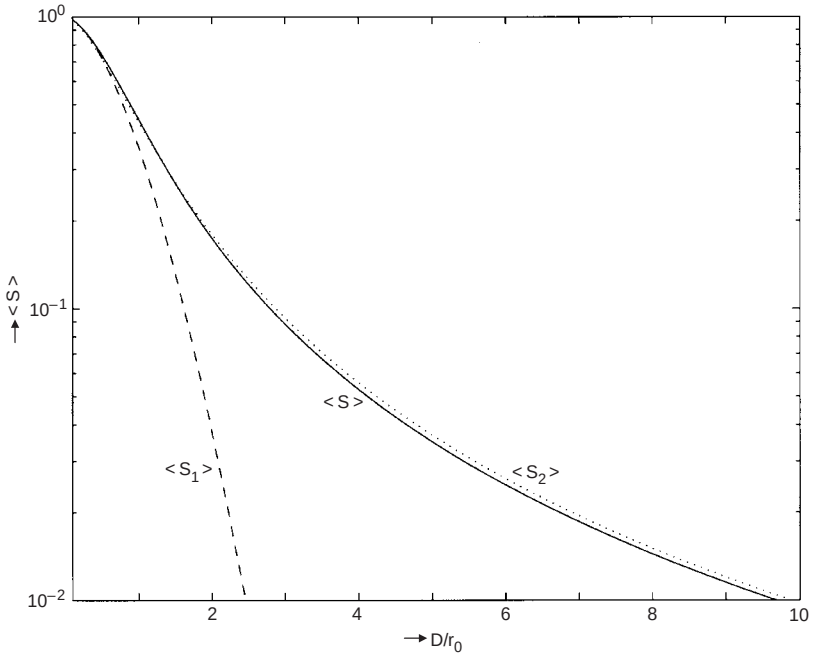


Figure 11-3. Variation of time-averaged Strehl ratio with D/r_0. The solid curve represents the exact value $\langle S\rangle$, the dashed curve represents the approximate value $\langle S_1\rangle$, and the dotted curve represents the approximate value $\langle S_2\rangle$.

From Section 8.2, the aberration-free central irradiance is given by $PS_p/\lambda^2 R^2$, where P is the total power, $S_p = \pi\left(D^2/4\right)$ is the area of a circular pupil of diameter D, and R is the distance between the planes of the exit pupil and the image. It is clear that, for a fixed total power, this irradiance relative to its aberration-free value for a pupil of diameter r_0 increases as $\left(D/r_0\right)^2$. The corresponding aberrated central irradiance is given by

$$\eta\left(D/r_0\right) = \left(D/r_0\right)^2 \langle S \rangle \quad . \tag{11-18}$$

Figure 11-4 shows a comparison of the aberration-free and aberrated central irradiances as a function of D/r_0. We note that for small values of D, η increases as for an aberration-free system, indicating a small effect of turbulence. As D increases, η increases much more slowly, and the increase for $D/r_0 > 5$ is very small. As $D/r_0 \to \infty$, $\eta \to 1$. The two asymptotes of $\eta\left(D/r_0\right)$ intersect at $D/r_0 = 1$. Indeed Fried[7] defined r_0 in a way so as to yield this result. He called the quantity η the *normalized resolution*.

In astronomical observations, the power P increases as D increases. However, if the observation is made against a uniform background, then the background irradiance in the image also increases as D^2. Hence, the detectability of a point object is limited by turbulence to a value corresponding to a pupil of diameter r_0, no matter how large the actual diameter D is. In the case of a laser transmitter with a fixed value of laser power, the central irradiance on a target is again limited to its aberration-free value for a beam of diameter r_0, no matter how large the actual transmitter diameter is.

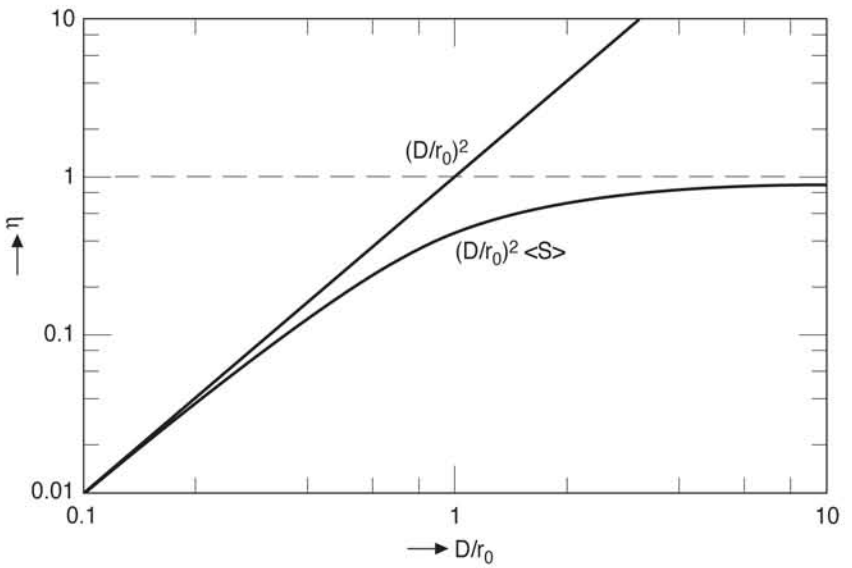


Figure 11-4. Central irradiance η for a fixed total power as a function of D/r_0. Its aberration-free value increases as $\left(D/r_0\right)^2$, but its aberrated value approaches unity as $D/r_0 \to \infty$.

Figure 11-5 shows the irradiance distribution normalized to unity at the center for several values of D/r_0. The diffraction rings disappear even for small values of D/r_0 such as 1, and the PSFs become smooth and may be approximated by Gaussian functions. The corresponding encircled power is also shown in this figure. As D/r_0 increases, a given fraction of the total power is contained in an increasingly larger circle. As an example, whereas 84% of the total power is contained in a circle of radius $r_c = 1.22$ when there is no turbulence, it is contained in a circle of radius 1.9 when $D/r_0 = 1$.

11.3.2.2 *Imaging With Annular Pupils*

Similar results are obtained for systems with annular pupils.[10] Figure 11-6 shows how, for fixed total power, the aberrated central irradiance relative to its aberration-free value for a circular pupil of diameter r_0,

$$\eta(\epsilon; D/r_0) = (1-\epsilon^2)(D/r_0)^2 \langle S(\epsilon; D/r_0)\rangle \quad , \tag{11-19}$$

varies with D/r_0. The aberration-free central irradiance varying as $(1-\epsilon^2)(D/r_0)^2$ is illustrated by the straight lines for several values of ϵ. For small values of D/r_0, η increases with D/r_0, as does its aberration-free counterpart. However, for larger values of D/r_0, it increases slowly with a negligible increase beyond a certain value of D/r_0, depending on the value of ϵ. The saturation effects of atmospheric turbulence occur at larger and larger values of D/r_0 as ϵ increases. Irrespective of the value of ϵ,

$$\eta(\epsilon; D/r_0) \to 1 \text{ as } D/r_0 \to \infty \quad , \tag{11-20}$$

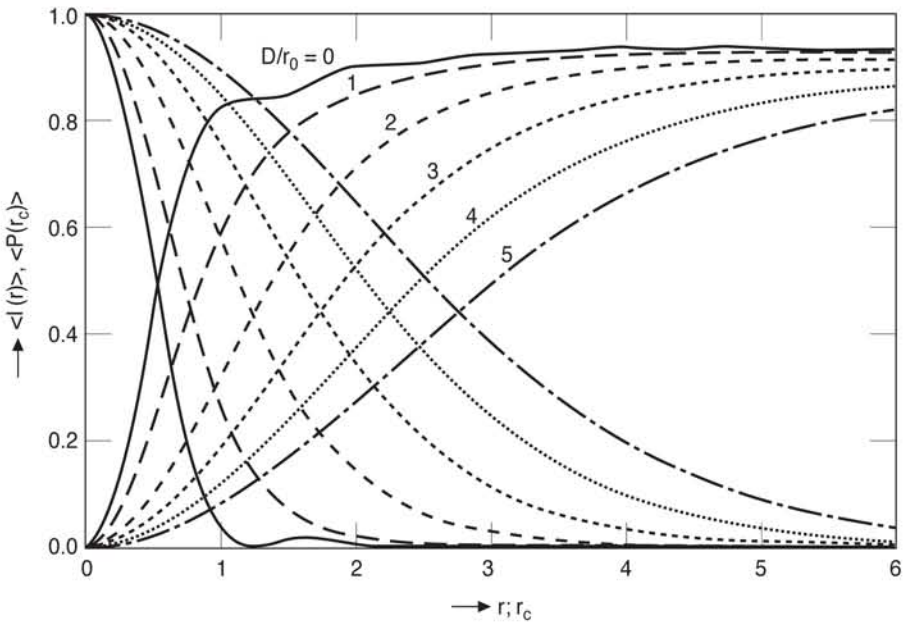


Figure 11-5. Time-averaged irradiance and encircled-power distributions for different values of D/r_0.

as in the case of circular pupils. The two asymptotes of $\eta(\epsilon; D/r_0)$ for a given value of ϵ intersect at the point $D/r_0 = (1-\epsilon^2)^{-1/2}$. Hence, regardless of how large D is, the central irradiance is less than or equal to the aberration-free central irradiance for a system with a circular pupil of diameter r_0, equality approaching as $D/r_0 \to \infty$. The limiting value of the central irradiance is independent of the value of ϵ. Some typical values of the Strehl ratio are given in Table 11-3 for several values of D/r_0 and ϵ.

