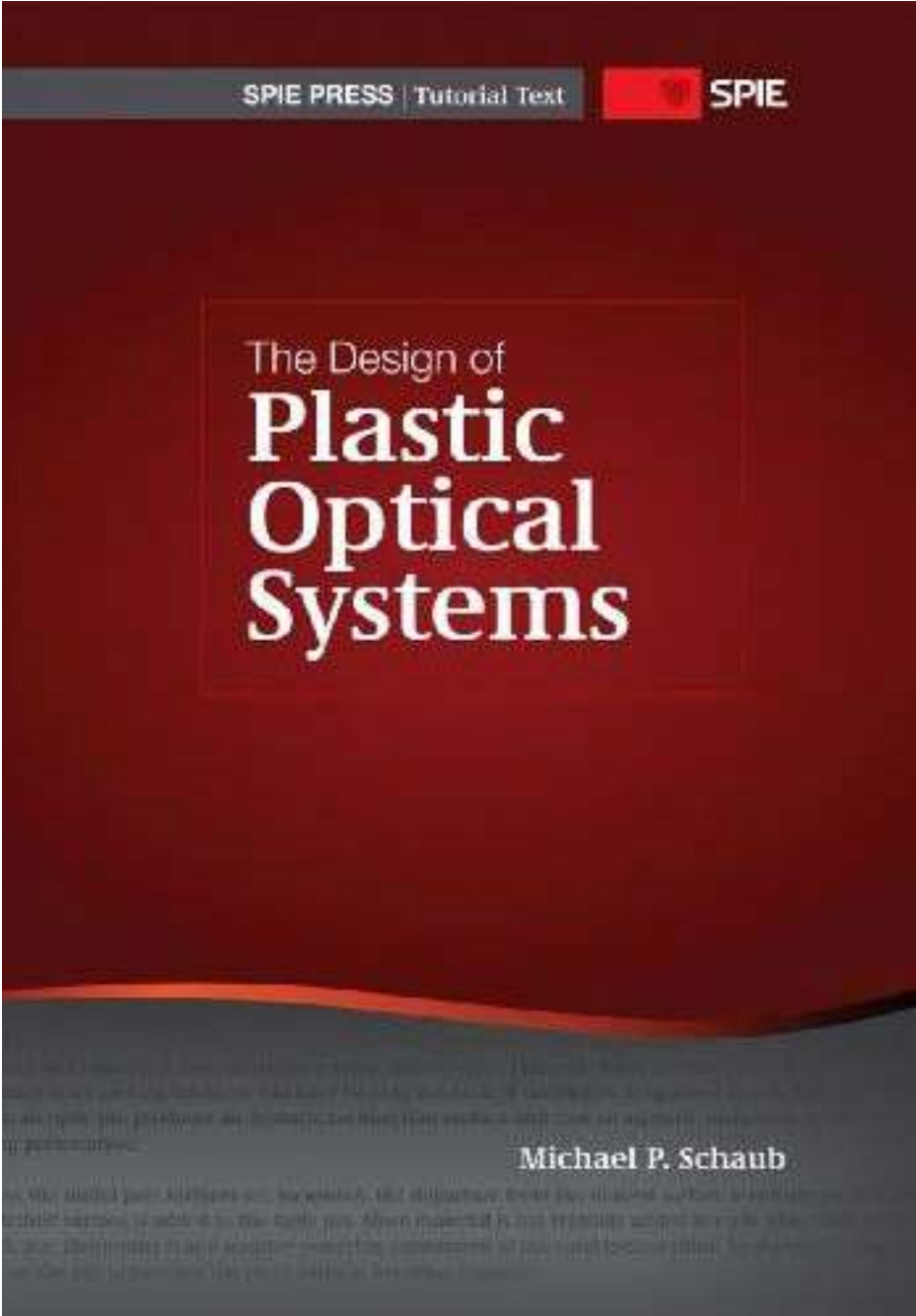


1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45



Petitioner has added vertical line numbering to all pages, as well as romanettes to pages (i)-(ii). Otherwise, it leaves the original page numbering.

MOTOROLA MOBILITY
EXHIBIT 1007

Library of Congress Cataloging-in-Publication Data

Schaub, Michael P.

The design of plastic optical systems / Michael P. Schaub.

p. cm. -- (Tutorial texts series ; TT 80)

Includes bibliographical references and index.

ISBN 978-0-8194-7240-3

1. Plastic lenses. 2. Optical instruments--Design and construction. 3. Plastics--Optical properties. 4. Optical materials. I. Title.

TS517.5.P5S33 2009

681'.4--dc22

2009028012

Published by

SPIE

P.O. Box 10

Bellingham, Washington 98227-0010 USA

Phone: 360.676.3290

Fax: 360.647.1445

E-mail: Books@spie.org

www.spie.org

Copyright © 2009 Society of Photo-Optical Instrumentation Engineers

All rights reserved. No part of this publication may be reproduced or distributed in any form or by any means without written permission of the publisher.

The content of this book reflects the thought of the author(s). Every effort has been made to publish reliable and accurate information herein, but the publisher is not responsible for the validity of the information or for any outcomes resulting from reliance thereon.

Printed in the United States of America.



Contents

Preface	ix
Acknowledgments.....	xi
Chapter 1 Introduction	1
1.1 Background	1
1.2 When Are Plastic Optics Appropriate?	5
Chapter 2 Optical Plastics	15
2.1 Plastic Versus Glass Maps	15
2.2 Material Properties	19
2.3 Material Selection	28
2.4 Material Specification	29
Chapter 3 Manufacturing Methods	31
3.1 Casting.....	31
3.2 Embossing and Compression Molding.....	35
3.3 Machining.....	36
3.4 Injection Molding	39
Chapter 4 Design Guidelines	65
4.1 Design Basics	65
4.2 Tolerances	81
4.3 Plastic Versus Glass	87
4.4 Shape and Thickness	90
4.5 Aspheric Surfaces.....	94
4.6 Diffractive Surfaces.....	100
4.7 Athermalization	105
4.8 Coatings.....	110
4.9 Optomechanical Design	113
4.10 Stray Light.....	118
4.11 Special Considerations for Small and Large Parts	128
4.12 Drawings	132
4.13 Vendors and Vendor Interaction	140

Chapter 5	Design Examples	143
	5.1 Singlet Lens.....	143
	5.2 Webcams	151
	5.3 Cell Phone Camera.....	164
	5.4 Infrared Multiorder or Harmonic Diffractive Lens	167
Chapter 6	Testing	177
	6.1 Parameters, Equipment, and Techniques.....	177
	6.2 Making Testing Easier	189
Chapter 7	Prototyping.....	193
	7.1 Optics	193
	7.2 Mechanical Parts	197
	7.3 Assembly and Test	198
Chapter 8	Production.....	201
	8.1 Transition to Production.....	201
	8.2 Steady-State Production	206
	References	207
	Index	213

Preface

We routinely come into contact with and utilize plastic optical systems in our daily lives. As an illustration of this, consider the following events that may occur during a typical weekday. Traveling to work in our car, traffic signals change color to regulate the flow of vehicles. Arriving at work, as we enter the building, motion sensors turn on the hallway lights. At our desk, a fingerprint reader grants us access to our laptop computer. In the lab, we take and send pictures of the latest prototype, using the camera in our cell phone, enabling others to see the hardware. After work, stopping at the store, the bar code reader brings up the prices of our items. Back at home, we enjoy the latest movie released on DVD.

All of these devices—the traffic signals, motion sensors, fingerprint readers, cell phone cameras, bar code scanners, and DVD players—rely upon plastic optical systems to perform their function. As a result, there is a growing need for individuals who are knowledgeable in the design, development, and production of such systems. This tutorial text is written with this need in mind. The book is an elaboration upon the material covered in the SPIE short course “The Design of Plastic Optical Systems.” It is meant to provide an overview of the design of plastic optical systems and is structured along the lines of a typical development project. Following a brief background discussion, the advantages and disadvantages of plastic optics are considered. Next, the available materials and their properties are described, as well as the issues of material selection and specification. Various manufacturing methods are reviewed, followed by a chapter on design guidelines, leading into several design examples. Following the examples, the prototyping and testing of a design is covered. Finally, bringing the design to production is discussed.

There are several groups that should be able to benefit from the material presented. The first group is optical engineers, who often have received training in optical system design, particularly with glass optics, but who are unfamiliar with the special design characteristics of plastic optics. The second group is technical management, who need to understand the advantages and limitations of plastic optical systems. The third, and by far the largest group, is engineers of other disciplines who find they need to design and develop plastic optical systems but lack the knowledge or training to do so.

The text is written at an introductory level. No familiarity with plastic optical parts or their design is assumed. Any background knowledge of optics and optical design will be useful, but not required, to understand most of the subjects covered. Discussions of many of the subjects covered in this text can be found distributed amongst various publicly and/or commercially available sources. There is a wealth of information on plastic optics available in articles, conference proceedings, trade journals, books, the Internet, and various patent databases. We reference many of these throughout the text and encourage the reader to utilize

1
2
3
4
5
6
7 them as additional sources of information, knowledge, and possibly, creative
8 inspiration.¹⁻⁸

9 The plastic optics industry tends to be somewhat secretive, with the details of
10 many vendors' processes, techniques, and tooling methods considered
11 proprietary. In spite of this, there are many well-established relationships
12 amongst both individuals and companies in the field. It is a relatively small
13 community, and it is not unusual for individuals to have worked for several
14 plastic optics companies, repeatedly coming into contact with people throughout
15 their career. Referral to and assistance from competing companies is not
16 uncommon. While the proprietary aspects of the field may be frustrating for
17 some readers, not knowing all the intimate details of a particular process will not
18 prevent the successful design of a plastic optical system.

19 Having completed this book, readers should understand the benefits and
20 limitations of plastic optical systems and be able to determine if this technology
21 is appropriate for their applications. They will have the basic knowledge to
22 undertake the design of these systems, should they choose to do so themselves, or
23 they will be able to have the appropriate conversations with the individuals or
24 companies they ask to perform the work.
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45

Acknowledgments

As with most people in the field of plastic optics, I entered the profession somewhat by chance. During my time in graduate school, a local company (Donnelly Optics, which became Applied Image Group Optics, which later became Aurora Optical) announced a part-time position for an optical engineer. The work involved designing plastic optical systems, with an emphasis on diffractive optics. This was aligned with my dissertation research. After successfully interviewing for the position, I began my career in plastic optics. Part-time work turned to full-time work, to the annoyance of my advisor, Bob Shannon, whose retirement I probably delayed a few years. After graduation, I remained at the company for several more years before moving to my current position at Raytheon Missile Systems (RMS). I continue to be involved with the design of plastic optical systems at work, as a consultant, and through teaching my SPIE short course.

The irony is not lost on me that after an undergraduate optics degree from the University of Rochester, several years of work as a contractor at Kodak, an MSc from Oxford, graduate school at the Optical Sciences Center, resulting, finally, in a PhD with emphasis on optical design, the first design I had go into high volume production was a single-element, lens-in-barrel system that went into a version of the “Barbie Cam,” a silver plastic toy digital camera with a large pink flower surrounding the lens. Maybe not the greatest resume addition or claim to fame, but there must be some positive karma from, to quote my wife, “all those little girls you made happy.” In retrospect, in that simple design were many of the elements of successful plastic optical systems—appropriate material selection, design for molding and assembly, integral mounting features, toleranced performance, and testability. Regardless of the simplicity of the system, there is a certain joy in seeing thousands upon thousands (if not millions) of “my design” being produced.

I have had the opportunity to work with, and learn from, a large group of people throughout my career in plastic optics. I want to acknowledge the influence and work of the following individuals: Chris Bond, Jim Busche, John Daly, Ken Davis, Patricio Durazo, Jeff Fontaine, Rob Gallagher, Dan Joseph, Bill Kuhn, Bob Kuo, Judy Love, Paul McClellan, Paul Merems, Tom Miller, Bill Naylor, Jon Nisper, Rich Pfisterer, Tom Piedmont, Tom Rooney, Dave Strickland, Al Symmons, Jay Tome, Vince Vancho, and Zefy. The list is certainly incomplete. No slight is intended to anyone whose name does not appear.

Additional thanks go to John Schaefer for sharing his expertise in diamond turning, as well as information from some of his presentations on the subject. It is a little known fact that he worked at Donnelly Optics for a short period of time, setting up their entire diamond turning facility.

I thank all those who have taught me about optical design, particularly Bob Shannon, Jose Sasian, and George Lawrence, who are or were at the Optical Sciences Center. I also thank many of my Raytheon colleagues, particularly Dan

1
2
3
4
5
6
7 Brunton, Pat Coronato, Eric Fest, Greg Hanauska, Jim Hicks, Justin Jenia, Dave
8 Jenkins, Rick Juergens, Dave Markason, Patrick McCarthy, Dan Melonis, Bill
9 Reynolds, Ron Roncone, Wayne Sunne, Byron Taylor, and Dan Tolleson for
10 useful discussions on a variety of optical, optomechanical, and nonoptical topics.
11 Dr. Roncone supplied much of the background information on harmonic
12 diffractive lenses, although any errors in the design example are mine alone.

13 I especially want to thank Barry Broome, whom I worked with during his
14 time as an optics consultant at Donnelly Optics. He is well known in the plastic
15 optics industry and has extensive experience in the design and production of
16 plastic optical systems. He previously taught the SPIE plastic optics short course,
17 referred me as his replacement in that role, and generously offered much of the
18 material associated with the course.

19 I thank my parents, Charles and Dona Schaub, for raising me to value
20 education and the possibilities that can come from it. My father never had the
21 educational opportunities that he and my mother worked so hard to provide for
22 their three children, and if he were still alive, I think he would be happy that a
23 tutorial text I authored is being published.

24 Finally, I thank my wife Elsa, minha querida, for generally putting up with
25 me, and our cat, Shadowzinha, for keeping me company during the writing of
26 this text.