Figure 11-7 shows how the irradiance distribution or the PSF and encircled power change as D/r_0 increases for $\epsilon = 0.5$. The PSFs are normalized to unity at the center, the actual central value being the Strehl ratio given in Table 11-1. As D/r_0 increases, the diffraction rings disappear and the PSFs become smooth, and a given fraction of total power is contained in a circle of larger and larger radius.

The aberration variance increases monotonically from $1.03(D/r_0)^{5/3}$ for a circular pupil $(\epsilon = 0)$ to $1.84(D/r_0)^{5/3}$ for an infinitesimally thin ring pupil $(\epsilon \to 1)$.[11]

11.3.3 Short-Exposure Image

The aberration may be decomposed into various aberration types in terms of Zernike polynomials.[12] It is found, for example, that 87% of the aberration variance consists of wavefront tilt, i.e., random image motion. The variance of a short-exposure image (so it is not degraded by image motion) is given by

$$\sigma_\Phi^2 = 0.134(D/r_0)^{5/3} \quad . \tag{11-21}$$

Figure 11-8 shows short-exposure PSFs corresponding to different values of D/r_0. In Figure 11-8a, D is kept fixed while r_0 decreases from a value equal to D to $D/3$ and to $D/10$, e.g., $D = 1$ m, $r_0 = 1$ m, 33.3 cm, and 10 cm, respectively. We note that each image is broken up into small spots called *speckles,* which is a characteristic of random aberrations. The size of a speckle is determined by D, its angular radius being approximately equal to λ/D. The size of the total image is determined by r_0, its angular radius being approximately equal to λ/r_0. The image becomes progressively worse as r_0 decreases showing the effects of what astronomers call *seeing.* In Figure 11-8b, the value of r_0 is kept fixed while the value of D increases from a value equal to r_0 to $3\,r_0$ and to $10\,r_0$, e.g., $r_0 = 10$ cm, $D = 10$ cm, 30 cm, and 1 m, respectively. Now the size of a

Table 11-1. Time-averaged Strehl ratio for various values of ϵ and D/r_0.

$\epsilon \backslash D/r_0$	1	2	3	4	5
0	0.445	0.175	0.089	0.053	0.035
0.25	0.430	0.169	0.088	0.054	0.036
0.50	0.391	0.160	0.090	0.058	0.040
0.75	0.344	0.152	0.095	0.067	0.050

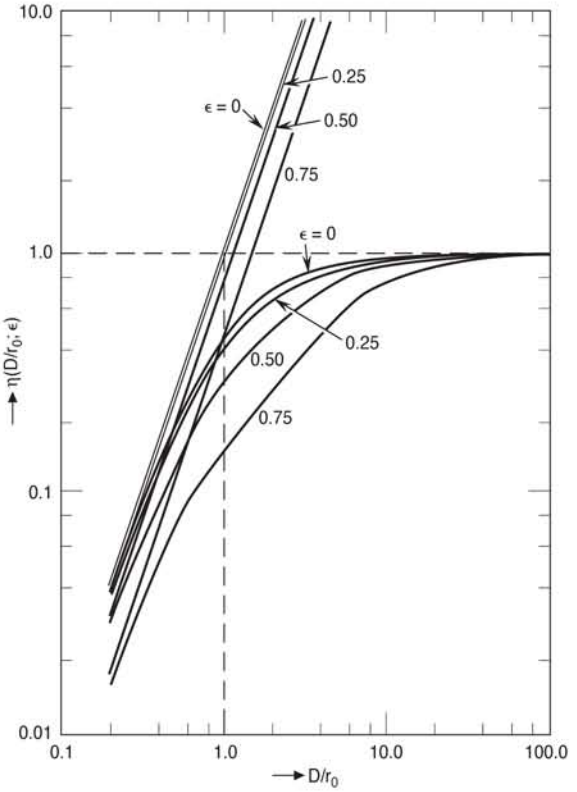


Figure 11-6. Variation of $\eta(\epsilon; D/r_0)$ with D/r_0 for several values of ϵ. Its aberration-free value given by $(1-\epsilon^2)(D/r_0)^2$ is represented by the straight lines. Its aberrated value approaches unity as $D/r_0 \to \infty$, regardless of the value of ϵ.

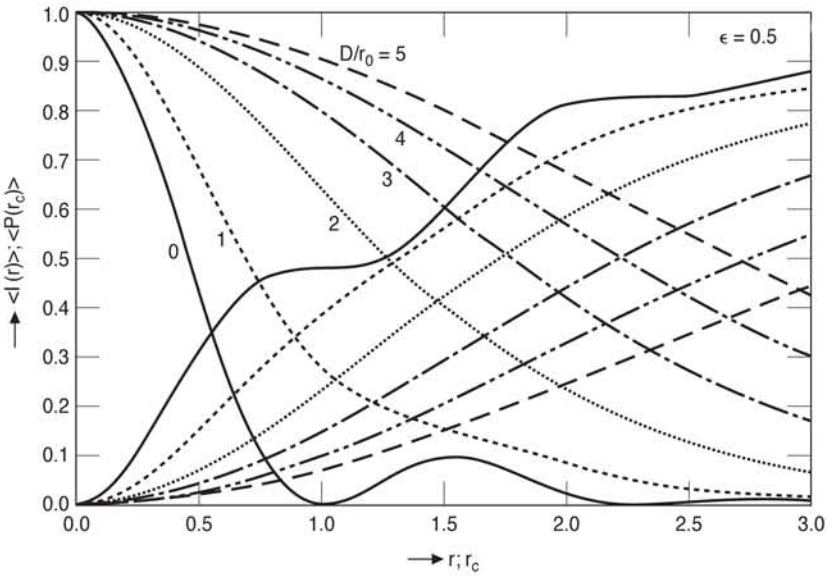
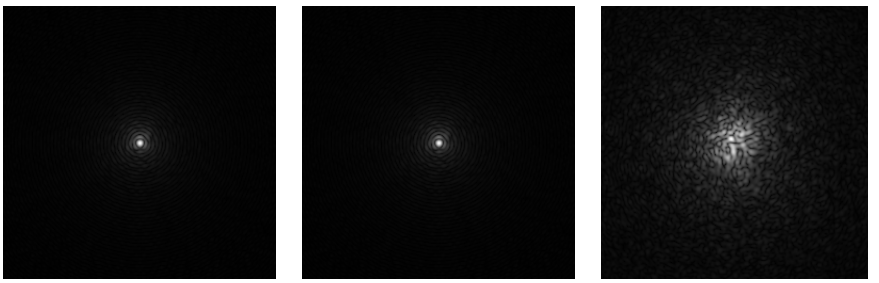


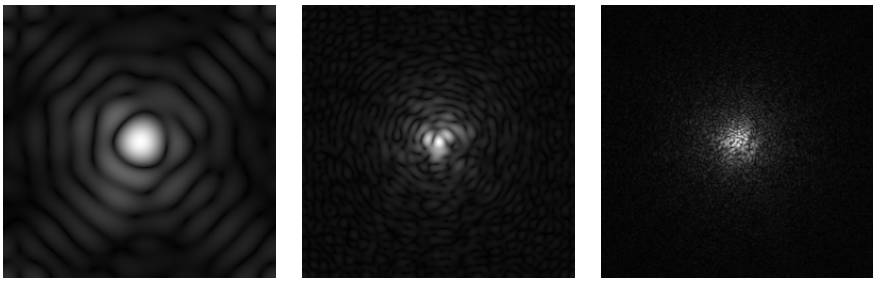
Figure 11-7. Time-averaged PSF and encircled power for several typical values of D/r_0 and $\epsilon = 0.5$.

speckle decreases as D increases, but the image size is approximately constant. Thus, an increase in D does not significantly improve the resolution of the system (as determined by the total size of the image). For convenience, the PSFs in Figure 11-8b are shown reduced by a factor of 1.5 compared to those in Figure 11-8a. Thus, for example, the pictures corresponding to D/r_0 =10 in these two parts are otherwise similar. (The aberration function used for this case and the corresponding interferogram are shown in Figure 12-4.) The approximate expressions of Eqs. (8-13)–(8-15) are not suitable for calculating the average Strehl ratios for random aberrations. For example, even for D/r_0 =1, Eq. (8-15) gives a Strehl ratio of 0.357, compared to a true value of 0.445. For larger values of D/r_0, Eq. (8-15) underestimates the average Strehl ratio by larger factors.



$D/r_0 = 1$ $D/r_0 = 3$ $D/r_0 = 10$

Figure 11-8a. Short-exposure PSFs aberrated by atmospheric turbulence. D is kept fixed and r_0 is varied. For example, $D = 1$ m and $r_0 = 1$ m, 33.3 cm, and 10 cm, giving $D/r_0 = 1$, 3, and 10. The value of D determines the size of a speckle, while r_0 determines the overall size of the image.



$D/r_0 = 1$ $D/r_0 = 3$ $D/r_0 = 10$

Figure 11-8b. Short-exposure PSFs aberrated by atmospheric turbulence. r_0 is kept fixed, and D is varied. For example, $r_0 = 10$ cm, and $D = 10$ cm, 30 cm, and 1 m, giving $D/r_0 = 1$, 3, and 10. The value of D determines the size of a speckle while r_0 determines the size of the overall image. For convenience, the PSFs shown here have been reduced by a factor of 1.5 compared to those in Figure 11-8a.

11.3.4 Lucky Imaging and Adaptive Optics

Since the aberrations introduced by turbulence are random in nature and since the wavefront tilt represents a large portion of these aberrations, it is conceivable that at a certain instant in time a short exposure (SE) image is practically aberration free. Thus an observer can take a series of short-exposure images over a time period before the turbulence characterized by r_0 changes significantly. The good-quality images are selected based on their full width at half maximum or peak pixel value normalized by the total illumination, centered on the peak pixel, and summed in a process called shift and add. This approach has been referred to as *lucky imaging*.[13,14]

Figure 11-9a shows a sample of a poor-quality SE star image, and 11-9b shows a corresponding good-quality image. Figure 11-9c shows an LE image obtained by adding all of the 50,000 SE images taken at a rate of 40/sec over a period of about 21 minutes.

The image quality improves when the centroids of the SE images are aligned, as in Figure 11-9d. This image is equivalent to an LE image where the tip and tilt of the wavefront have been corrected in real time. Figures 11-9e through 11-9g illustrate images of increasing quality as 50, 10, and 1 percent of the best selected images based on their full width at half maximum are aligned on their peak value and added.

The image quality can also be improved by correcting the wavefront errors in (near) real time with adaptive optics.[4-6] In practice, a steering mirror with only three actuators is used to correct the large tip and tilt of the wavefront. The residual aberration is corrected by a deformable mirror, which is deformed by an array of actuators attached to it. The signals for the actuators are determined either by sensing the wavefront errors with a wavefront sensor in a closed loop to minimize the variance of the residual errors, or the actuators are actuated to produce Zernike modes (e.g., focus, two modes of astigmatism, two modes of coma, etc.) iteratively until the sharpness of the image is maximized.[15-17] The signals are independent of the optical wavelength provided atmospheric dispersion is negligible. The two approaches are referred to as *zonal* and *modal approaches*, respectively. The zonal approach has the advantage that the rate of correction is limited only by the rate at which the wavefront errors can be sensed and the actuators can be actuated. However, the amount of light that is used by the wavefront sensor is lost from the image. In practice, however, the image beam is split into two parts. The centroid of the image of one part is measured with a quad cell, and the tilt indicated by it is corrected with a steering mirror. The resulting tilt-corrected image of the other part with the residual aberration is corrected with a deformable mirror in a closed-loop manner. In the modal approach, there is no loss of light, but the rate or the bandwidth of correction can be slow due to its iterative nature, especially when turbulence is severe and a large number of modes must be corrected. Moreover, for imaging an extended object, wavefront sensing requires a point source in its vicinity, but the modal approach is applicable to the extended object itself. Adaptive optics has also been used in lucky imaging to achieve diffraction-limited image quality.[18]

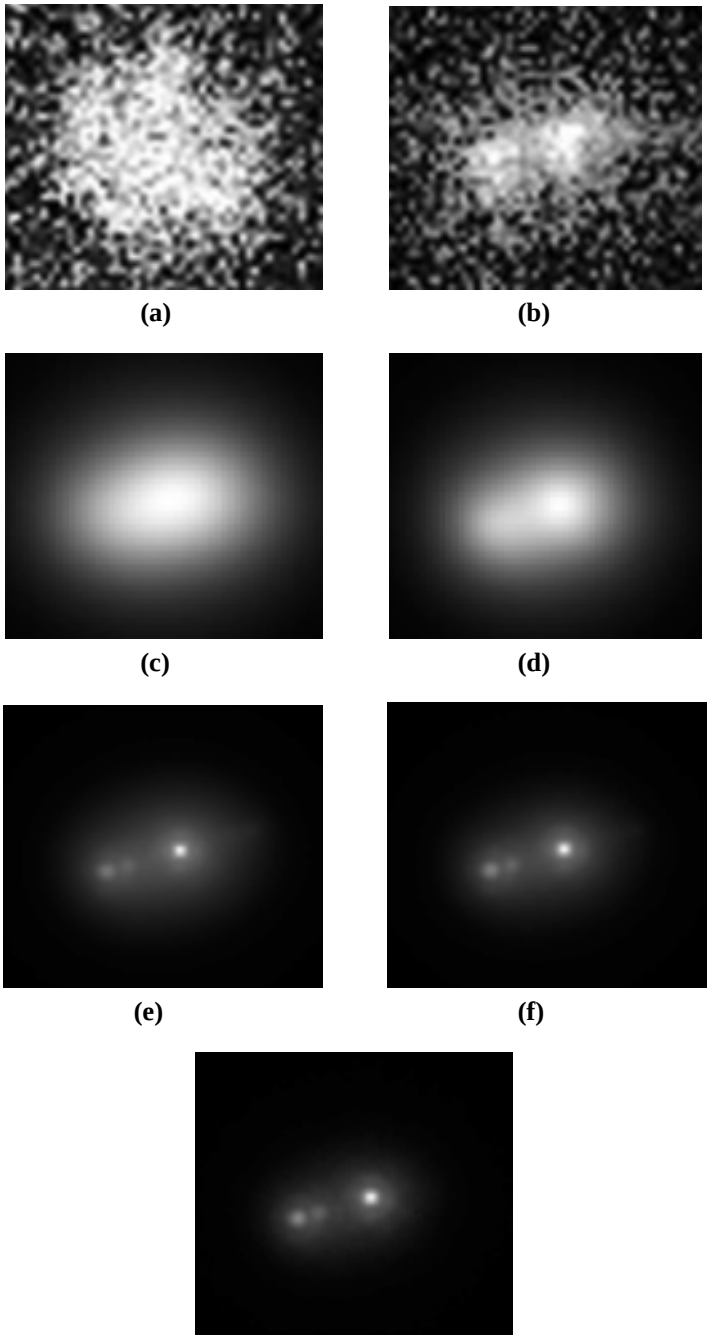


Figure 11-9. Lucky imaging. (a) Poor-quality SE star image, (b) corresponding good-quality image, (c) LE image obtained by adding all of the 50,000 SE images taken at a rate of 40/sec over a period of about 21 minutes, (d) LE image obtained by aligning the centroids of the SE images, (e) 50% of the selected SE images aligned by peak pixel value and added, (f) 10% of SE images added, and (g) 1% of SE images added.

Of course, adaptive optics can improve the image quality only if the object lies within an isoplanatic angle of turbulence. In the case of a ground-to-space laser illuminating a satellite, the angular travel (point-ahead angle) of the satellite during the round-trip time of the beam to the satellite must be less than the isoplanatic angle of turbulence.

11.4 FABRICATION ERRORS AND TOLERANCES

In Chapters 1–6, we have shown how to calculate the aberrations of various optical imaging systems. Although it was not pointed out explicitly, it was understood that the elements of a system had their *prescribed shapes*, i.e., the elements did not have any fabrication errors. The aberrations of a system thus calculated are referred to as its *design aberrations.* In practice, when the elements of a system are fabricated, their exact shapes *will* deviate however slightly from their prescribed shapes. These *fabrication* or *manufacturing errors* are generally referred to as their *surface* or *figure errors.* They are typically random in nature in that if an element is fabricated in large quantities, its errors will vary randomly from one sample to another. However, these errors have certain statistical properties that depend on the fabrication process. For example, the width (*correlation length*) of the polishing irregularities of an element depends on the size of the tool used to polish it. The figure errors of an element of a system contribute to its aberrations. For example, if θ and θ' are the angles of incidence and refraction of a ray incident on a refracting surface separating media of refractive indices n and n', and if δF is the deviation of the surface at the point of incidence of the ray along the surface normal at that point from the prescribed shape, the change in its optical path length is given by

$$\delta W = (n\cos\theta - n'\cos\theta')\delta F \quad . \tag{11-22}$$

Thus, under normal incidence, a plane-parallel plate of refractive index n introduces wavefront errors that are $(n-1)$ times its corresponding figure errors. In the case of a reflecting surface in air, Eq. (11-22) reduces to

$$\delta W = 2\cos\theta\,\delta F \quad . \tag{11-23}$$

Thus, a conservative estimate of the wavefront errors in this case is equal to twice the figure errors. The wavefront errors arising from *thermal distortions* of the elements and their *misalignments* and *spacing errors* may also be calculated by using Eqs. (11-22) and (11-23).

Because of the random nature of the figure errors, the total expected wavefront error of the system can be obtained by a square root of the sum of the variances of the wavefront errors contributed by its elements. Indeed this is how optical tolerances on the figure errors of the elements of a system are allocated. For example, if we are interested in a system Strehl ratio of 0.8 so that the total budget for the *standard deviation* of the wavefront errors is $\lambda/14$, the figure errors of the elements can be allocated equally or preferentially among them such that the root sum square of the standard deviations of their wavefront errors is $\lambda/14$.