27 This tutorial text is dedicated in memory of Henry J. Kreis, better known to
28 many of his Tucson friends and colleagues as "Uncle Henry." He was a good
29 friend, a great man, and an extraordinary mold-process engineer. After many
30 years at Polaroid, he moved to Arizona to take on the challenge of leading the
31 mold-processing group at a newly formed plastic optics company. I had the good
32 fortune of having the cubicle next to his, which allowed an easy, on-going
33 discussion regarding manufacturability during the design process, as well as
34 frequent brainstorming sessions. He taught me much of what I know about the
35 design of robust, producible plastic optical systems. On the production floor, as
36 lead process engineer, he taught the art of optical mold processing to the next
37 generation of mold-process engineers and technicians. In addition, he was
38 involved in the design, development, and evaluation of improved tooling
39 methods.

40 Henry loved, among other things, his family, golf (which I sometimes think
41 is the real reason he came to Tucson), a good meal, and tasty beverages. I am
42 happy to be able to say that many times I had the pleasure of sharing the latter
43 two with him. His passing not too long ago, besides leaving a void in the lives of
44 his friends and family, reduced the already small number of highly skilled and
45 experienced plastic optics mold-process engineers. It is safe to say that there are,
and will be, very few people who know as much about processing a plastic optic
injection mold as he did. Thankfully, due to his efforts, some of his knowledge
lives on in the people he trained, the coworkers with whom he shared his
experience, and hopefully, in this tutorial text.

Michael Schaub
July 2008

Chapter 1

Introduction

In this chapter, we provide some background information on plastic optics, including a brief history of their development, as well as recent advances in materials, machining, and manufacturing processes. We also discuss issues that may determine if plastic optics should be used for a particular application, as well as their potential advantages and disadvantages.

1.1 Background

Plastic optics have existed for longer than most currently practicing engineers. In 1936, after several years of development, the Rohm and Haas Company announced the commercial availability of polymethylmethacrylate (PMMA), now commonly referred to as acrylic, under the trade name Plexiglas.⁹ In the same year, DuPont began commercial production of its acrylic material, known as Lucite.¹⁰ In Britain, the Imperial Chemical Industries PLC introduced its version, which was called Perspex.¹¹ The following year, Dow Chemical introduced polystyrene in the United States under the name STYRON.¹²

Plastic lenses appeared around the same time, along with colorful marketing campaigns and disagreements over their invention. In London, on March 20, 1934, the KGK Syndicate Ltd. was formed between Peter Maurice Koch de Gooreynd and Arthur Kingston; just under four years later, in February 1938, the partnership was dissolved. The Wellcome Library, also in London, houses copies of documents concerning the partnership, as well as a copy of a letter from Kingston's lawyer regarding a dispute between the partners over who deserved credit for inventing the plastic optical lens.¹³ It seems that a BBC radio broadcast stated that Koch de Gooreynd was the inventor of the plastic lens, a claim that Kingston disagreed with.

In 1937, *TIME* magazine published an article describing competition between U.S., British, and German firms to substitute glass lenses with plastic lenses in cameras, eyeglasses, and binoculars. It is noted that shortly before the article's printing, Gooreynd (of the KGK Syndicate) appears in New York and bounces lenses on the table to show their fracture resistance. A few weeks later, in Los Angeles, a Rohm and Haas customer named E. G. Lloyd goes even further, putting on a dramatic exhibition for the press by taking a hammer to his lenses.¹⁴ Despite the impressive displays, these early lenses suffered from several problems, such as low scratch resistance and discoloration. Nevertheless, the field of plastic optics had begun in earnest.

With materials available, the manufacturers searched for markets to buy them. For instance, Rohm and Haas produced a few acrylic musical instruments, including a flute and a violin. The flute reportedly had good tone, while the violin

did not.⁹ Most early successful uses of optical plastic were either for eyeglasses or window applications. There were various attempts to create the equivalent of modern laminated windows, commonly known as “safety” glass, where a plastic film is sandwiched between glass plates to keep glass shards from flying when the window is broken.

The use of plastic windows in airplanes provided an initial market for the acrylic manufacturers and allowed them to continue development of the material. Acrylic initially was used for small windows, and as the material became more accepted and understood, larger windows; ultimately, even cockpit canopies were produced. With the beginning of World War II, the use of acrylic for military aircraft skyrocketed. Other war-time uses of plastic optics included antitank gun sights and experimental aerial cameras.¹⁵

Demand was high throughout the war period but abruptly declined with the end of the war and cancellation of government contracts. Manufacturers worked to replace the lost sales, putting acrylic into everyday items. Consumers, however, were not particularly interested in plastic products, due to the return of other materials that were not available during the war. In 1947, jukeboxes were the largest market for Plexiglas.⁹ Plastic optical elements began to appear in the auto industry, with acrylic replacing elements previously made of glass, such as tail lights. This association with the auto industry turned into a significant market for plastic optics and remains so today. Another application that developed shortly after the war, which remains with us today, is the illuminated sign.

Polycarbonate was introduced in the 1950s, with GE Plastics developing LEXAN¹⁶ and Bayer developing Makrolon.¹⁷ However, these early versions of polycarbonate were not suitable for most optical uses due to the “golden” tint of the material, as it was described by Bayer’s advertising.¹⁷ Also in the 1950s, Kodak began producing cameras with acrylic viewfinders and objective lenses.¹⁸ Cameras continued to be a large market for plastic lenses in the 1960s, with over 50 million Kodak INSTAMATIC cameras made during the decade.¹⁹ Acrylic and polystyrene lenses were often used together to achieve color correction in these types of systems. Polycarbonate elements were still being used more often in nonimaging applications, such as car tail lights and traffic signals. In 1965, ICI developed polymethylpentene, often known by the trade name TPX. It is used today in some specialty plastic optical applications. In the 1970s, continued development of polycarbonate removed most of its tint, not necessarily enough for imaging applications, but enough to enable it to be used for automotive headlamps.

Significant progress in the use of polycarbonate elements occurred in the early 1980s with the introduction of optical-grade polycarbonate. The optical grade material was used for a variety of applications, such as eyeglass lenses and safety visors, as well as for canopies on fighter jets such as the F-14. Encouraged by the demand for optical data storage discs, optical-grade polycarbonate was used for CDs, first introduced in 1982, as well as for lenses to read them. This provided, in the view of the chemical companies producing the material, an optical application with significant plastic material usage. Having similar optical

properties to polystyrene, optical-grade polycarbonate could be used as an alternative in color-correcting optical systems.

The 1990s saw an increase in available optical plastics. Two newly developed materials to emerge were cyclic olefin copolymers (COCs), commonly known by the trade name Zeonex,ⁱ and cyclic olefin polymers (COPs), known by the names Apelⁱⁱ and Topas.ⁱⁱⁱ These materials are probably the first plastics designed specifically with optical products in mind. They have similar optical properties to acrylic but have higher thermal capability and lower water absorption. We will discuss these materials in more detail in the next chapter.

Besides these new polymers, optical grades of polyetherimide (PEI), best known by the trade name ULTEM,^{iv} and polyethersulfone (PES), known by the trade name RADEL,^v also became available. These materials have excellent temperature capability, beyond the other optical plastics, as well as relatively high refractive indices. Their potential downside is a deep amber color, which limits their use for many visible applications.

Another fairly recently introduced material is optical polyester. One example is OKP4,^{vi} which is produced by the Osaka Gas Chemicals Co., Ltd. This new material has a refractive index over 1.6, birefringence properties similar to acrylic, and a low Abbe number.

In addition to improvements in plastic optical materials, there have also been significant advances in areas that affect the production of plastic optics. In particular, there have been improvements in injection molding machines, as well as in machining technologies, such as single-point diamond turning. Improvements in molding machines have been driven by both product and process innovation.²⁰ Product innovation refers to changes such as reducing the number of components in a system by making the components perform multiple functions. This often results in more complex parts, pushing the capability of the molding machines and spurring their advance. Process innovation refers to optimizing a production process to make it as efficient and capable as possible.

There has been a big push for process control and improvement across many industries, driven by the acceptance of manufacturing philosophies such as statistical process control and “lean” and “just-in-time” manufacturing. These philosophies have created demand for injection molding machines with greater process control and feedback.

Probably the largest change in injection molding equipment, particularly for use in the production of plastic optics, has been the switch from hydraulic-operated machines to hybrid or electric machines. Hybrid machines, as the name suggests, use a combination of hydraulic and electrical drives. Electric machines, sometimes referred to as “all electric” machines, completely eliminate the use of

ⁱ ZEONEX is a registered trademark of ZEON Corporation.

ⁱⁱ Apel is a registered trademark of Mitsui Chemicals.

ⁱⁱⁱ Topas is a registered trademark of Topas Advanced Polymers.

^{iv} Ultem is a registered trademark of SABIC Innovative Plastics.

^v RADEL is a registered trademark of Solvay Advanced Polymers, LLC.

^{vi} OKP4 is a product of Osaka Gas Chemicals.

hydraulics, relying on electrical motors and drives. The transition to electric machines has occurred for several reasons. The primary reason is repeatability. The repeatability of an electric machine (or its potential for repeatability) is inherently higher than for a hydraulic machine.²¹ Given the precision required in optical elements, repeatability of a manufacturing process is a highly desired characteristic. In addition to this, moving from a hydraulic to an electric machine eliminates the use of hydraulic oil in the machine clamp mechanism. Hydraulic oil is associated with several problems. It can cause environmental issues, particularly ground contamination from leaking oil. It creates health and safety concerns such as oil vapors, falls, and fire hazards. There is also a cost associated with its storage and disposal. This argument can only be pressed so far though, for as we shall see, although switching to electric machines eliminates hydraulic oil in the molding machine, oil is still often required for temperature regulation of the mold itself. It can fairly be said that the use of electric machines does not completely remove the use of oil, but it does reduce the total quantity of oil used on the molding floor. In addition, electric machines tend to use less power than hydraulic machines, resulting in energy cost savings.

The general advances in processors and computing, which have also had great influence on the process of optical design, have influenced the development and use of injection molding machines. Most machines are now fully computer controlled, with a flat screen on the side of the machine displaying the current process parameters, as well as time-based graphs of pressure and other parameters for each injection cycle. Multiple sensors provide feedback, for display as well as for closed-loop control. The molding machine can be connected to a personal computer to continuously send process and sensor data for analysis and tracking, and the data can be stored for later download and further evaluation.

Besides the injection molding machines themselves, improvements have also taken place in the ancillary equipment associated with them. This includes improvements in material dryers, thermal conditioning equipment, and automated robotic pickers and degating systems. Improvements in automation have reduced the need for human intervention during steady-state production, which results in reduced cost, as well as less chance of human error or safety incidents. It is no longer necessary to have a person standing by each machine during production.

In addition to the advances in molding machine technology, there have also been significant advances in machining technologies. This affects the precision and complexity of molds that can be built and allows the direct machining of plastic optical parts, either as prototypes or for actual production. Advances in computing and computer control, as well as in other areas, have pushed machining capabilities to previously unthinkable levels. Nowhere is this more evident than in the process of single-point diamond turning, often simply referred to as diamond turning.

Diamond turning is a machining process combining the use of gem-quality diamond tools with precision computer numerically controlled (CNC) equipment. According to Schaefer, advances that have been made over the last 30 years

include going from air bearing to oil hydrostatic slide ways, changing from belt drive to direct brushless DC spindle drives, changing from laser positioning to glass scale positioning systems, using PC DSP instead of tape CNC controllers, using epoxy granite instead of granite machine bases, and switching from ball lead screws to linear motor drive axis drives.²²

The result of these advances has been a factor of eight improvement in surface finish and a factor of four improvement in surface form. In addition, the achievable spindle speeds have increased ten fold and the cost of the diamond-turning machines has dropped by about half. The machines are commercially available with a number of options, including machines with multiple axes.

These machining improvements have enabled the creation of high-precision complex parts. Off axis, nonsymmetric, and microfeatured surfaces can all be produced using modern diamond-turning equipment. For the optical designer, this allows additional freedom in selecting surface forms or in the integration of features such as multizonal surfaces. Proper machine setup, processing, and fixturing can eliminate the need for postpolishing. This means that plastic optical prototypes can be directly machined and that optical mold inserts can be built to high optical surface accuracy and surface roughness requirements. Combining this precision machining with advances in testing, mold compensation is readily performed.

Designers can, and should, take advantage of the improvements that have been made in plastic optical materials, injection molding machines, and machining technologies such as diamond turning. Conversations with vendors and/or suppliers, who need to stay current with these technologies, is usually the best way to be aware of the current state of capability, as well as ongoing development.

1.2 When Are Plastic Optics Appropriate?

As with all optical technologies, plastic optics are not the solution to every problem. When deciding whether to use plastic optics, the first question that needs to be answered is “Are plastic optics appropriate for my application?” There are a number of considerations in determining the answer to this question. Finances, project schedule, and supply chain issues can all play a role. Production volumes, along with cost and weight, are the top reasons that people decide to use plastic optics, while thermal issues such as storage temperature limits or temperature-dependant focus shift are the main reason that they are not chosen. Other considerations, such as the spectral band used, product lifetime, or even biological compatibility, may contribute to the decision whether or not they are appropriate. These issues, along with advantages, disadvantages, and application examples, are briefly discussed in the following paragraphs.

If there is a large anticipated production volume, that is, when many systems are expected to be manufactured, plastic optics may be appropriate. “What’s the volume?” is a question typically heard early in discussions with plastic optics

manufacturers. Certain plastic optics manufacturing methods are suitable for producing large quantities of parts. In certain consumer applications it is not unusual for millions of lenses and lens assemblies to be produced each month by a single manufacturer. As an example, in 2007 worldwide production of cell phone cameras topped one billion units, most of which contained plastic optics. Some manufacturing methods, such as injection molding, allow for the simultaneous production of multiple copies of a part. Molds can also be built with the ability to later increase the number of parts simultaneously produced to allow for additional production. Several plastic optics manufacturing methods will be discussed later in the text.