As a numerical example, consider a three-mirror system. For simplicity, let the allowed standard deviation of the figure errors of each mirror be σ_F with a corresponding wavefront error of $2\sigma_F$. The total wavefront error contributed by the three mirrors is

$$\sigma_W^2 = 3(2\sigma_F)^2 \quad . \tag{11-24}$$

Accordingly, for a Strehl ratio of 0.8, the figure error tolerance is $\lambda/48$.

11. 5 SUMMARY

Previous chapters have considered systems with deterministic aberrations in the sense that they are known either by calculation or by measurement. This chapter has considered random aberrations in the sense that we know their statistics but not their detailed distributions. Examples of such aberrations are random image motion, those due to fabrication errors, and those introduced by atmospheric turbulence. Figures 11-1 and 11-2 show respectively the effect of a random transverse and longitudinal image motion on the Strehl ratio of an image as a function of the standard deviation of the motion. In the case of transverse image motion, the image moves up, down, and sideways in the image plane. However, in the case of longitudinal image motion, the image is defocused as it moves randomly along the optical axis. As expected, the Strehl ratio decreases monotonically as the image motion increases. However, the decrease in the case of transverse image motion is smaller for a larger value of the obscuration ratio due to its larger depth of focus. This effect is opposite to that of the transverse image motion, where the drop in Strehl ratio with increasing image motion increases as the obscuration increases due to the narrower central disc of its PSF.

The aberrations introduced by atmospheric turbulence not only reduce the Strehl ratio and broaden the image, but also break the image into speckles. Whereas the size of the image is determined by the coherence length of atmospheric turbulence, that of a speckle is determined by the pupil diameter. A large portion (87%) of the aberration is a random wavefront tilt whose effect can be avoided by taking short-exposure images and adding them with proper registration. This is done in lucky imaging by discarding bad images. The image quality can also be improved by using a steering mirror to overcome the wavefront tilt, and a deformable mirror to overcome the aberrations.

References

1. V. N. Mahajan, "Degradation of an image due to Gaussian image motion," *Appl. Opt.* **17**, 3329–3334 (1978).

2. V. N. Mahajan, "Degradation of an image due to Gaussian longitudinal motion," *Appl. Opt.* **46**, 3700–3705 (2007).

3. V. N. Mahajan and G.-M. Dai, "Imaging through atmospheric turbulence," in *Handbook of Optics*, 3rd ed., M. Bass, Ed., Chapter 4, Vol. V (McGraw-Hill, New York, 2009).

4. R. Fugate, "Adaptive optics," in *Handbook of Optics*, 3rd ed., M. Bass, Ed., Chapter 5, Vol. V (McGraw-Hill, New York, 2009)

5. J. W. Hardy, *Adaptive Optics for Astronomical Telescopes* (Oxford, New York, 1998).

6. R. K. Tyson, *Introduction to Adaptive Optics* (SPIE Press, Bellingham, WA, 1999).

7. D. Fried, "Optical resolution through a randomly inhomogeneous medium for very long and very short exposures," *J. Opt. Soc. Am.* **56**, 1372–1379 (1966).

8. D. Fried, "Evaluation of r_0 for propagation down through the atmosphere," *Appl. Opt.* **13**, 2620–2622 (1974); errata 1, *Appl. Opt.* **14**, 2567 (1975); errata 2, *Appl. Opt.* **16**, 549 (1977).

9. D. L. Walters and L. W. Bradford, "Measurement of r_0 and θ_0: two decades and 18 sites," *Appl. Opt*. **36**, 7876–7886 (1997).

10. V. N. Mahajan and B. K. C. Lum, "Imaging through atmospheric turbulence with annular pupils," *Appl. Opt.* **20**, 3233–3237 (1981).

11. G.-m Dai and V. N. Mahajan, "Zernike annular polynomials and atmospheric turbulence," *J. Opt. Soc. Am. A* **24**, 139–155 (2007).

12. R. J. Noll, "Zernike polynomials and atmospheric turbulence," *J. Opt. Soc. Am.* **66**, 207–211 (1976).

13. N. M. Law, C. D. Mackay, and J. E. Baldwin, "Lucky imaging: high angular resolution imaging in the visible from the ground," *Astron. & Astrophys.* **446**, 739–745 (2006).

14. C. Mackay, J. Baldwin, N. Law, and P. Warner, "High resolution imaging in the visible from ground without adaptive optics: New techniques and results," *Proc. SPIE* **5492**, 128–135 (2004) [doi: 10.1117/12.550443].

15. R. A. Miller and Buffington, “Real-time wavefront correction of atmospherically degraded telescopic images through image sharpening,” *J. Opt. Soc. Am*. **61**, 1200–1210 (1974).

16. A. Buffington, F. S. Crawford, R. A. Miller, A. J. Schwemin, and R. G. Smits, “Correction of atmospheric distortion with an image-sharpening telescope,” *J. Opt. Soc. Am*. **67**, 298–305 (1977).

17. V. N. Mahajan, J. Govignon, and R. J. Morgan, “Adaptive optics without wavefront sensors,” *Proc. SPIE* **228**, 63–69 (1980).

18. C. Mackay, N. Law, and T. D. Stayley, “Diffraction limited imaging in the visible from large ground-based telescopes: New methods for future instruments and telescopes,” *Proc. SPIE* **7014**, 7014C– 7014C-7 (2010) [doi: 10.1117/12.787439].

CHAPTER 12
Observation of Aberrations

12.1 INTRODUCTION

In this chapter, we describe briefly how the primary aberrations of an optical system can be observed. The emphasis of our discussion is on how to recognize a primary aberration and not on how to measure it precisely. Since the optical frequencies are very high ($10^{14} - 10^{15}$ Hz), optical wavefronts, aberrated or not, cannot be observed directly; optical detectors simply do not respond at these frequencies. We have seen in Chapter 8 that the image of a monochromatic point object formed by an aberrated system is characteristically different for a different aberration. Another and more powerful way to recognize an aberration is to form an *interferogram* by combining two parts of a light beam, one of which has been transmitted through the system.

12.2 PRIMARY ABERRATIONS

Consider an optical imaging system with a circular exit pupil of radius a. Letting (r, θ) be the polar coordinates of a point in the plane of its exit pupil, the functional form of the *primary phase aberrations* may be written

$$\Phi(\rho,\theta) = \begin{cases} A_s\rho^4 + B_d\rho^2, & \text{Spherical combined with defocus} \quad (12\text{-}1) \\ A_c\rho^3\cos\theta + B_t\rho\cos\theta, & \text{Coma combined with tilt} \quad (12\text{-}2) \\ A_a\rho^2\cos^2\theta + B_d\rho^2, & \text{Astigmatism combined with defocus} \quad (12\text{-}3) \\ A_d\rho^2, & \text{Field curvature} \quad (12\text{-}4) \\ A_t\rho\cos\theta, & \text{Distortion,} \quad (12\text{-}5) \end{cases}$$

where A_i or B_i is a peak aberration coefficient representing the maximum value of the corresponding aberration across the pupil, and $\rho = r/a$ is a normalized radial variable. When $\Phi(\rho,\theta) = 0$ for a certain point object, the wavefront passing through the center of the exit pupil is spherical centered at its Gaussian image point. Let its radius of curvature be R. For an aberrated system, $\Phi(\rho,\theta)$ represents the optical deviation of the wavefront at a point (ρ,θ) from being spherical.

In Eq. (12-1), when $B_d = 0$, the aberration is *spherical*. Nonzero B_d implies that the aberration is combined with defocus; i.e., the aberration is not with respect to a reference sphere centered at the Gaussian image point but with respect to another sphere centered at a distance z from the plane of the exit pupil according to Eq. (8-6). As discussed in Chapter 7, the reference sphere is centered at the marginal image point, center of the circle of least confusion, and the point midway between the marginal and Gaussian image points when $B_d/A_s = -2$, -1.5, and -1, respectively. The midway point corresponds to minimum variance of the aberration and, therefore, to maximum Strehl ratio (for small aberrations), as may be seen by comparing the aberration thus obtained with the Zernike circle polynomial $Z_4^0(\rho)$.

In Eq. (12-2), when $B_t = 0$, the aberration is coma. Nonzero B_t implies that the aberration is combined with tilt, or that it is with respect to a reference sphere centered at a point $(2FB_t, 0)$ in the image plane, where F is the focal ratio or the f-number of the image-forming light cone. The variance of the aberration is minimum when $B_t/A_c = -2/3$, as in the Zernike circle polynomial $Z_3^1(\rho, \theta)$.

In Eq. (12-3), when $B_d = 0$, the aberration is *astigmatism.* Nonzero A_d implies that it is combined with defocus. The variance of the aberration is minimum when $B_d/A_a = -1/2$, as in the Zernike circle polynomial $Z_2^2(\rho, \theta)$. When $B_d/A_a = 0$ or -1, we obtain the so-called tangential and sagittal images of a point object. Equations (12-4) and (12-5) represent defocus or field curvature and tilt or distortion aberrations, respectively. Figure 12-1 shows isometric plots of the various aberrations.

12.3 INTERFEROGRAMS

There are a variety of interferometers that are used for detecting and measuring aberrations of optical systems.[1] Figure 12-2 illustrates schematically a *Twyman-Green interferometer* in which a collimated laser beam is divided into two parts by a beam splitter *BS*. One part, called the *test beam*, is incident on the system under test, indicated by the lens *L*, and the other, called the *reference beam*, is incident on a plane mirror M_1. The focus *F* of the lens system lies at the center of curvature *C* of a spherical mirror M_2. As the angle of the incident light is changed to study the off-axis aberrations of the system, the mirror is tilted so that its center of curvature lies at the current focus of the beam. In this arrangement the mirror does not introduce any aberration since it is forming the image of an object lying at its center of curvature (see Section 4.2).

The two reflected beams interfere in the region of their overlap. Lens L' is used to observe the interference pattern on a screen S placed in a plane containing the image of L formed by L'. A record of the interference pattern is called an interferogram. Note that since the test beam goes through the lens system *L* twice, its aberration is twice that of the system.

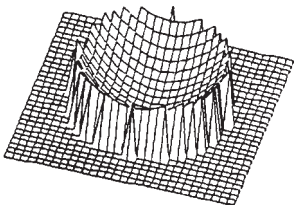
If the reference beam has a uniform phase and the test beam has a phase distribution $\Phi(x, y)$, and if their amplitudes are equal to each other, the irradiance distribution of their interference pattern is given by

$$\begin{aligned} I(x,y) &= I_0\left|1 + \exp\left[i\Phi(x,y)\right]\right|^2 \\ &= 2I_0\left\{1 + \cos\left[\Phi(x,y)\right]\right\} \quad , \end{aligned} \tag{12-6}$$

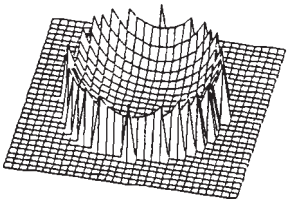
where I_0 is the irradiance when only one beam is present. The irradiance has a maximum value equal to $4I_0$ at those points for which

$$\Phi(x, y) = 2\pi n \tag{12-7a}$$

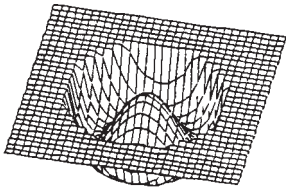
and a minimum value equal to zero wherever



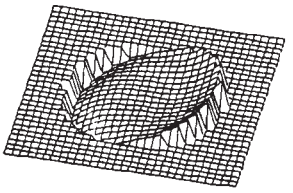
Defocus: ρ^2



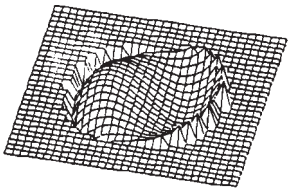
Spher ca : ρ^4



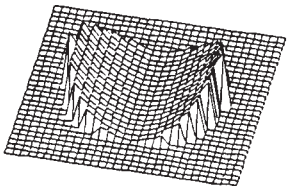
Ba anced Spher ca : $\rho^4 - \rho^2$



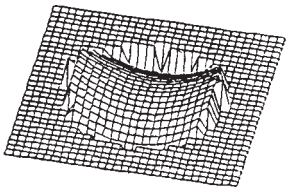
Coma: $\rho^3 \cos\theta$



Ba anced Coma: $(\rho^3 - \frac{2}{3}\ \rho)\cos\theta$



Ast gmat sm: $\rho^2 \cos^2\theta$



Ba anced Ast gmat sm: $\rho^2 \cos^2\theta - \frac{1}{2}\rho^2$

Figure 12-1. Isometric plot of primary aberrations representing the difference between an ideal wavefront (typically, spherical) and an actual wavefront.

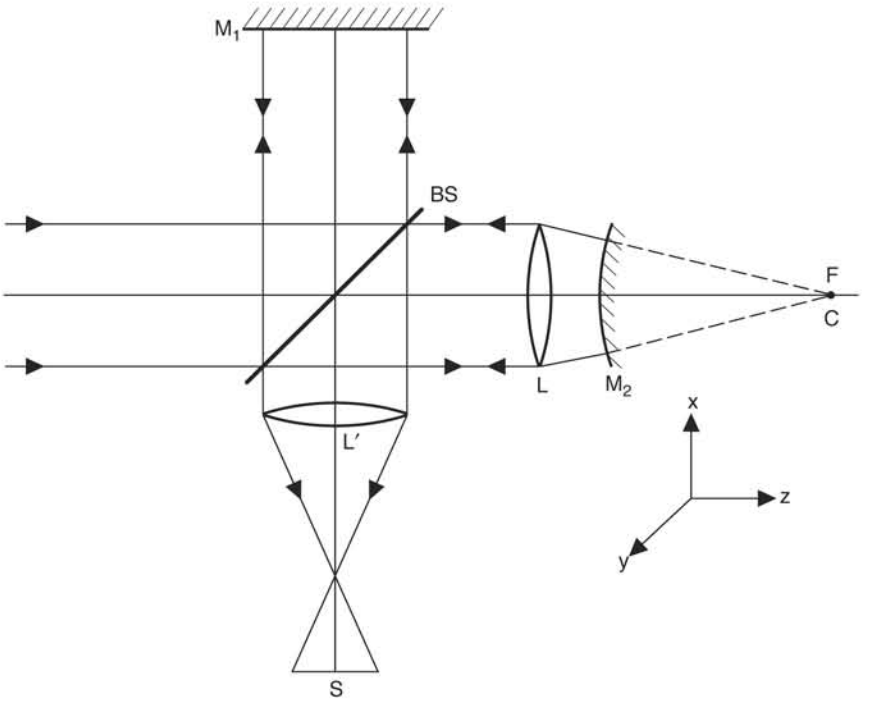


Figure 12-2. Twyman-Green interferometer for testing a lens system L. F is the image-space focal point of L, and C is the center of curvature of a spherical mirror M_2. The interfering beams are focused by a lens L', and the interference pattern is observed on a screen S.

$$\Phi(x, y) = 2\pi(n + 1/2) \quad , \tag{12-7b}$$

where n is a positive or a negative integer, including zero. Each fringe in the interference pattern represents a certain value of n, which in turn corresponds to the locus of (x, y) points with phase aberration given by Eq. (12-7a) for a bright fringe and Eq. (12-7b) for a dark fringe. If the test beam is aberration free $[\Phi(x, y) = 0]$, then the interference pattern has a uniform irradiance of $4I_0$.

Figure 12-3 shows the interferograms when the lens system L under test suffers from 3λ of a primary aberration, corresponding to 6λ of an aberration of the interfering test beam. In our discussion, we give the value of an aberration coefficient in wavelength units, rather than in radians, as is customary in optics. For defocus and spherical aberration, the interference pattern consists of concentric circular interference fringes. The fringe spacing depends on the type of the aberration. Figure 12-3a shows the interferogram obtained when the system is aberration free but it is misfocused, i.e., when its focus F lies to the left or the right of the center of curvature C of the spherical mirror M_2 by an amount corresponding to 3λ of defocus aberration. [See Eqs. (1-3a) and (1-3b) for a relationship between the longitudinal defocus, i.e., the axial spacing between F and C, and the peak defocus aberration B_d, which is 3λ in our example.] Figure 12-3b