While large production volumes are often appropriate for plastic optics, low production volumes can also be appropriate, particularly in the case of highly complex parts. It is possible to make a wide variety of complex plastic optical components, some of which may challenge traditional glass fabrication methods, such as the part shown in Fig. 1.1. Multifaceted surfaces, multiple zone lenses, and lens arrays are examples of complex components that are routinely fabricated in plastic. That being said, the advantage of making complex parts from plastic rather than glass is diminishing somewhat. The advances in manufacturing technologies that have improved the quality and performance of plastic optics, particularly computer-controlled machining, can and are being applied to the fabrication of glass optics. Glass elements are regularly molded, although typically at lower rates than comparable plastic elements.^{23–25}

Cost is often a factor in the decision to use plastic optics. In fact, it is typically the primary reason that they are considered. In general, and especially for large production volumes, plastic optics cost less than equivalent glass elements. This is mainly due to the time required to fabricate the element, but also partly due to material costs. The cost of a plastic optical element will be determined by a number of factors, including the fabrication method, the number of parts being made, the complexity of the part, the precision to which it must be made, its size and volume, and the material it is made of.



Figure 1.1 Complex molded plastic optical part. (Photograph courtesy of G-S Plastic Optics.)

Weight is another important reason to consider using plastic optics. For this particular characteristic, plastic optics have a significant advantage over traditional glass elements. Put simply, plastic is much less dense than glass, resulting in less weight for an equivalent volume of material. It can be argued that lenses of a given focal length, one made from glass and one made from plastic, do not necessarily require the same volume. High-index optical glasses allow longer radii of curvature, reducing lens volume, but the reduction does not typically make up the weight differential.

Head-mounted optical systems are an application where the weight advantage of plastic optics is often utilized. Heads-up displays for pilots, night vision goggles for soldiers, and microdisplays of technical manuals for mechanics or technicians are all examples of this. Studies have shown that wearing head-mounted displays can alter neck posture, increasing the stress on the musculoskeletal system of the neck and head.²⁶ Reducing the weight that must be borne will decrease the amount of stress imposed.

Other situations where weight can be a deciding factor for the use of plastic optics are those driven by battery lifetime or fuel limits. For instance, in a camcorder zoom lens, several elements typically move to change the magnification while maintaining focus. The number of times that these elements are capable of being driven before the battery must be recharged or replaced will depend upon the weight of the elements. In this case, replacing a glass element with a plastic element may increase the usage time the consumer experiences with a given battery charge. Another example would be an optical system meant to be carried by an airborne vehicle, such as a micro-UAV (unmanned aerial vehicle). The amount of time that the unit can remain aloft is in part a function of its weight, and plastic optics can be used to reduce overall system weight.

Plastic optics possess superior fracture resistance when compared to glass. Dropping a plastic lens may result in a scratch or dent to the element, but the lens will most likely remain in one piece. As an example of fracture resistance, consider that polycarbonate is often used in the production of "bulletproof glass" for vehicles or buildings. "Safety glasses," which are used for eye protection, are typically made from plastic, not glass. In this case, plastic can provide the fracture resistance necessary to meet the American National Standards Institute (ANSI) requirement that the lenses withstand the impact of a quarter-inch steel ball traveling at 150 feet per second.

Another potential advantage of plastic optics is part-to-part repeatability. Many plastic optics production methods rely upon replication processes. That is, the plastic part is formed by reproducing the shape of a master tool. An optical insert in an injection mold is one example of such a master. As mentioned above, it is possible to simultaneously produce multiple copies of a part. If a set of parts are simultaneously produced, each of the parts within the set may not be completely alike, since they may be formed from individually fabricated masters. However, the subsequent sets of parts will be very similar to the prior set. For a given part parameter, such as surface decentration, this can lead to a set of several narrow distributions. In contrast, traditional glass fabrication methods

Chapter 4

Design Guidelines

In this chapter, we discuss guidelines that apply to the design of plastic optical systems. We begin by reviewing some basics of optical design, consider tolerances and their effects, and cover various element parameters, such as thickness, shape, and different surface types. We also discuss guidelines for optomechanical design, stray light prevention and analysis, issues associated with drawings, and interacting with vendors.

We noted in the last chapter that injection molding is the most popular method of producing plastic optical elements. As such, most of the guidelines discussed are associated with the injection-molding production process. In some cases, the guidelines for injection-molded optics may differ from optics made from other processes, such as diamond turning. In these situations, I will attempt to point out the differences.

It should be kept in mind that what we discuss in this chapter are guidelines, not hard and fast rules. The guidelines presented are based mostly on practical experience (often bad experiences) and an understanding of the manufacturing process. Various designers and vendors may have different opinions regarding any given guideline, and trades can often be made between cost, performance, and producibility.

4.1 Design Basics

Optical design is a skill that previously was learned under the guidance of an experienced designer in somewhat of an apprentice relationship. The optical design community was relatively small; there were few courses of instruction and limited opportunities to enter the field. With the advent of personal computers and a number of commercially available optical design programs, access to tools to easily perform optical design has greatly increased. Currently, almost any company or individual with a computer and enough money to buy or lease an optical design program can get into the field. In this section, we present a brief discussion of the basics of optical design, beginning with a short, general review of geometric optics and aberrations. There are entire books and classes devoted to geometric optics and optical design, and given the length of this text, we will not cover the subject in great depth. We refer the reader to any (or all) of several well-known works for more detailed study.^{56–58}

A basic assumption in our discussion is that light can be represented by rays, and that light rays travel in straight lines (at least in homogeneous media). In Chapter 2, we discussed the refraction of a ray at the boundary between two media of different refractive indices. The rule that governs the refraction of the ray is known as Snell's law. We explained that the refraction of the ray is

determined by the angle of the ray with respect to the normal to the surface and by the ratio of the refractive indices of the two media. We also showed that the higher the ratio of refractive indices between the two materials, the larger the amount of refraction (or bending) of the ray.

In the refraction example discussed above, we considered a single ray at an interface. Although in Fig. 2.3 we drew a planar interface, the boundary between the two media could take any shape. In this general case, Snell's law still applies, with the ray angles taken with respect to the normal to the surface at the point of intersection of the ray. We now consider the situation of a ray crossing an interface that is spherical, such as when a light ray is incident on the surface of a conventional spherical lens. Figure 4.1 illustrates the geometry of this situation, where we have shown the lens as a sphere. We assume the material to the left of the interface to have a refractive index of 1 and the lens material to have a refractive index of 1.5. Similar to our earlier example, we find the angle between the ray and the normal to the surface (at the point of intersection of the ray), and with the values of the refractive indices, we determine the angle between the surface normal and the refracted ray. We can imagine extending the ray further to the right, where it will intersect the second surface of the lens. At this interface, the rear surface of the lens, we can again find the angle of the ray with respect to the surface normal, calculate the angle of refraction, and determine the direction of the ray after it passes through the surface. We have thus computed the path of a ray through a lens.

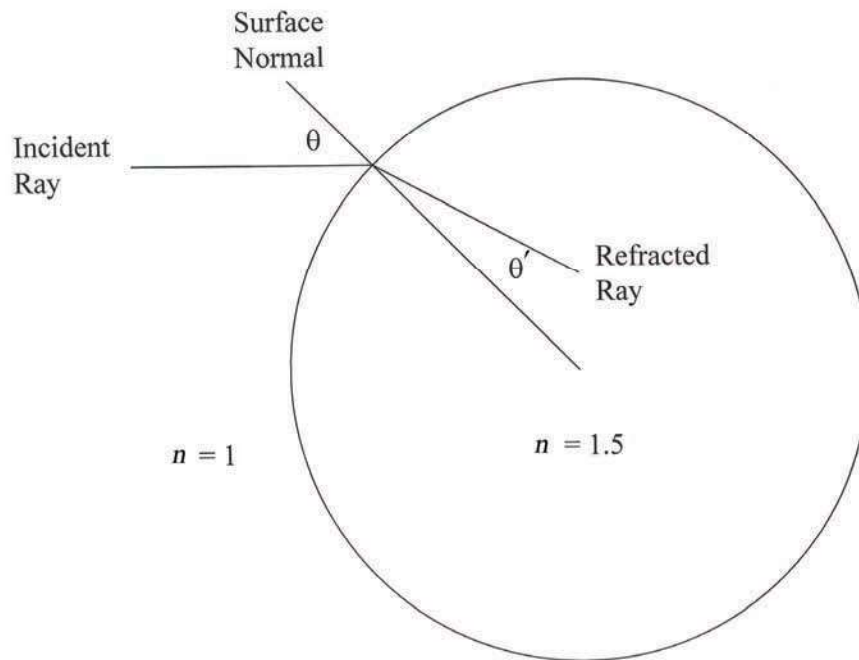


Figure 4.1 Ray refracting at a spherical surface.

If we now increase from one ray to several rays, we have the case of a finite-sized beam of light incident on and passing through a lens, as was shown in Fig. 2.4. We can see that the parallel ray bundle incident on the lens is converging after having passed through it. Because of this, the lens is known as a converging (positive) lens. If the ray bundle had been diverging after passing through the lens, it would be known as a diverging (negative) lens.

A lens, to the first order, is described by the location of a set of points known as the cardinal points. The cardinal points consist of the principal points, the nodal points, and the focal points. Each of the cardinal point types come in pairs. There is a front principal point and a rear principal point. The same is true for the nodal and focal points. Figure 4.2 shows a biconvex lens with the cardinal points labeled. The principal points are points of unit magnification. If a plane is drawn perpendicular to the axis, through a principal point, it is known as a principal plane. The principal plane can be considered the representative plane of bending of the rays in the lens. This can be seen by extending the input and output rays and noting that their intersection occurs at the (rear) principal plane. If we were to bring a ray from the rear side of the lens, it will appear to bend at the front principal plane. In a real system, the principal surfaces (which represent the ray-bending surfaces) are not truly planes, but for our discussions it is acceptable to consider them as such.

The nodal points are the points of unit angular magnification. That is, if a ray is aimed at the front nodal point, it will appear to emerge from the rear nodal

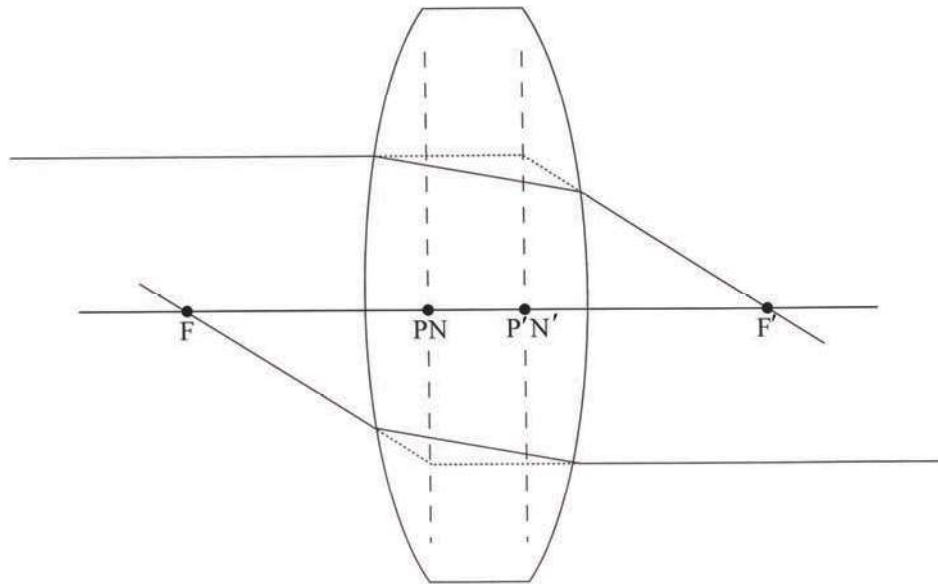


Figure 4.2 Biconvex lens with cardinal points labeled.

point at the same angle. If the lens system is in air, the nodal points will coincide with the principal points.

The focal points are the positions where an input ray that is parallel to the optical axis crosses the axis after passing through the lens. This can be seen for the input ray coming from the left in Fig. 4.2, which crosses the axis at the rear focal point, F' . A ray parallel to the axis coming from the right of the rear surface of the lens would cross the axis at the front focal point, F .

The distance from the principal point to the focal point is known as the effective focal length (EFL) of the lens. Most people have some familiarity with the concept of effective focal length. For a thin lens, the focal length is approximately the distance from the center of the lens to the location along the axis where an input collimated beam comes to focus. For a thick lens, or for a system made of multiple lenses, the focal length is not measured from the center of the lens system. To find the effective focal length, we extend an input ray (parallel and near to the axis of the system) as well as the corresponding output ray and find their intersection. The intersection of the two rays is the principal plane, and as before, the distance from the principal plane to where the output ray crosses the axis is the effective focal length. Depending on the powers and positions of the elements within the system, the principal plane may be in front of, inside of, or behind the physical system.

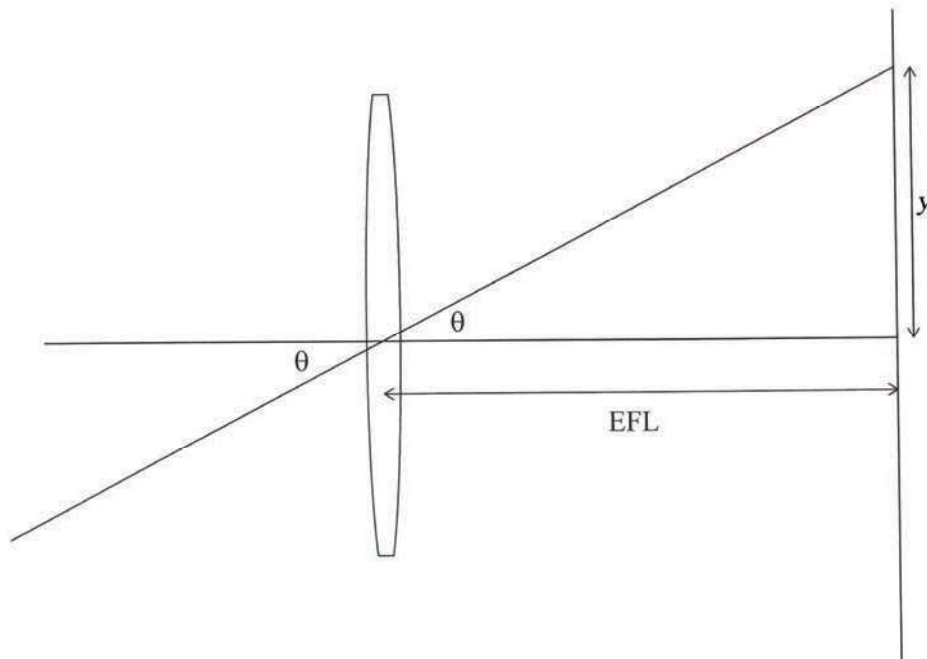


Figure 4.3 Relationship between focal length and image height.