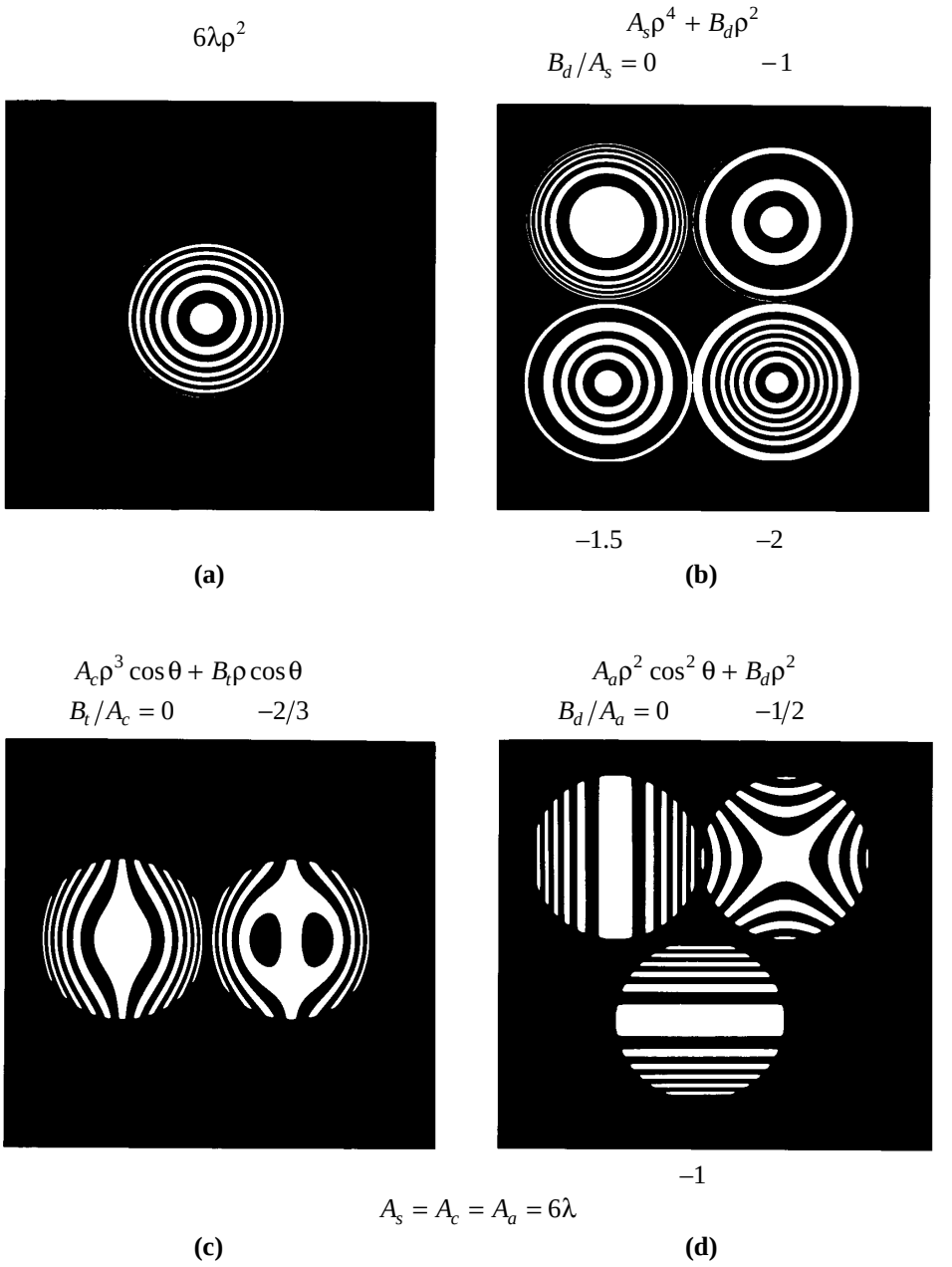


Figure 12-3. Interferograms of primary aberrations: (a) defocus, (b) spherical combined with defocus, (c) coma combined with tilt, (d) astigmatism combined with defocus. The aberrations in the interferograms are twice their corresponding values in the system under test because the test beam goes through the system twice.

shows the interferograms obtained when the system has 3λ of spherical aberration (i.e., $A_s = 3\lambda$) and a certain amount of defocus. The case $B_d = 0$ (i.e., F and C coincident) represents such a system with an image of a certain object being observed in its Gaussian image plane. Similarly, the interferogram obtained for $B_d/A_s = -2$ represents the system

when the image is observed in its marginal image plane. For a system with positive spherical aberration, its marginal focus lies farther from its exit pupil than its paraxial focus (see Figure 7-1). Hence, this interferogram is obtained when the points F and C are separated from each other axially, according to Eq. (1-3d), by $-48\lambda F^2$, i.e., when F lies to the left of C by $48\lambda F^2$. The other two interferograms, $B_d = -A_s$ and $B_d = -1.5\,A_s$, represent the system when the image is observed in the minimum-aberration-variance plane (or maximum Strehl ratio for small values of A_s) and the circle-of-least-confusion plane, respectively.

Figure 12-3c shows the interferograms obtained when light is incident at a certain angle from the axis of the system such that it suffers from 3λ of coma. The fringes in this case are cubic curves. The case $B_t = 0$ corresponds to two parallel interfering beams (F and C are coincident in this case). The case $B_t = -2A_c/3$ represents the system corresponding to minimum aberration variance. A tilt aberration with a peak value of B_t may be obtained by transversally displacing C from F by $(-2FB_t, 0)$ so that C lies at the diffraction focus of the comatic diffraction pattern of the system (see Section 8.3.3 for a discussion of the diffraction focus). It may also be obtained by tilting the plane mirror M_1 by an angle B_t/a, where a is the radius of the test beam [see Eq. (1-5c) and note the factors of 2 because of the reflection of the reference beam by mirror M_1 and doubling of the system aberration in the test beam].

Figure 12-3d shows the interferograms obtained when the system suffers from 3λ of astigmatism. When $B_d = 0$ or $-A_a$, representing the system with an image being observed in a plane containing one or the other astigmatic focal line, respectively, we obtain an interferogram with straight line fringes, since the aberration depends on either x or y (but not both). However, the fringe spacing is not uniform. When $B_d = -A_a/2$, the fringe pattern consists of rectangular hyperbolas. If the system under test is aberration free, but the two interfering beams are tilted with respect to each other, representing a wavefront tilt error, we obtain straight line fringes that are uniformly spaced. The fringe spacing is inversely proportional to the tilt angle.

So far we have discussed interferograms of primary aberrations when only one of them is present. These interferograms are relatively simple and the aberration type may be recognized from the shape of the fringes. It should be evident that a general aberration consisting of a mixture of these aberrations and/or others will yield a much more complex interferogram. As an example of a general aberration, Figure 12-4a shows a possible aberration introduced by atmospheric turbulence, as in ground-based astronomical observations. It corresponds to $D/r_0 = 10$, as discussed in Section 11.3.3. On the average, the standard deviation of the instantaneous aberration introduced is given by $[0.134\,(D/r_0)^{5/3}]^{1/2}$, which is 2.5 radians or 0.4λ for $D/r_0 = 10$. The interferogram for this aberration is shown in Figure 12-4b. When 25λ of tilt are added to the aberration, the interferogram appears as in Figure 12-4c. Doubling of the aberration, as in a Twyman–Green interferometer, is not considered in Figure 12-4.

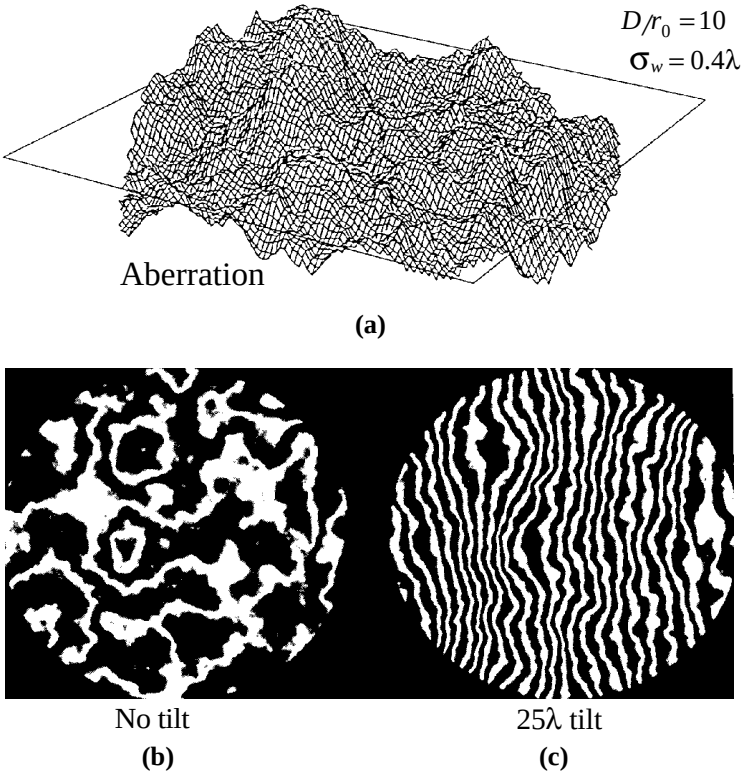


Figure 12-4. Aberration introduced by atmospheric turbulence corresponding to $D/r_0 = 10$. (a) Isometric plot of the aberration (b) Aberration interferogram. The standard deviation of the tilt-free aberration introduced by turbulence is 0.4λ. (c) Interferogram with 25λ of tilt.

12.4 SUMMARY

Because of the high optical frequencies and the comparatively slow time response of the photodetectors, the aberrations or phase errors of a light wave cannot be observed or measured directly. They are determined by forming an interferogram, where two parts of a light beam are combined after one part has passed through the system under test. In this chapter, we have shown isometric plots of the primary aberrations (see Figure 12-1), representing, for example, the shape of the surface of a deformable mirror for a certain aberration, and the interferograms, as may be seen in optical testing (see Figure 12-3). An interferogram for a random aberration is also shown (see Figure 12-4). The purpose is to acquaint the reader with what may be seen in practice when working in a laboratory.

References

1. D. Malacara, Ed., *Optical Shop Testing*, 3rd ed., Wiley, New York (2007).

Bibliography

M. Born and E. Wolf, *Principles of Optics,* 7th ed., Cambridge University Press, New York, 1999.

A. E. Conrady, *Applied Optics and Optical Design,* Parts I and II, Oxford, London, 1929 (Reprinted by Dover, New York, 1957).

H. H. Hopkins, *Wave Theory of Aberrations,* Oxford, London, 1950.

F. A. Jenkins and H. E. White, *Fundamentals of Optics,* 4th ed., McGraw-Hill, New York, 1976.

M. V. Klein and T. E. Furtak, *Optics,* 2nd ed., Wiley, New York, 1986.

E. H. Linfoot, *Recent Advances in Optics,* Clarendon, Oxford, 1955.

V. N. Mahajan, *Optical Imaging and Aberrations,* Part I: *Ray Geometrical Optics*, SPIE Press, Bellingham, WA, 1998; Second Printing, 2001.

V. N. Mahajan, *Optical Imaging and Aberrations,* Part II: *Wave Diffraction Optics*, SPIE Press, Bellingham, WA, Second Edition, 2011.