The distance from the last optical surface of the system to the image is known as the back focal length (BFL), while the distance from the last piece of mechanical structure to the image is known as the flange focal length (FFL). In some systems there will be a required length of the BFL and/or FFL in order to provide necessary clearance for other elements, such as a fold mirror.

The focal length of a single lens can be calculated from knowledge of its radii of curvature, its thickness, and its refractive index. The equation for the focal length of a lens in air is shown in Eq. (4.1):

$$\frac{1}{\text{EFL}} = (n-1) \left[\frac{1}{R_1} - \frac{1}{R_2} + \frac{t(n-1)}{nR_1R_2} \right], \quad (4.1)$$

where n is the refractive index of the lens material, R_1 is the radius of the first surface of the lens, R_2 is the radius of the second surface of the lens, and t is the center thickness of the lens. The sign convention for the radii is that the distance is positive if the center of curvature is to the right of the vertex of the surface and negative if the center of curvature is to the left of the vertex. As an example for the biconvex lens shown in Fig. 4.2, the front surface of the lens (left-hand surface) has a positive radius, while the rear surface has a negative radius.

The focal length of a lens or lens system provides the scaling factor between input angles for collimated beams and the height of the image formed by the lens, as is shown in Fig. 4.3. Collimated input beams can be considered to come from object points at an infinite distance (or very far) from the lens. We can see from the figure that the height of the image is related to the focal length and input angle through the equation

$$y' = \text{EFL} \tan \theta. \quad (4.2)$$

Using this equation for any given focal length lens, we can determine the image height as a function of input angle. Alternatively, if we have selected a particular detector and a desired field of view, we can calculate the focal length required to achieve it.

In the case of finite object distances, we can determine the location and size of an image by knowing the focal length of the system and the distance of the object from the front focal point. An example of a finite imaging situation is shown in Fig. 4.4. The relationship between the object and image distances is given by

$$xx' = -ff', \quad (4.3)$$

where x is the distance of the object from the front focal point, x' is the distance of the image from the rear focal point, f is the front focal length, and f' is the rear focal length of the system. For a system in air, f and f' are equal. Setting $f = f'$ and solving for the image distance x' , we obtain

$$x' = -\frac{f^2}{x}. \quad (4.4)$$

The height of the image h' can be determined from

$$h' = \frac{hf}{x} = \frac{-hx'}{f}, \quad (4.5)$$

where h is the height of the object. The ratio of the image height to the object height (h'/h) is known as the magnification and is usually denoted by m . Using Eqs. (4.4) and (4.5), we can determine the image location and size for any object height and position when imaged with a given focal length lens or lens system.

We next consider the concepts of the aperture stop and the pupils. In any optical system, there is some aperture that limits the size of light beam that can pass through the system. In some cases, this aperture may be the diameter of a lens. Alternately, it could be the mounting flange or retaining ring of a lens. In many cases an aperture (such as an iris) is specifically positioned to set the beam size. The limiting aperture is called the aperture stop of the system. In most photographic cameras, there is an adjustable aperture that is the aperture stop. When the aperture is closed down to a smaller opening, it is referred to as “stopping down” the system.

The entrance and exit pupils are nothing more than the images of the aperture stop when viewed through all the elements in front of and behind the aperture stop, respectively. All of the light entering the system appears to go into the

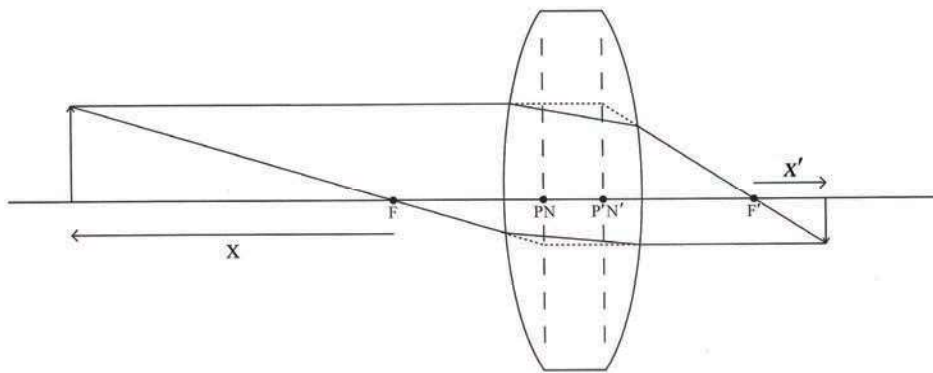


Figure 4.4 Example of imaging at finite conjugates.

entrance pupil, and all of the light exiting the system appears to come out of the exit pupil. When we look into the front of a lens system and see the limiting aperture, we are looking at the entrance pupil. Similarly, when looking into the rear of the lens system, we see the exit pupil.

The ratio of the focal length of the lens system to the entrance pupil diameter is the f number ($f/\#$) of the system. The $f/\#$ describes the system's light capturing ability. Smaller values of $f/\#$ (such as $f/2$ versus $f/4$) are associated with more light capturing ability. Smaller $f/\#$ systems are referred to as "faster" systems; that is, an $f/2$ system is faster than an $f/4$ system. The term comes from early film cameras, where a faster system required the shutter to be open for a shorter period of time than a slower system (i.e., the picture could be taken faster).

Up to this point in our discussion, we have considered only the first-order properties of lenses. We have discussed that we can determine the size and location of an image, but we have not yet concerned ourselves with the quality of that image. By image quality, we refer to how much the image of a point object looks like a point. If we consider a generic object to be made up of a collection of points, the quality of the image will depend on how well the lens system images each point making up the object to a corresponding point in the image.

In general, a point in the object is not imaged to a point in the image but (at best) to a small blur. This is due to diffraction as well as the aberrations of the lens system. Diffraction deals with the wave nature of light and sets a lower limit on how small the image of a point can be in the absence of aberrations. At the moment, we do not concern ourselves with diffraction, only with aberrations. Aberrations are the departure from perfect imaging. When an optical system has aberrations, the image of a point does not look like a point, and/or the image of a point is in the wrong location. We next consider some common aberrations seen in optical systems as well as some general techniques for controlling them. More extensive discussion of aberrations, aberration theory, and techniques for aberration control are found in several works.^{59–61}

We begin our discussion of aberrations by considering chromatic aberrations. Chromatic aberrations can be thought of simply as the variation in first-order characteristics of the lens (or lens system) with the wavelength of light. Consider, for instance, a singlet lens used with the visible spectrum. We know that the different colors in the visible spectrum are associated with light of different wavelengths, with red being longer and blue being shorter wavelengths. Earlier, we discussed the fact that materials have different refractive indices for different wavelengths, which is known as dispersion. We also saw in this section [from Eq. (4.1)] that the focal length of a single lens depends in part on its refractive index. It follows that since the lens has different refractive indices for different wavelengths, it will also have different focal lengths for different wavelengths. This is shown in Fig. 4.5. The rays representing the blue light, for which the lens material has a higher index, focus nearer to the lens than the rays representing the red light, for which the lens has a lower refractive index. The higher blue-light refractive index results in a shorter focal length than the lower red-light refractive

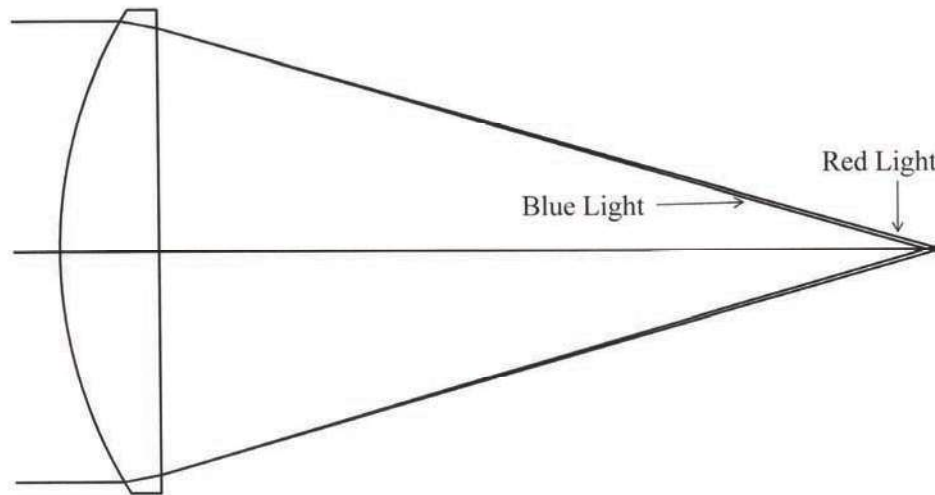


Figure 4.5 Lens exhibiting axial chromatic aberration.

index. This axial separation of the foci of the different wavelengths of light is known as axial chromatic aberration or axial color. The effect of axial color is to create color blurring in the image. If we were to place a screen at the axial position of the focus of the blue light, we would (in geometric terms) see a bright blue spot surrounded by a red blur. The red blur would be due to the fact that the red light had not yet come to focus. In reality, we would not just see a red blur because all the wavelengths between blue and red would be striking the screen, somewhat out of focus.

Now that we understand the potential variation of focal length with wavelength, we can predict what happens due to this variation in the case of an off-axis image point. We stated earlier in Eq. (4.5) that the height of the image depends on the focal length of the lens. If the focal length varies with wavelength, then the image height will also vary with wavelength, and for a given object, different wavelengths (colors) will end up at different image heights. Using the same lens as the previous case, we have a shorter blue focal length and a longer red focal length. This results in the blue light being imaged to a different height than the red light. This separation in height of the different colors is known as lateral chromatic aberration or lateral color.

In addition to chromatic aberrations, there are geometric and nongeometric aberrations. Geometric aberrations result in the rays from a point object not coming together to form a point image, regardless of the image surface location or shape. With nongeometric aberrations, we can adjust the image surface position and shape such that we can obtain a point image for a given point object. We consider the geometric aberrations known as spherical aberration, coma, and astigmatism, and the nongeometric aberrations of Petzval curvature and distortion. Each of these aberrations can vary with color, so we could have spherochromatism, which is the variation of spherical aberration with

wavelength. The variation of these aberrations with wavelength is a higher-order consideration, which we will not discuss in detail here. In our discussion of geometric and nongeometric aberrations, we assume we are using light of only one wavelength.

The first geometric aberration we consider is spherical aberration. Spherical aberration can be thought of (in the case of a simple lens where the pupil and lens are colocated) as a variation in focus with lens aperture radial position. For more complex multielement lenses, it is the variation of focus with respect to the radial position of a ray in the entrance pupil. Figure 4.6 shows a lens with a planar front and a spherical rear surface that is exhibiting spherical aberration. In this case, rays from the outer edge of the lens focus closer to the lens than do rays from the central portion of the lens. The existence of spherical aberration can be fairly simply understood by considering Snell's law. As the rays move farther out from the axis of this lens, the angle of incidence on the rear lens surface increases. Correspondingly, the angle of refraction increases, but not at the rate that would be required for the rays to all pass through the same point on the axis. In the case shown, the rays refract at larger angles than would be desired for perfect imaging. When the rays from the outer portion of the lens (or pupil) focus closer to the lens than the rays near the center, the spherical aberration is said to be "undercorrected." If the reverse were true, and the rays from the outer portion focused further away, the spherical aberration would be "overcorrected." The amount of spherical aberration, that is, how far the rays from the edge of the pupil intersect the image plane compared to the intersection of the ray through the center of the pupil, varies as the third power of the pupil size, but it is constant with field angle.

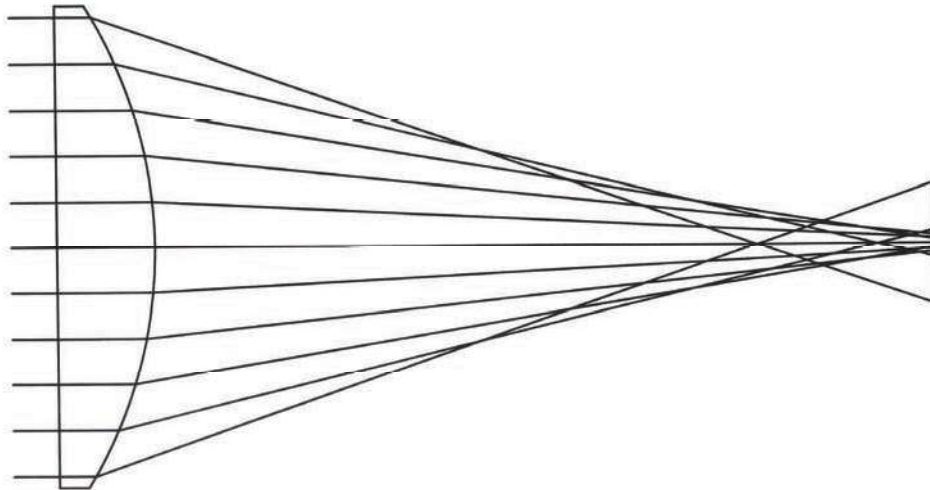


Figure 4.6 Plano-convex lens with spherical rear surface exhibiting a large amount of spherical aberration.

There are several design methods to control spherical aberration. One method is to change the shape of the lens such that the angles of incidence on the surfaces are adjusted (typically reduced). Changing the lens shape while maintaining its focal length is referred to as “bending” the lens. Figure 4.7 shows a lens with a different shape but with the same focal length as the lens shown in Fig. 4.6. We can see that this “bent” lens shape has much less spherical aberration. Another method to reduce spherical aberration is to “split” the lens, dividing the power of the lens amongst multiple elements. This method also reduces the angles of incidence on the surfaces, which reduces the amount of spherical aberration. Yet another method to control the spherical aberration is to use an aspheric surface. By selecting the correct surface shape, we can adjust the angles of incidence such that the rays are all properly refracted to go through the same axial image point. An example of this is shown in Fig. 4.8, where the rear surface of the lens is a conic surface with a conic constant equal to the negative of the square of the refractive index of the lens.

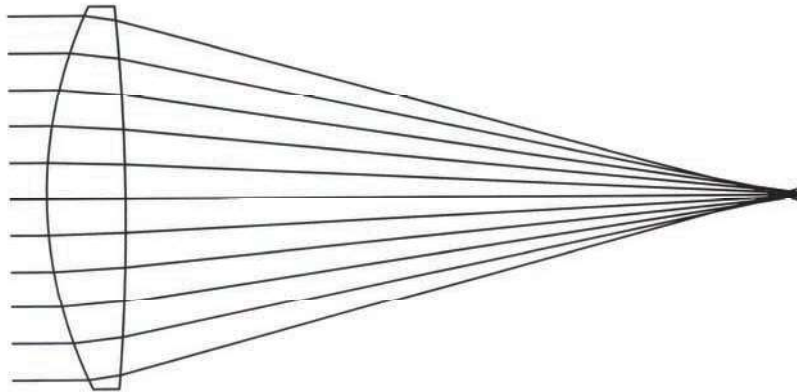


Figure 4.7 “Bent” lens showing reduced spherical aberration.

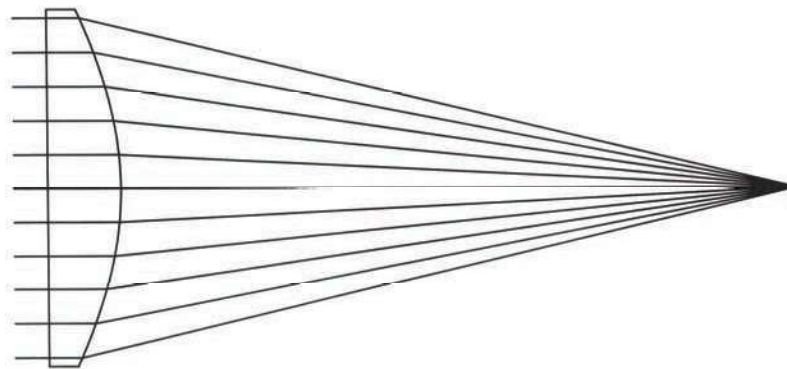


Figure 4.8 Plano-convex lens with rear conic surface, exhibiting no spherical aberration.

The second geometric aberration we discuss is coma. The name for coma comes from the comet-like appearance of a point image when this aberration is present. Coma is a variation in magnification as a function of radial position in the entrance pupil. Rays from different annular zones within the pupil strike the image plane at different heights, as shown in Fig. 4.9. We can see in this figure that the rays from the outer edge of the pupil come together at a different height than the ray from the center of the pupil. The amount of coma varies with the square of the pupil diameter and linearly with the field angle. Coma can be a particularly annoying aberration because it produces an asymmetric image of a point object. There are several methods used to control coma. One method is to move the axial location of the aperture stop, which adjusts where the beam of rays strike the lens. Another way to control coma is through the use of symmetry. If a lens system is symmetric about the aperture stop, the coma introduced by the lenses before the aperture stop will be cancelled by the lenses after the aperture stop. This is technically true only if the object and image distances are the same, which would give a magnification of one. However, even if the system is not used symmetrically, the coma is to a large extent cancelled between the two halves of the lens system.

The third geometric aberration we consider is astigmatism. Some readers may be familiar with astigmatism due to having it in their visual system. When astigmatism is present, rays in two orthogonal planes through the lens system

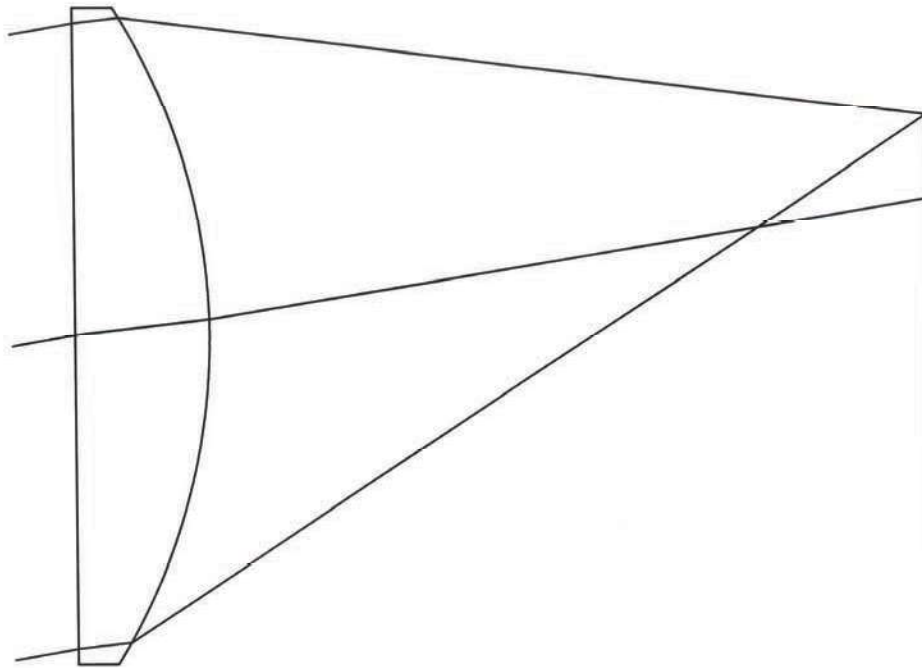


Figure 4.9 Plano-convex lens exhibiting coma.

focus at two different axial locations. At each focus, the image of a point object is a line, due to the rays in one plane being out of focus. In between these two foci, the image of a point object has an elliptical shape. Midway between the foci, the shape is circular. Astigmatism arises from the fact that as the beam strikes the lens at an off-axis position, the height and width of the beam are different. As a result, rays in the two directions have different angles of incidence and are refracted by different amounts. The amount of astigmatism depends upon the shape of the lens, as well as its distance from the aperture stop. For a given lens, astigmatism depends linearly on the pupil size and with the square of the field angle. By adjusting the shape of a lens and its distance from the aperture stop, the astigmatism of an element can be controlled. It can also be controlled through the use of an aspheric surface, which will change the angles of incidence that the beam sees when it strikes the surface.

Having discussed the geometric aberrations, we now consider the nongeometric aberrations. Again, we refer to them as nongeometric because they do not cause the image of a point object to blur; instead, they change where the point-like image of a point object is located. The first nongeometric aberration we discuss is Petzval curvature, which is an inherent curvature of the image surface of a lens. While we often refer to the “image plane” (because our film or detectors are typically planar), the preferred shape for a positive lens, if all other aberrations were corrected, would not be a plane but an inward-curving surface. This curved image plane is known as the Petzval surface and is illustrated in Fig. 4.10. The use of a planar image surface, instead of a curved one, results in the off-axis points being out of focus. The blur depends linearly on the pupil size and quadratically with field height. It should be noted that some optical systems, such as the human eye and the Schmidt camera, use a curved image surface.

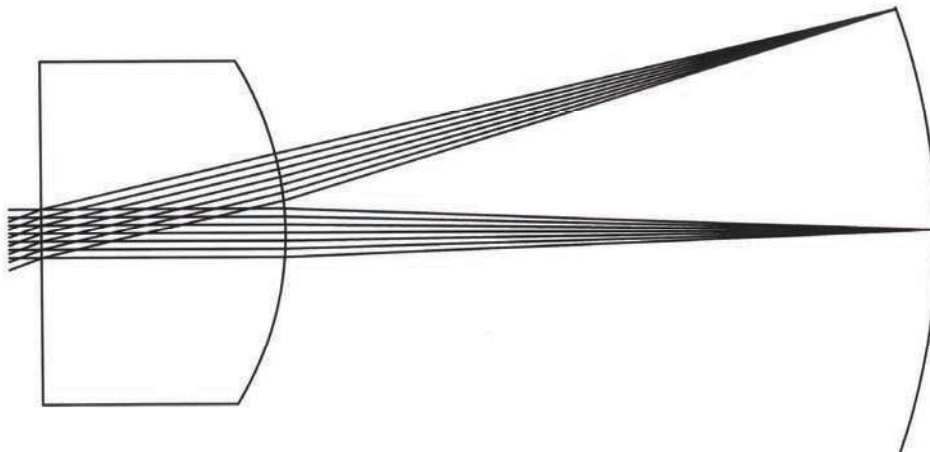


Figure 4.10 Lens showing a curved image surface, known as Petzval curvature.

The amount of Petzval curvature depends directly upon the power of a lens and inversely on its refractive index. Being an inherent property of the lens, the Petzval curvature contribution of a lens does not depend upon its position within a system of lenses. While the Petzval surface of a positive lens is inward curving, the opposite is true for a negative lens; a negative lens has an outward curving Petzval surface. Thus, it would seem logical that one way to reduce the amount of Petzval curvature would be to combine positive and negative lenses. If we were to take a positive and a negative lens with equal and opposite power and combine them in a doublet, we would end up with no Petzval curvature—that is, with a flat image plane. The problem with this arrangement (assuming thin lenses are in direct contact) would be that the powers of the lenses would cancel each other (they are equal and opposite), so we would be without any optical power. However, while the contribution of any lens element to the Petzval curvature does not depend on its location within the system, the contribution of its power to the total system power does. Therefore, we can separate the two components, which will create optical power, while maintaining the sum of their Petzval contributions, which is zero. Using separated positive and negative elements is a common optical design technique to reduce the Petzval sum. An example of this can be seen in the Cooke triplet, which consists of a positive element, a negative element, and a positive element, all of which are axially separated.

Another method of reducing the Petzval curvature, also known as “flattening the field,” is the use of a negative field lens placed near the image plane. A field lens is a lens that is placed near the image plane or at an intermediate image. We can understand how a negative lens placed near the image plane flattens the field by considering the fact that it provides increasing glass thickness as a function of field height. When a converging beam passes through a block of glass, the focus position of the beam is shifted by an amount related to the thickness and refractive index of the glass. In the case of a negative field lens, the glass thickness increases as a function of the field, so the image is increasingly shifted as a function of the field height. By choosing the correct surfaces on the field lens, that is, by setting the correct thickness variation, the image is shifted by the appropriate amount to make it lie on, or nearly on, a plane. One potential problem with a lens of this type, often referred to as a “field flattener,” is its location near the image plane, where the converging beams have a small cross-sectional area. Any small defect on the lens (such as a dig) or any contamination can block a large portion of the beam heading to a given field point.

The last aberration we discuss is distortion. Most people are familiar with distortion, having seen the distorted image from a wide-angle camera lens. We discussed earlier how to compute the height of an image, given a focal length and field angle. This image height assumes that there is no distortion in the system. If distortion is present, the image of a point object will not be at the expected height, but it will be displaced. The distance that the actual image is displaced from the predicted image is the amount of distortion of the system. The amount of distortion is often quoted as a percentage—the ratio of the displacement of the image to the predicted image height times 100.

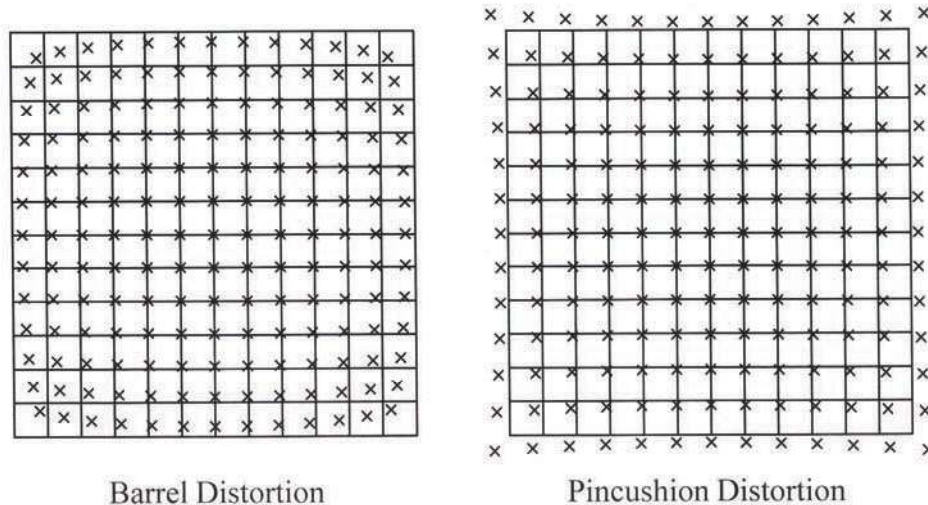


Figure 4.11 Effect of barrel and pincushion distortion on the image of a grid object.

The distortion contribution of an element depends upon the distance of the element from the aperture stop of the system as well as on the element's thickness. The distortion of a system is constant with pupil size and varies with the third power of the field. Because of this cubic dependence, the image of a grid object has the familiar “barrel” or “pincushion” appearance associated with distortion, depending on whether the distortion is positive or negative. Examples of these are shown in Fig. 4.11, where the crosses in the plots show the location in the image of the corners of the squares forming the grid object.

Distortion, like coma, can be eliminated or reduced through the use of symmetric lens element arrangement. The distortion introduced by the front half of the system will be cancelled by the contributions of the rear half of the system. Distortion can also be controlled by the use of aspheric surfaces. In many imaging applications, small amounts of distortion (up to about 2%) can easily be tolerated.