L. C. Martin and W. T. Welford, *Technical Optics,* Vol. I, 2nd ed., Pitman, London, 1966.

P. Mouroulis and J. Macdonald, *Geometrical Optics and Optical Design*, Oxford, New York, 1997.

E. L. O'Neill, *Introduction to Statistical Optics,* Addison-Wesley, Reading, Massachusetts, 1963.

D. J. Schroeder, *Astronomical Optics,* Academic Press, New York, 1987.

W. J. Smith, *Modern Optical Engineering,* McGraw-Hill, New York, 1966.

W. T. Welford, *Aberrations of the Symmetrical Optical System,* Academic Press, New York, 1974.

References for Additional Reading

These references are the author's collection as the editor of Milestone Series 74 entitled *Effects of Aberrations in Optical Imaging*, published by SPIE Press in 1993.

Section One: Aberration-Free Systems

1. G. B. Airy, "On the diffraction of an object-glass with circular aperture," *Trans. Cambridge Philos. Soc*. **5**, 283–291 (1835).

2. Lord Rayleigh, "On images formed without reflection or refraction," *Philos. Mag*. **5**, 214–218 (1881).

3. L. Beiser, "Perspective rendering of the field intensity diffracted at a circular aperture," *Appl. Opt.* **5**, 869–870 (1966).

4. J. E. Harvey, "Fourier treatment of near-field scalar diffraction theory," *Am. J. Phys*. **47**, 974–980 (1979).

5. R. Barakat and A. Houston, "Reciprocity relations between the transfer function and total illuminance. I," *J. Opt. Soc. Am.* **53**, 1244–1249 (1963).

6. V. N. Mahajan, "Asymptotic behavior of diffraction images," *Canadian J. Phys*. **57**, 1426–1431 (1979).

7. P. P. Clark, J. W. Howard, and E. R. Freniere, "Asymptotic approximation to the encircled energy function for arbitrary aperture shapes," *Appl. Opt*. **23**, 353–357 (1983).

8. I. Ogura, "Asymptotic behavior of the response function of optical systems," *J. Opt. Soc. Am.* **48**, 579–580 (1958).

9. B. Tatian, "Asymptotic expansions for correcting truncation error in transfer-function calculations," *J. Opt. Soc. Am*. **61**, 1214–1224 (1971).

10. W. S. Kovach, "Energy distribution in the PSF for an arbitrary passband," *Appl. Opt.* **13**, 1769–1771 (1974).

11. H. S. Dhadwal and J. Hantgan, "Generalized point spread function for a diffraction-limited aberration-free imaging system under polychromatic illumination," *Opt. Eng*. **28**, 1237–1240 (1989).

Section Two: Defocused Systems

12. E. Wolf, "Light distribution near focus in an error-free diffraction image," *Proc. Royal Soc. A* **204**, 533–548 (1951).

13. J. C. Dainty, "The image of a point for an aberration-free lens with a circular pupil," *Opt. Comm.* **1**, 176–178 (1969).

14. R. E. Stephens and L. E. Sutton, "Diffraction image of a point in the focal plane and several out-of-focus planes," *J. Opt. Soc. Am.* **58**, 1001–1002 (1968).

15. P. A. Stokseth, "Properties of a defocused optical system," *J. Opt. Soc. Am.* **59**, 1314–1321 (1969).

16. D. S. Burch, "Fresnel diffraction by a circular aperture," *Am. J. Phys*. **53**, 255–260 (1985).

17. H. Osterberg and L. W. Smith, "Defocusing images to increase resolution," *Science* **134**, 1193–1196 (1961).

18. T. S. McKechnie, "The effect of defocus on the resolution of two points," *Optica Acta* **20**, 253–262 (1973).

19. T. S. McKechnie, "The effect of condenser obstruction on the two-point resolution of a microscope," *Optica Acta* **19**, 729–737 (1972).

20. D. K. Cook and G. D. Mountain, "The effect of phase angle on the resolution of two coherently illuminated points," *Optical and Quan. Elec*. **10**, 179–180 (1978).

21 A. Arimoto, "Intensity distribution of aberration-free diffraction patterns due to circular apertures in large F-number optical systems," *Optica Acta* **23**, 245–250 (1976).

22. Y. Li and E. Wolf, "Focal shifts in diffracted converging spherical waves," Opt. Comm. **39**, 211–215 (1981).

23. Y. Li, "Dependence of the focal shift on Fresnel number and f number," *J. Opt. Soc. Am.* **72**, 770–774 (1982).

24. V. N. Mahajan, "Axial irradiance and optimum focusing of laser beams," *Appl. Opt*. **22**, 3042–3053 (1983).

25. H. H. Hopkins, "The frequency response of a defocused optical system," *Proc. Royal Soc*. *A* **231**, 91–203 (1955).

26. W. H. Steel, "The defocused image of sinusoidal gratings," *Optica Acta* **3**, 65–74 (1956).

27. L. Levi and R. H. Austing, "Tables of the modulation transfer function of a defocused perfect lens," *Appl. Opt*. **7**, 967–974 (1968).

Section Three: Strehl and Hopkins Ratios

28. Lord Rayleigh, “Investigations in optics, with special reference to the spectroscope. Sec. 4: Influence of aberrations,” *Philos. Mag.* **8**, 403–411 (1879).

29. K. Strehl, “Ueber Luftschlieren und Zonenfehler,” *Zeitschrift fur instrumentenkunde* **22**, 213–217 (1902).

30. A. Maréchal, “Etude des effets combines de la diffraction et des aberrations geometriques sur l'image d'un point lumineux,” *Revue d'Optique* **26**, 257–277 (1947).

31. H. H. Hopkins, “The use of diffraction-based criteria of image quality in automatic optical design,” *Optica Acta* **13**, 343–369 (1966).

32. W. H. Steel, “The problem of optical tolerances for systems with absorption,” *Appl. Opt.* **8**, 2297–2299 (1969).

33. V. N. Mahajan, “Strehl ratio for primary aberrations: some analytical results for circular and annular pupils,” *J. Opt. Soc. Am.* **72**, 1258–1266 (1982).

34. V. N. Mahajan, “Strehl ratio for primary aberrations in terms of their aberration variance,” *J. Opt. Soc. Am.* **73**, 860–861 (1983).

35. G. Martial, “Strehl ratio and aberration balancing,” *J. Opt. Soc. Am. A* **8**, 164–170 (1991).

36. W. B. King, “Dependence of the Strehl ratio on the magnitude of the variance of the wave aberration,” *J. Opt. Soc. Am.* **58**, 655–661 (1968).

37. J. J. H. Wang, “Tolerance conditions for aberrations,” *J. Opt. Soc. Am*. **62**, 598–599 (1972).

38. H. H. Hopkins, “Geometrical-optical treatment of frequency response,” *Proc. Phys. Soc*. *B* **70**, 449–470 (1957).

39. H. H. Hopkins, “The aberration permissible in optical systems,” *Proc. Phys. Soc. B* **70**, 449–470 (1957).

40. W. B. King, “Correlation between the relative modulation function and the magnitude of the variance of the wave-aberration difference function,” *J. Opt. Soc. Am.* **59**, 285–290 (1969).

41. S. Szapiel, “Hopkins variance formula extended to low relative modulations,” *Optica Acta* **33**, 981–999 (1986).

Section Four: Aberration Balancing

42. B. R. A. Nijboer, "The diffraction theory of optical aberrations. Part I: General discussion of the geometrical aberrations," *Physica* **10**, 679–692 (1943).

43. B. R. A. Nijboer, "The diffraction theory of optical aberrations. Part II: Diffraction pattern in the presence of small aberrations," *Physica* **13**, 605–620 (1947).

44. K. Nienhuis and B. R. A. Nijboer, "The diffraction theory of optical aberrations. Part III: General formulae for small aberrations: experimental verification of the theoretical results," *Physica* **14**, 590–608 (1949).

45. B. Tatian, "Aberration balancing in rotationally symmetric lenses," *J. Opt. Soc. Am.* **64**, 1083–1091 (1974).

46. V. N. Mahajan, "Zernike annular polynomials for imaging systems with annular pupils," *J. Opt. Soc. Am.* **71**, 75–85; 1408 (1981).

47. V. N. Mahajan, "Zernike annular polynomials for imaging systems with annular pupils," *J. Opt. Soc. Am. A* **1**, 685 (1984).

48. S. Szapiel, "Aberration-balancing technique for radially symmetric amplitude distributions: a generalization of the Maréchal approach," *J. Opt. Soc. Am.* **72**, 947–956 (1982).

49. K. Pietraszkiewicz, "Determination of the optimal reference sphere," *J. Opt. Soc. Am*. **69**, 1045–1046 (1979).

50. A. Magiera, K. Pietraszkiewicz, "Position of the optimal reference sphere for apodized optical systems," *Optik* **58**, 85–91 (1981).

Section Five: Zernike Polynomials

51. S. N. Bezdid'ko, "The use of Zernike polynomials in optics," *Soviet J. Opt. Tech.* **41**, 425–429 (1974).

52. S. N. Bezdid'ko, "Calculation of the Strehl coefficient and determination of the best-focus plane in the case of polychromatic light," *Soviet J. Opt. Tech*. **42**, 514–516 **(**1975).

53. S. N. Bezdid'ko, "Determination of the Zernike polynomial expansion coefficients of the wave aberration," *Soviet J. Opt. Tech.* **42**, 426–427 (1975).