Having briefly reviewed geometric optics, as well as aberrations and some techniques to control them, we now consider the basic process of optical design, the role of the designer, and the role of optical design software. There are several excellent texts that discuss these topics in greater depth than will be covered here.^{62–64} The process of optical design begins with an understanding of the requirements that the completed design must meet. Basic requirements such as focal length, field of view, $f/\#$, and wavelength range are typically imposed. In many cases, specific numbers are not available, but a general desired range is known. Some type of performance requirement, such as encircled energy, modulation transfer function (MTF) value, or resolution is also usually stated. In addition to these basic optical requirements, there are often additional constraints,

such as packaging volume. For plastic optical systems, the factors that drive their selection (such as cost and weight) can be additional constraints. One of the functions of the designer is to understand the effects of the various constraints, as well as their interactions, and to determine if they can all simultaneously be met.

In many cases, the customer requesting the design is not familiar with the specification of optical systems. They know what they want but do not necessarily know how to put their desires in the form of a specification. Here the designer must play the role of interpreter, translating the customer's desires into requirements that the design can be evaluated against. For example, in the design of a web camera, the customer may not know the field of view required, but they know that they want to be able to see a person's head and shoulders when using the camera. Using this information, the optical designer can calculate approximately what field of view is needed, and if a specific detector is to be used and what the focal length of the lens needs to be.

In some instances, the customer cannot even be this specific. They may want to be sure to see the person's head and shoulders but cannot decide how much beyond their shoulders they should see. This may result in the need for a trade study, where, for example, different field-of-view systems are designed, and the performance, cost, and other factors compared. The customer may have multiple, competing, and sometimes unachievable desires. In this case, the designer must communicate closely with the customer, explaining what is and what is not achievable. Additionally, the designer must describe the cost, be it monetary, performance, or something else, that is necessary to meet a certain requirement.

Once the basic requirements, desires, and constraints are understood, a basic lens form, or starting point, can be selected. Often, the design that is required is similar to one that already exists; it is difficult to come up with something completely new. In this case, the existing design may be used as a starting point for the new design. Various parameters (such as focal length) can be adjusted in order to meet the requirements of the new design.

One question that often arises is, "Where do I get such an existing design?" In the design of glass optical systems, there are several sources that can provide a wide range of design forms. Multiple books exist that present the design details of various systems, as well as discussions of how they work, and why they look the way that they do.^{65,66} Patent databases are another excellent source of potential design starting points. Of course, when using a patented design as a starting point, the designer must ensure that the new design does not infringe on the patents. A third source of starting-point designs is the optical design software program used to perform the design work. Most of the optical design programs come with a database of designs.

In the design of plastic optical systems, it can be more difficult to obtain an existing design to use as a starting point. Many plastic optic manufacturers have a database of designs they have developed over the years, but these are typically not shared with the general public. The patent database can be an excellent place to look for a starting point, although at times it can be difficult to find exactly what is desired. Conference or journal articles can also be good places to look for

designs. In many instances, the exact prescription is not available, but the general lens form can be enough to begin with. If comparing glass and plastic solutions, beginning with a glass design and changing the glasses to optical plastics may be one place to begin. If no comparable design can be found to use as a starting point, the designer can always begin by using a single lens and adding lenses as needed.

With a starting point defined, the development of the design can begin. Today, most optical design work is performed on a computer with the assistance of an optical design software program. Readers who are interested in old-school (precomputer) design methods are referred to the book by Conrady.⁶⁷ There are multiple design programs available commercially, the best known of which are CodeV,^x OSLO,^{xi} and ZEMAX.^{xii} These programs are highly sophisticated and enable the designer to perform a wide range of designs and analyses. The choice of which software program to use is often dependent on a number of variables, such as personal preference and cost. For most plastic optical designs, any of the well-known commercial design codes can perform adequately.

There is currently no closed-loop algorithm to design an optical system. While the same result can be arrived at through different paths, wildly divergent designs may also arise from the same starting point. This is partly due to the large number of variables that are available in the design process. Variables are the parameters, such as radius of curvature, that the design program will vary in an attempt to improve the design. Typically, in a plastic optical design, each lens surface (radius and possibly aspheric coefficients), lens material, and to some extent thickness, as well as lens location, are available as variables. It is the job of the designer to determine what parameters are to be assigned as variables as well as their allowed range.

Once the variables are defined, an algorithm for evaluating the lens is developed. This algorithm is commonly known as the “merit function.” The merit function provides a numerical representation of the “goodness” of a design. A small value of the merit function means that a lens is considered better than another lens with a larger value (assuming the same merit function algorithm is used). The optical design codes normally have one or more default merit functions that can be modified by the designer as appropriate. The default merit functions typically use some composite of the performance (such as spot size) at each of the defined field angles to generate the merit function value. In addition to the spot size (or other performance metrics), constraints such as focal length, allowed packaging length, or distortion are also evaluated.

With the variables and constraints defined, the design may now be “optimized.” Optimization is usually performed by slightly changing each of the variables and determining which way they should be adjusted to improve the design—that is, lower the merit function value. At the same time, the adjustment

^x CodeV is a registered trademark of Optical Research Associates.

^{xi} OSLO is a registered trademark of Lambda Research Corporation.

^{xii} ZEMAX is a registered trademark of ZEMAX Development Corporation.

of the variables must be such that the constraints are met. Once the necessary changes are determined, the variables are adjusted (usually by a small amount), the merit function is evaluated, and the cycle is repeated until the improvement in the merit function value is less than some predefined amount. The goal of the optimization is to drive the system to the design form having the lowest value of the merit function. Because the merit function is a function of multiple variables, the solution found may be a local optimum, but not a global one.

Most of the optical design programs feature some type of “global optimization” function. The goal of the global optimization is to find the “best” lens solution given the constraints and variables entered. It must be remembered that “best” is determined solely by the inputs to the computer program. Any constraint that is not included will not necessarily be met. Some constraints (such as cost and manufacturability) are not always easily entered as constraints, although work in this area is ongoing.⁶⁸ The global optimizers can be useful for generating different design forms, which may themselves be used for local optimization. In some cases, the global optimization run will result in the preferred solution; in others, an ideal trade amongst the various constraints will not be achieved, and the designer will have to continue to work on the design.

Up to this point, we have discussed the design process as working towards determining the lens form with the lowest merit-function value. One thing to remember is that the merit function typically only evaluates the nominal lens system. In reality, we do not necessarily want to manufacture a lens with the best nominal performance, but a lens that provides the best “as-built” performance, taking into account the tolerances (and cost) that are associated with building it. The evaluation of the performance (and cost) of the various tolerances on the parts and their assembly is known as “tolerancing” the design and is discussed in the next section. The tolerancing of a design is a critical part of the design process. No design should be considered complete until a tolerance analysis has been performed.

Sometimes, the most difficult decision a designer must make during the design process is the determination that a design is finished. This is particularly true for novice designers. With a deadline approaching, this may not be a problem, as the design is done when time has run out. In other cases, when the design schedule is more open ended, it can be difficult to shake the feeling that “just a little more” performance can be squeezed from the system. Experience (or sheer exhaustion) will often allow the designer to recognize the need to bring the design effort to its conclusion.

4.2 Tolerances

There is potentially nothing more frustrating for an experienced optical designer to hear than the statement that “the design is all done, I just need to tolerance it.” In reality, nothing could be further from the truth. An optical design is not complete until all necessary tolerances have been assigned. The tolerancing of a

design can often take longer than the generation of the nominal optical prescription. Until a design has been toleranced, it is not known if its components can be made accurately enough, or how much it will cost to make the system meet its requirements. In this section, we discuss tolerances that can be achieved on plastic optic elements and consider the basic process of tolerancing a plastic optical design.

While some individuals and vendors in the plastic optics industry do not like to state explicitly what tolerances they can hold (due to part design dependence), we feel that for many typical optical elements there is a fairly standard set of achievable tolerances that a design should initially be evaluated against. Of course, every optical element needs to be individually evaluated, with the assistance of the molder, to assess the chances and/or cost of achieving the desired tolerances. Table 4.1 shows typical tolerances that can be achieved in molded plastic optical elements. The labeling of the columns can be debated as to the accuracy of each term. The terms used are similar to those used in other works discussing tolerances. Coming from a precision plastic optics molding house, the author believes that the tolerances labeled state-of-the-art are actually typical of the tolerances most optics molders hold, and that even slightly tighter tolerances may be able to be achieved. There may, however, be cost savings associated with alternate vendors and looser tolerances, and we suggest that the designer determine what tolerances are actually required to make their system perform adequately, as opposed to what tolerances can be achieved. We now briefly discuss each of the tolerances in the table.

Table 4.1 Typical tolerances for injection-molded optical elements.

	Commercial	Precision	State-of-the-Art
Radius	$\pm 5\%$	$\pm 2\%$	$\pm 0.5\%$
EFL	$\pm 5\%$	$\pm 2\%$	$\pm 1.0\%$
Thickness (mm)	± 0.13	± 0.05	± 0.020
Diameter (mm)	± 0.13	± 0.05	± 0.020
Surface Figure	$< 10f (5\lambda)$	$< 6f (3\lambda)$	$< 2f (1\lambda)$
Surface Irregularity	$< 5f (2.5\lambda)$	$< 3f (1.5\lambda)$	$< 1f (0.5\lambda)$
Surface Roughness (RMS)	$< 100\text{\AA}$	$< 50\text{\AA}$	$< 20\text{\AA}$
Surface S/D Quality	80/50	60/40	40/20
Wedge (TIR) (mm)	< 0.025	< 0.015	< 0.010
Radial Displacement (mm)	< 0.100	< 0.050	< 0.020
Aspect Ratio	$< 8:1$	$< 6:1$	$< 4:1$
Repeatability[†]	$< 2\%$	$< 1\%$	$< 0.5\%$
DOE[§] Depth (μm)	---	± 0.25	± 0.10
DOE Min. Groove (μm)	---	25	10

* diameter/thickness ratio [†] part to part in one cavity [§] diffractive optical element

NOTE: Above tolerances are for 10- to 25-mm-diameter elements.

The tolerance on the radius of curvature depends on the precision of the optical insert, the injection molding process, and the shape of the surface. Steeper surfaces can typically be held to a tighter tolerance than weak shallow surfaces. This is due to the increased self support that a steeper surface provides. A tighter tolerance on radius may also be achieved through compensation of the optical insert.

The tolerance on the EFL includes within it the tolerances on both surfaces of the element as well as the refractive index of the material. As such, tighter EFL tolerance may require tighter radii tolerances. In some cases, it is the back focal length (BFL), not the effective focal length (EFL), that is of interest. In this case, the tolerance should be called out on the BFL, perhaps with a loosened requirement on EFL.

The thickness of an element is set by adjusting the positions of the optical inserts within the mold. The inserts typically rest on thick, precision-ground spacers. The spacers are often initially built too thick and then are ground to the correct length after the mold processing has been performed. Any change to the optic insert, such as re-turning the surface to remove a scratch or swapping an insert due to damage, may require adjustment to the spacers. In some cases, the accuracy of the thickness measurement sets a lower bound on the thickness adjustment.

The diameter of the element is typically formed by the feature (hole) that is machined into the mold base. This feature size, along with the material shrinkage, sets the final diameter. Improved machining has resulted in diameter features in the mold that are held extremely tightly. However, in some cases the diameter may increase when the part is removed from the mold. Unless the mold has been machined in a “steel safe” condition (with extra material left on), it may not be easy to correct the diameter size.

Glass optical surfaces are typically specified using the terms “power” and “irregularity.” The terms come from the test process, where a lens surface is compared against a highly accurate spherical test plate. When putting the test plate and lens together, and using a nearly monochromatic light source, a series of rings (fringes) would be observed from the interference of the light reflected from the two surfaces. If the lens surface was perfectly spherical, but of the wrong radius, the fringes would be perfectly circular and the number of fringes seen would relate to the difference in the radii of the lens surface and test plate. Any departure of the lens surface from true spherical shape would result in changes to the circular shape of the fringes, referred to as irregularity. When dealing with plastic optics and aspheric surfaces in particular, the definitions of surface figure and irregularity are not universally agreed upon.

For our discussion, we consider surface form (figure) to be how closely the surface matches the desired surface, in a symmetric way, while irregularity describes how rotationally asymmetric the surface is. This description is most appropriate for surface testing using a contact profilometer. Sometimes the term cylindrical irregularity is used, which allows the radius to be adjusted between measurements instead of using one radius value for all measurements. For

interferometric surface measurement, having only a radius tolerance and a form tolerance is the preferred method. This format can also be used with surface profilers. Whatever specification method is chosen, it is important that the surface description used in the tolerance analysis accurately reflects the specification, and that the designer and vendor agree upon the definitions and test methods.

If we have specified a spherical surface, we would like the surface produced to be spherical as well. However, due to edge break or variation in shrinkage, the surface may end up being slightly aspheric. Provided the asphericity is less than the surface figure tolerance, the surface will conform to its specification. The surface form tolerance is often specified in fringes. Conversion to distance units can be obtained by noting that two fringes correspond to a distance of one wavelength (typically chosen as 632.8 nm).

The surface figure tolerance is typically used in conjunction with the radius tolerance. That is, the radius is first adjusted (up to its allowed tolerance) to minimize the form error, then the form error (surface figure) is evaluated against its tolerance. Surface figure can be adjusted by altering the mold process or by compensating the mold. In some cases, in order to produce a spherical surface, an aspheric optical insert will be used. Surface figure typically gets harder to hold as the size of the part increases. The values in the table are for parts up to about 25 mm in diameter. For symmetric parts up to 75 mm in diameter, Beich⁶⁹ suggests that less than two fringes per 25 mm can be achieved.

While the surface figure tolerance sets how well the surface must match the desired surface, the surface irregularity tolerance sets how symmetric the surface must be. As opposed to the terms surface form and irregularity, other authors use the terms irregularity and astigmatism.²² Plastic optical parts may have some amount of asymmetry in them due to the nonuniform flow of the material during the molding process or due to ejection of the part.