54. S. N. Bezdid'ko, "Numerical method of calculating the Strehl coefficient using Zernike polynomials," *Soviet J. Opt. Tech.* **43**, 222–225 (1977).

55. S. N. Bezdid'ko, "Use of orthogonal polynomials in the case of optical systems with annular pupils," *Opt. Spectroscopy* **43**, 2000–2003 (1977).

56. G. Conforti, "Zernike aberration coefficients from Seidel and higher-order power-series coefficients," *Opt. Lett*. **8**, 407–408 (1983).

57. R. K. Tyson, "Conversion of Zernike aberration coefficients to Seidel and higher-order power-series aberration coefficients," *Opt. Lett*. **7**, 262–264 (1982).

58. R. J. Noll, "Zernike polynomials and atmospheric turbulence," *J. Opt. Soc. Am.* **66,** 207–2111 (1976).

59. J. Y. Wang and D. E. Silva, "Wave-front interpretation with Zernike polynomials," *Appl. Opt.* **19**, 1510–1518 (1980).

Section Six: Aberrated Systems

60. R. Barakat, "Total illumination in a diffraction image containing spherical aberration," *J. Opt. Soc. Am.* **51**, 152–167 (1961).

61. R. Barakat and A. Houston, "Diffraction effects of coma," *J. Opt. Soc. Am.* **54**, 1084–1088 (1964).

62. V. N. Mahajan, "Aberrated point-spread functions for rotationally symmetric aberrations," *Appl. Opt*. **22**, 3035–3141 (1983).

63. S. Szapiel, "Aberration-variance-based formula for calculating point-spread functions: rotationally symmetric aberrations," *Appl. Opt*. **25**, 244–251 (1986).

64. V. N. Mahajan, "Line of sight of an aberrated optical system," *J. Opt. Soc. Am. A* **2**, 833–846 (1985).

65. H. H. Hopkins, "Image shift, phase distortion and the optical transfer function," *Optica Acta* **31**, 345–368 (1984).

Section Seven: Annular Apertures

66. G. B. Airy, "On the diffraction of an annular aperture," *Philos. Mag*. **18**, 1–10, 132–133 (1841).

67. H. F. A. Tschunko, "Imaging performance of annular apertures," *Appl. Opt*. **18**, 3770–3774 (1974).

68. A. T. Young, "Photometric error analysis. X: Encircled energy (total illuminance) calculations for annular apertures," *Appl. Opt*. **9**, 1874–1888 (1970).

69. I. L. Goldberg and A. W. McCulloch, "Annular aperture diffracted energy distribution for an extended source," *Appl. Opt*. **8**, 1451–1458 (1969).

70. V. N. Mahajan, "Included power for obscured circular pupils," *Appl. Opt*. **17**, 964–968 (1978).

71. J. J. Stamnes, H. Heier, and S. Ljunggren, "Encircled energy for systems with centrally obscured circular pupils," *Appl. Opt*. **21**, 1628–1633 (1982).

72. E. H. Linfoot and E. Wolf, "Diffraction images in systems with an annular aperture," *Proc. Phys. Soc*. *B* **66**, 145–149 (1953).

73. T. Asakura and H. Mishina, "Irradiance distribution in the diffraction patterns of an annular aperture with spherical aberration and coma," *Japanese J. Appl. Phys*. **7**, 751–758 (1968).

74. E. L. O'Neill, "Transfer function for an annular aperture," *J. Opt. Soc. Am*. **46**, 285–288 (1956).

Section Eight: Gaussian Beams

75. A. L. Buck, "The radiation pattern of a truncated Gaussian aperture distribution," *Proc. IEEE* **55**, 448–450 (1967).

76. J. P. Campbell and L. G. DeShazer, "Near fields of truncated-Gaussian apertures," *J. Opt. Soc. Am*. **59**, 1427–1429 (1969).

77. G. O. Olaofe, "Diffraction by Gaussian apertures," *J. Opt. Soc. Am* **60**, 1654–1657 (1970).

78. R. G. Schell and G. Tyras, "Irradiance from an aperture with truncated-Gaussian field distribution," *J. Opt. Soc. Am.* **61**, 31–35 (1971).

79. V. P. Nayyar and N. K. Verma, "Diffraction by truncated-Gaussian annular apertures," *J. Opt*. [Paris] **9**, 307–310 (1978).

80. D. A. Holmes, J. E. Korka, P. V. Avizonis, "Parametric study of apertured focused Gaussian beams," *Appl. Opt*. **11**, 565–574 (1972).

81. Y. Li and E. Wolf, "Focal shift in focused truncated Gaussian beams," *Opt. Comm*. **42**, 151–156 (1982).

82. K. Tanaka, N. Saga, and K. Hauchi, "Focusing of a Gaussian beam through a finite aperture lens," *Appl. Opt.* **24**, 1098–1101 (1985).

83. D. D. Lowenthal, "Maréchal intensity criteria modified for Gaussian beams," *Appl. Opt*. **13**, 2126–2133, 2774 (1974).

84. D. D. Lowenthal, "Far-field diffraction patterns for gaussian beams in the presence of small spherical aberrations," *J. Opt. Soc. Am.* **65**, 853–855 (1975).

85. R. Herloski, "Strehl ratio for untruncated aberrated Gaussian beams," *J. Opt. Soc. Am. A* **2**, 1027–1030 (1985).

86. V. N. Mahajan, "Uniform versus Gaussian beams: a comparison of the effects of diffraction, obscuration, and aberrations," *J. Opt. Soc. Am. A* **3**, 470–485 (1986).

87. S. C. Biswas and J.-E. Villeneuve, "Diffraction of a laser beam by a circular aperture under the combined effect of three primary aberrations," *Appl. Opt.* **25**, 2221–2232 (1986).

Section Nine: Random Aberrations

88. D. L. Fried, "Optical resolution through a randomly inhomogeneous medium for very long and very short exposures," *J. Opt. Soc. Am.* **56**, 1372–1379 (1966).

89. D. L. Fried, "Optical heterodyne detection of an atmospherically distorted signal wave front," *Proc. IEEE* **55**, 57–67 (1967).

90. J. Y. Wang, "Optical resolution through a turbulent medium with adaptive phase compensations," *J. Opt. Soc. Am.* **67**, 383–390 (1977).

91. V. N. Mahajan and B. K. C. Lum, "Imaging through atmospheric turbulence with annular pupils," *Appl. Opt.* **20**, 3233–3237 (1981).

92. V. N. Mahajan, "Degradation of an image due to Gaussian motion," *Appl. Opt*. **17**, 3329–3334 (1978).

Section Ten: Coherent Systems

93. W. H. Steel, "Effects of small aberrations on the images of partially coherent objects," *J. Opt. Soc. Am.* **47**, 405–413 (1957).

94. R. Barakat, "Diffraction images of coherently illuminated objects in the presence of aberrations," *Optica Acta* **17**, 337–347 (1969).

95. R. Barakat," Partially coherent imagery in the presence of aberrations," *Optica Acta* **16**, 205–223 (1970).

96. J. P. Mills and B. J. Thompson, "Effect of aberrations and apodization on the performance of coherent optical systems. I. The amplitude impulse response," *J. Opt. Soc. Am. A* **3**, 694–703 (1986).

97. J. P. Mills and B. J. Thompson, "Effect of aberrations and apodization on the performance of coherent optical systems. II. Imaging," *J. Opt. Soc. Am. A* **3**, 704–716 (1986).

98. D. B. Allred and J. P. Mills, "Effect of aberrations and apodization on the performance of coherent optical systems. 3: The near field," *Appl. Opt*. **28**, 673–681 (1989).

Index

The terms annular, circular, and Gaussian refer to a pupil.

S

T

V

W

Z

Virendra N. Mahajan was born in Vihari, Pakistan, and educated in India and the United States. He received his Ph.D. degree in optical sciences from the College of Optical Sciences, University of Arizona. He spent nine years at the Charles Stark Draper Laboratory in Cambridge, Massachusetts, where he worked on space optical systems. Since 1983, he has been at The Aerospace Corporation in El Segundo, California, where he is a distinguished scientist working on space-based surveillance systems. Dr. Mahajan is an adjunct professor in the College of Optical Sciences at the University of Arizona, and the Department of Optics and Photonics at the National Central University in Taiwan, where he teaches graduate courses on imaging and aberrations. He also teaches short courses on aberrations at meetings of the Optical Society of America and SPIE. He has published numerous papers on diffraction, aberrations, adaptive optics, and acousto-optics. He is a fellow of OSA, SPIE, and the Optical Society of India. He is an associate editor of OSA's 3rd edition of the Handbook of Optics, and a recipient of SPIE's Conrady award. He has served as a Topical Editor of Optics Letters, chairman of OSA's Astronomical, Aeronautical, and Space Optics technical group, and a member of several committees of both OSA and SPIE. Dr. Mahajan is the author of *Optical Imaging and Aberrations*, Part I: *Ray Geometrical Optics* (1998), *Optical Imaging and Aberrations*, Part II: *Wave Diffraction Optics*, *Second Edition* (2011), and editor of *Selected Papers on Effects of Aberrations in Optical Imaging* (1994), all published by SPIE Press.