While the surface form and irregularity deal with the surface variation on a larger scale, the surface roughness controls the variation of the surface on a small scale. The main effect of surface roughness is scattering of the light passing through it, which can be particularly troublesome in laser-based systems. In most systems, low enough surface roughness can be achieved to reduce any impact on system performance. In the case of molded plastic optics, the surface roughness largely depends on the surface roughness of the optical insert it is replicated from. Polished stainless steel inserts can achieve excellent surface roughness. The more commonly used nickel-plated inserts, which are diamond turned, also can achieve low-surface roughness values and can be postpolished if necessary.

The surface roughness is called out as a root mean square (rms) value. The actual value of surface roughness reported depends on what spatial frequencies are evaluated in the measurement. There is typically some upper frequency bound set by the wavelengths that the system uses and some lower frequency bound set by the measurement equipment. This is covered in the discussion of testing in a later chapter.

The scratch-dig specification sets how many and how large the scratches and digs (pits or pit-like defects) on the optical surface can be. The nomenclature of the scratch-dig specification is two numbers separated by a slash. The numbers refer to the allowed size of scratch or dig. For instance, a 40/20 callout means that the maximum allowed scratch width is 40 μm , and the maximum pit diameter is 0.2 mm (200 μm). Scratch-dig inspection is usually performed visually using a calibrated set of reference samples. Scratch-dig specifications are typically referenced to some defining document, such as MIL-SPEC-13830, which gives additional guidance on their interpretation. For small elements, scratch-dig issues can be important, as a dig may take up an appreciable amount of the surface area.

The tolerance on wedge in molded plastic optics is controlled in the production of the mold. Because of the way the molds are made, wedge is usually not a large problem. The wedge specification is often called out as a total indicated run-out (TIR) quantity, which should not be confused with total internal reflection. The total indicated run-out is the full range that a dial indicator moves when held against the part during the measurement.

Radial displacement is another tolerance that is largely controlled by the quality of the mold. As we discussed earlier, molds are typically made up of various inserted pieces. Each inserted piece will have some fit tolerance with the piece it goes into as well as centration of its features to its diameter, for example, the centration of the optic surface to the diameter of the optic pin. These tolerances combine to create the final radial displacement tolerance. In actuality, there are two radial displacements that must be considered. These are the displacement of an individual optic surface to the diameter of the lens and the displacement of the two optic surfaces to each other. These values are regularly held to approximately 20 μm , with surface-to-surface values in the 5- μm range held for small elements in cell phone cameras.

Aspect ratio, or the ratio of the diameter of the lens to its thickness, is not really a tolerance but more of a rule of thumb. As this ratio decreases, the parts look less and less like a standard lens and more “cube-like.” If the aspect ratio goes significantly below the values listed, discussion with the molder should take place before proceeding much further with the design.

Repeatability is usually one of the positive characteristics of molded plastic optics. Once the mold has been placed in the injection-molding machine and the process is stable, the parts that are produced from a given cavity will be very much like each other. Part size and shape have some influence on part-to-part repeatability. For instance, a surface with a long radius will tend to vary more than a surface with a short radius, and irregularity on large parts will tend to vary more than irregularity on small ones.

The tolerances on diffractive optical surface features are typically set by the accuracy of the master. With proper mold processing, accurate replication of the master can be achieved. Because of their small feature sizes, shrinkage is usually not a concern. For diamond-turned masters, the depth tolerance is set by the accuracy of the diamond-turning machine, while the minimum groove width is

typically set by the diamond tool size. Smaller feature sizes can be obtained through the use of other mastering techniques. As the features grow smaller, attention must be paid to the ratio of the depth to the width of the feature. As the depth-to-width ratio grows larger, the feature becomes more difficult to replicate.

One tolerance that was not listed in the table was variation in the refractive index. Molders do not typically want to set a refractive index tolerance because it is difficult to measure, and more importantly, because it is largely out of their control. While changes in the mold processing can affect the refractive index, the main contributor to index variation is the plastic optical material itself and how well it is controlled by the manufacturer. As we discussed earlier, several of the plastics were not specifically intended for use in optical devices, and the manufacturers do not necessarily attempt to control the refractive index variation. To be on the safe side, an index variation of ± 0.002 is often assumed. If a design is highly sensitive to refractive index variation, careful consideration of the design is suggested.

Having discussed the achievable tolerances, we now consider the tolerancing of the design. More detailed discussions of tolerancing can be found in Refs. 60 and 63. There are two main types of tolerancing performed: sensitivity analysis and predicted performance analysis. Sensitivity analysis aims to determine the sensitivity of some system characteristic to each of the individual tolerances. The characteristic may be related to image quality, such as MTF, or it may be some other metric, such as system boresight. Predicted performance analysis (as the name implies) aims to predict the performance of the system or the range of its performance, given a set of tolerances and their distributions. As in the sensitivity analysis, various parameters can be used as representations of the system performance. In each of the tolerance analyses, compensating changes, if they exist, must be specified. For instance, in many optical systems, there will be a final focus adjustment. This may occur by moving the lens barrel with respect to the image plane before it is fixed in place. This adjustment, known as a compensator, should be included in the analysis. Without the compensator, the system performance will appear much more sensitive or much worse than it really is.

To conduct the sensitivity analysis, a small change is made to each system parameter (radii, thickness, etc.) associated with a tolerance, one at a time, and the change in the system characteristic is evaluated. After running through each toleranced parameter, a sensitivity table is displayed. This table tells the designer which parameters the characteristic is most sensitive to. If the system performance is extremely sensitive to a particular parameter, the tolerance on that parameter will need to be tightly controlled, or the design needs to be adjusted to decrease its sensitivity.

To perform the performance prediction analysis, the designer must specify ranges for each of the tolerances as well as the distribution of the tolerances. The distribution of tolerances for a plastic optical system may be different from the distribution of tolerances in a glass system. The reason for this is the difference in manufacturing methods. Consider the thickness of a single element. In a

plastic optic, the center thickness will be set by adjusting the optic pin locations, as discussed earlier. The adjustment will be made with the attempt to reach the nominal thickness value, but the final thickness may fall anywhere within the allowed range. For a glass optic, often the center thickness will be on the high side. This is because it is safer for the optician to leave on extra material in case additional polishing is required. Once the surfaces meet their requirements and the center thickness is within the tolerance range, work on the lens will usually stop, leaving the center thickness on the high side of nominal. In this case, the predicted distribution of glass lenses would be skewed toward the high side, while the thickness of the plastic lenses would be more symmetrically distributed.

Because of the repeatability of plastic optics, it is often best to assume that their tolerances will take values near the end of their ranges. This is a more conservative approach than assuming a uniform distribution of values, and it will generally predict reduced performance compared to the uniform distribution. The predicted performance analysis can be run with various tolerance probability distributions to see the effect of distribution choice.

Predicted performance analyses are typically conducted in one of two ways: either through the use of statistical analysis or through the use of Monte Carlo techniques. In the statistical analysis method, the tolerance sensitivities are computed, along with their derivatives, and the performance is predicted by statistically combining the tolerance effects. In the Monte Carlo technique, tolerance values are randomly selected from within their distribution, and the lens is evaluated. This procedure is repeated a large number of times until the performance distribution is obtained. The statistical method is far faster than the Monte Carlo method, but it depends upon the underlying statistical assumptions. The Monte Carlo method, if properly set up, may be more representative of the actual built systems. Often it is useful to use the statistical method throughout much of the design cycle, and then perform a Monte Carlo analysis near the end to verify the performance predictions. We will show an example of Monte Carlo analysis and the effect of tolerance probability distribution in the next chapter.

4.3 Plastic Versus Glass

From a pure design viewpoint, standard techniques used in the design of glass optical systems can be utilized in the design of plastic optical systems. Many techniques, such as the balancing of lower- and higher-order spherical aberration, the use of spherochromatism to balance axial color, astigmatic field flattening, and even the use of chromatic distortion to balance lateral color, can be applied to both glass and plastic optical systems. However, in many cases, factors other than optical design may not allow a standard technique to be used, requiring alternate solutions. A number of classic design techniques are discussed in the lens design book by Smith.⁷⁰ We follow the general order used in that text in our

discussion of the differences in applying these design techniques to plastic and glass.

A common technique used in designing glass optical systems is the “splitting” of an element. Splitting a highly powered element into two (or more) elements, each of approximately equal power, and together having the same total power as the original element, reduces the angles of incidence on each of the surfaces, which results in a reduced amount of aberration. This is true whether the lens is made of glass or plastic. In the design of a plastic lens system, however, there may be a limit on the number of lenses that can be used due, for example, to cost or space constraints. The cost of a plastic optical system increases directly with the number of elements, so a given price point may set the number of elements allowed. An example of a space-constrained system would be a cell phone camera. In this design, a maximum overall length is usually imposed. In theory, increased numbers of elements could be shoved into the available space by decreasing the element thickness and spacing. In reality, there will be a minimum desired thickness such that the lenses can be manufactured. For production using injection molding, the lenses must be thick enough to allow the flow of the injected molten plastic.

A potential alternate solution to splitting the lens would be the use of an aspheric surface. Aspheric surfaces are readily manufactured on plastic optics, and the use of an appropriate aspheric surface may eliminate the need for the extra element created by splitting. Aspheric surfaces are discussed in a later section of this chapter.

A second standard technique in the design of glass systems is the use of an achromatic (or intentionally chromatic) doublet, a technique sometimes referred to as compounding an element. We can think of this technique as creating a new glass type, one that doesn’t exist on its own, by combining two different glasses. In many cases, the compounded element created is a cemented (as opposed to an air-spaced) doublet. The use of achromatic or partially achromatic doublets helps with the correction of chromatic aberrations as well as control of other aberrations (we have introduced at least one additional potentially cemented surface).

We can take the compounding technique further, in the case of a cemented doublet, by separating the cemented elements. By doing so, we introduce additional degrees of freedom. The shape of each element can now be independently changed, without the requirement of a matching cemented interface surface. Separating the elements may require tighter control of the decentration between them. The refraction at the new glass-air interface will typically be larger than at the previous glass-glass interface, with the possibility of total internal reflection without adjustment of the surfaces. This, however, can be dealt with in the design process and should not hinder separating the cemented elements.

As discussed above, cost or space reasons may not allow the use of sets of achromatic doublets. If doublets are used in a plastic optical system, they are usually not of the cemented variety. Some cemented plastic doublets have been

manufactured, but they are not the norm. Instead, air-spaced doublets, or the use of separated elements of varying materials similar to those used in a Cooke triplet, are more likely to be seen. As such, separation of the elements to achieve added performance will already have been used and will not be an additional available technique in the plastic optical system.

Similar to using an aspheric surface instead of splitting a lens, it may be possible in a plastic optical system to use a diffractive surface instead of adding an additional element as required to compound an element. A diffractive surface on a plastic lens, similar to an aspheric surface on a plastic lens, does not increase the cost of producing the element. There is, however, a potential performance cost, in terms of stray light, from using a diffractive surface.

Other standard techniques in the design of glass optical systems, based on material selection, are to raise the index of any positive singlet elements and lower the index of any negative singlet elements, or to raise the refractive indices of all the elements in general. A higher refractive index in a positive element will reduce the amount of Petzval curvature it introduces as well as reduce other aberrations through the use of longer radii surfaces, which typically decreases the angles of incidence on the surfaces. For a negative element, a lower refractive index will increase its (usually correcting) contribution to the Petzval sum. However, it will also result in a reduced radius of curvature, which may have a negative effect on aberrations. The balance of Petzval curvature and aberration correction must be considered when lowering the index of negative singlets. Raising the refractive index of all the elements will again reduce the required curvatures, thus reducing the aberration contribution of the surfaces.

As noted earlier, there are significantly fewer optical plastics to choose from compared to optical glasses. This limits the number of possible material choices available to the designer. In fact, some of the optical plastics, such as polystyrene and polycarbonate, have similar optical properties to each other, effectively reducing the choices even further. Depending on other constraints, such as thermal requirements or spectral band, there may be no higher-index plastic materials that are available. In this case, aspheric surfaces may again be used to provide the aberration reduction that would have been obtained from changing to a higher-index material. In terms of Petzval correction, if alternate materials are not available, it may be that a certain amount of Petzval curvature must be tolerated.

Another design technique in glass systems (which we have already proposed as an alternate solution in plastic systems) is the use of aspheric surfaces. In the design of glass optical systems, the use of an aspheric surface should be carefully considered. While improved manufacturing techniques [such as glass molding, computer-controlled grinding and polishing, and magnetorheological finishing (MRF)^{xiii}] have made it possible to produce aspheric glass elements, they still typically cost several times a similar spherical glass element. In some cases, however, the increased cost of a glass asphere is well worth the performance or

^{xiii} MRF is a registered trademark of QED Technologies.

packaging improvements that it brings. In the case of plastic optical elements, an aspheric surface costs no more than a spherical surface and should be used appropriately as needed.

In addition to the classic design techniques discussed above, there are also some “rules of thumb” used in the design of glass optics that do not apply to plastic optics. Probably the best known of these is the conversion (in a glass lens) of a weakly powered long-radius surface to a planar surface. For instance, in the discussion of lens bending, we showed that the optimum form for minimum spherical aberration (for $n = 1.5$), with an infinite object distance, has a weakly powered rear surface. In moving to a production design, the optician making the lens would commonly ask if the surface could be made planar, which would be easier and less costly to manufacture. In the case of molded plastic optics, just the opposite is true. Instead of turning the weak surface into a planar surface, we would want to add more power to it. The reason for this comes from the molding process. A curved surface will have more structural support (or surface tension) than a planar surface. This will help the surface to be stable while the part is cooling, resulting in less variation in radius as well as less surface irregularity. As discussed in the section on molding, when planar surfaces of high quality are required, it is sometimes necessary to move to an injection-compression molding method.

If the lens is not going to be molded but rather diamond turned, it is not necessary to add power to the weak surface. In fact, it may be easier to make the surface flat, which will allow it to be easily fixtured when diamond turning the curved surface.

Another rule in glass lenses that is often not required in plastic optic elements is symmetry. In the case of glass biconvex or biconcave lens elements that are almost symmetric, the rule of thumb is to force them to be symmetric. This serves several purposes, such as reduced tooling and test plates in the optical shop, and it eliminates the possibility of the lens being installed backwards in the assembly. For plastic optics, the optical inserts are usually formed by independently machined optical inserts. The inserts themselves do not necessarily have the same length or diameter on the two halves of the mold, which reduces the possibility of them being inserted into the wrong side. Also, the flanging on many plastic optical elements is not symmetric (often intentionally so), due to mounting or molding design. The use of automated assembly equipment, with indexed features, should reduce the incorrect insertion of elements. In addition, for manual or automated assembly, orientation marks (such as indentations on the flange) may be molded into the part to provide a visual reference.

4.4 Shape and Thickness

In addition to the rule-of-thumb differences between glass and plastic optics mentioned above, there are also often differences in the preferred shape and

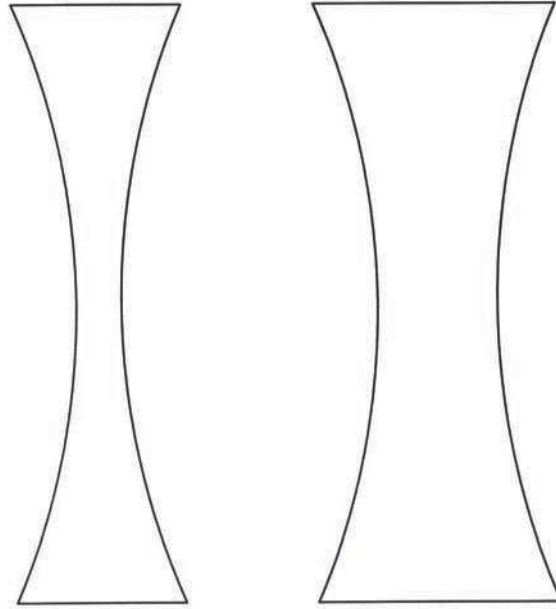


Figure 4.12 Negative lens with thin (left) and with preferred (right) center thickness.

thickness of plastic versus glass optics. These differences are driven by the manufacturing method, in this case injection molding. A fundamental requirement in the molding of plastic optics is sufficient thickness for the molten plastic to flow through. This generally constrains two parameters: the center thickness of the element, particularly for negative lenses, and the edge thickness of the element. Regarding the center thickness, consider the element shown in the left-hand side of Fig. 4.12. This would be a fairly standard glass component, but it is not a preferred plastic element. The reason for this is the flow of plastic during the injection molding process. The gate for injection of the plastic will typically be on the edge of the part. As the plastic enters the mold cavity, it will want to take the path of least resistance, which for this lens will be around the thicker outside portion of the lens. Any flanging added to the part will potentially exacerbate the problem. If the plastic first flows around the periphery of the lens, reaching the center last, the joining of the flow fronts in the central portion of the lens will result in a knit line. In addition to the knit line, the large thickness variation over the part will potentially result in a significant shrinkage variation over the clear aperture, possibly requiring several mold compensation cycles to achieve the desired surfaces.

Increasing the center thickness of the part, as shown in the right-hand side of Fig. 4.12, can eliminate both of these molding problems. With this increased thickness, the mold processor can get the plastic to flow through the center of the part, eliminating the knit line or at least moving it out of the clear aperture. The

reduced thickness variation also helps with the variation in shrinkage over the part. Some plastic optics vendors do not like to specify a standard edge-to-center thickness ratio but prefer to look at each part individually. Others will quote a rule of thumb for the edge-to-center thickness ratio to be less than either three or five, depending on the individual asked.

With regard to edge thickness of the lens, the requirement is to provide enough thickness for a suitable gate. The gate size required will depend on the individual part size and shape, but it is always preferable to have room to enlarge the gate if needed. Small parts may be able to use a smaller gate than large parts, but there will generally be some lower limit on minimum gate size. Gates with a width as small as 0.5 mm have been used on small lenses for cell phone cameras.

Edge thickness can be constrained in the optical design program during the optimization process. It is important that the designer understands how the edge thickness value is calculated in the particular software being used. If it is calculated from the sags of the surfaces at the heights of maximum ray intersection, the value calculated may meet the constraint value entered but not the real manufacturing requirement. This is because the clear aperture of the surface will need to be larger than the height of the ray intersections to allow for edge break in the molding process. Molders typically will ask for at least 0.5 to 1 mm of radial distance between the used portion of the optical surface and the transition into the element flange or diameter. If possible, the designer should attempt to provide greater than this amount, particularly on larger parts or parts with large amounts of thickness variation.

In the case of parts with a relatively large center thickness and a small edge thickness, as is sometimes seen on biconvex parts, adequate edge thickness must be allowed so that the gate can be large enough to prevent “jetting.” Jetting occurs when the plastic entering the cavity sprays across the open cavity space instead of smoothly flowing through the cavity. It can result in numerous knit lines or unusable parts. Limiting the power of the element, or increasing the center thickness, can generally alleviate the thin edge thickness. Jetting is obviously not a concern for parts that are to be diamond turned.

To be clear in regard to our discussion of edge thickness, it is the minimum cross-sectional area that the molten plastic must flow through that is typically the limiting factor. Occasionally, an engineer will attempt to solve the edge-thickness problem by adding a flange to the lens that is wider than the thickness at the edge of the clear aperture. While adding a flange may enable a wider gate, the plastic still must pass through the smaller thickness area. This is not to say that adding a flange will not help. Having the larger gate available may open the process window for the mold engineer; however, in the case of undersized edges it will not completely solve the flow problem.

Due to variations in shrinkage with thickness, from the viewpoint of a molder, the preferred shape of a molded element is one that has a fairly uniform thickness. In practical designs, in order to obtain sufficient optical power, most optical elements will not have this form. The goal of the optical designer should be to create parts that are as moldable as possible while still meeting the system

performance requirements. In some cases, there may need to be a trade between the performance and ease of production of the system. Figure 4.13 shows a series of parts, some of which should mold well, and some of which should, if possible, be avoided. Referring to the figure, Lens A would be a fairly typical glass convex-plano lens. For plastic, however, it is not a preferred shape. The edge thickness is fairly thin, which would limit the gate size. In addition, the rear plano surface would not support itself well during the molding process, potentially leading to sink on the surface. A preferred version of the lens is shown as Lens B. The center thickness of the lens has been increased, which allows a better edge thickness. The plano surface has been changed to a fairly low-powered surface, which will provide better support and less variation in radius and irregularity. Lens C shows a typical meniscus lens. This lens should mold well, due to its relatively uniform thickness. Lens D shows a biconvex lens with mounting flanges added. The flanges have been extended far enough to protect the vertices of the surfaces. The flanges are of reasonable length and should not cause problems when molding the lens. Lens E shows a lens where the flange has been extended in order to act as a lens spacer. This is not a preferred lens shape because the long flange will be difficult to fill while maintaining the optical surfaces. It would be better to use a lens similar to Lens D and add a separate spacer if required.

Because of constraints such as cost or packaging, the number of elements a designer is allowed to use may be limited. In this situation, it may be necessary for each element to “work” as hard as it can. This, in addition to the use of aspheric surfaces, may result in lens shapes that are unfamiliar to those used to working with glass lenses. Figure 4.14 shows two examples of such lenses, which while not typically seen in glass designs, would be considered perfectly normal to those familiar with plastic optical design. The lens on the left comes from a web camera design, and the lens on the right comes from a cell phone

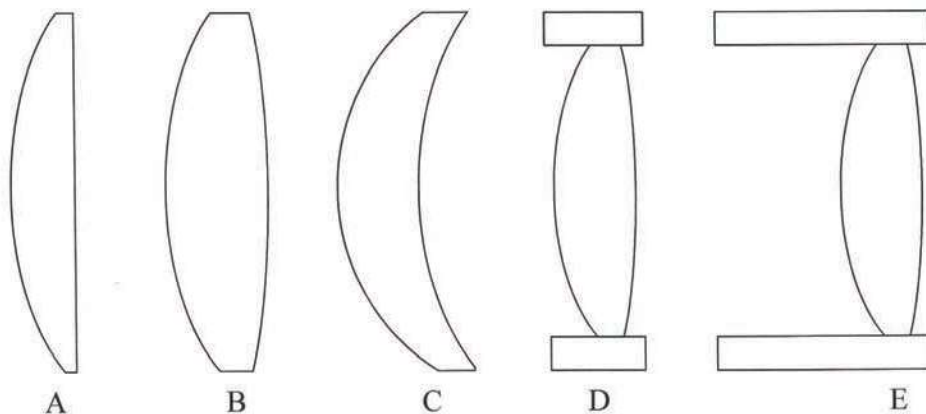


Figure 4.13 Various element shapes, some of which should be avoided (A, E).

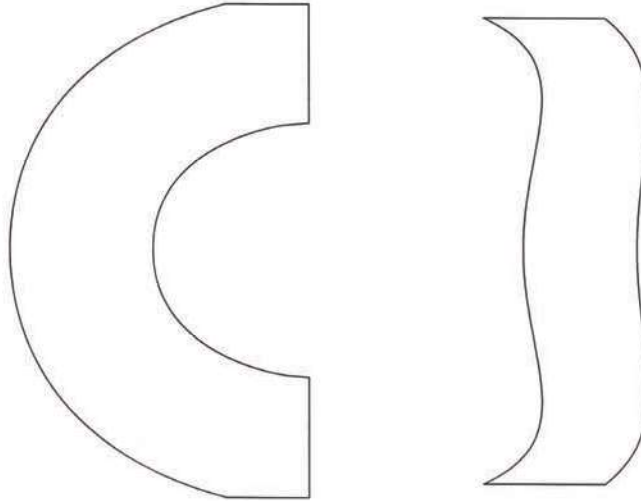


Figure 4.14 Two examples of lens shapes seen in plastic that would be unusual in glass.

camera design. These are not just design study curiosities, as both systems have been put into high-volume production. From a manufacturing standpoint, the lens on the left is considerably easier to produce in plastic than in glass. Manufacturing it in glass requires each lens to be polished individually, as opposed to blocking up a number of lenses on a spindle, which is normally done. In plastic, a multiple-cavity tool produced several copies of the lens with each mold cycle.

Experienced designers of glass optical systems often have familiarity with certain lens shapes as well as an understanding of why systems look the way that they do. They can often get an idea of what the system is doing just by looking at the lenses and comparing it to their knowledge of the “classical” design forms. The same is often true with plastic optical systems. With experience, familiarity with certain plastic lens shapes will develop, and what previously seemed strange will be completely normal.

4.5 Aspheric Surfaces

We have mentioned several times that with plastic optics, unlike glass, creating an aspheric surface costs essentially nothing more than creating a spherical surface. Given this, it makes sense to take advantage of aspheric surfaces in plastic optical designs. An aspheric surface is simply a surface that is not spherical. While a spherical surface can be completely described by defining its radius of curvature, an aspheric surface requires a more complex representation.

We consider two main types of aspheric surfaces, those that are rotationally symmetric and those that are not. Consider first aspheric surfaces with rotational

symmetry, of which there are (at least) two well-known forms. The first form is the family of conics. Those who remember their geometry will already know that conics are plane curves that are created by the intersection of a plane and a right circular cone. If the plane is perpendicular to the axis of the cone, the intersection curve is a circle. As the plane is tilted, the intersection curve becomes an ellipse (of course, a circle can be considered a special case of an ellipse). Continuing to tilt the plane, such that one line in the cone is parallel to the plane, the intersection of the two is an open curve, which is a parabola. Continuing to tilt the plane, now two lines in the cone will be parallel to the plane, and the intersecting curve is a hyperbola. When these intersection curves are rotated about their symmetry axis, they form surfaces that are called conic surfaces. As an example, in the case of rotating the plane curve of a parabola, we would generate a parabolic surface or a paraboloid.

Conic surfaces are easily represented in a closed form, as shown in Eq. (4.6):

$$z = \frac{cr^2}{1 + \sqrt{1 - (1+k)c^2r^2}}, \quad (4.6)$$

where z is the sag of the surface, which is the distance along the axis of the surface from a vertex plane perpendicular to the axis of the surface; c is the base curvature, the inverse of the radius at the vertex of the surface; k is the conic constant; and r is the radial coordinate, which is the perpendicular distance from the axis. An illustration of the sag of an aspheric surface is shown in Fig. 4.15. The conic constant, along with the base radius, completely describe the conic

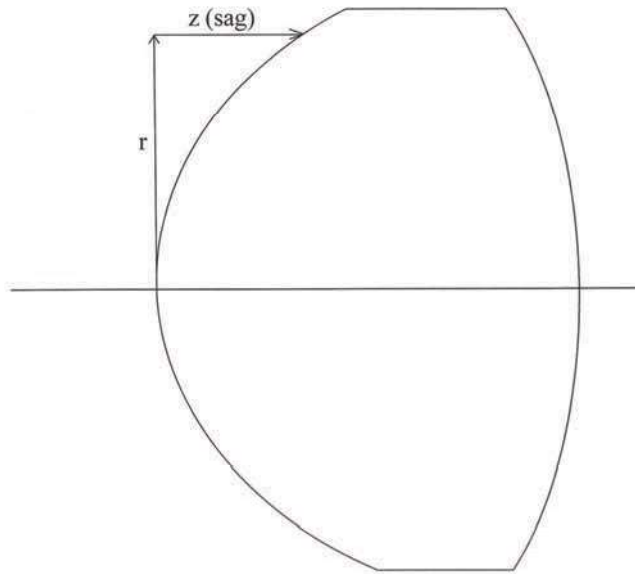


Figure 4.15 Illustration of sag for an aspheric surface.

surface. Using the description in Eq. (4.6), a surface that has a conic constant of 0 is spherical, a surface with a conic constant between 0 and -1 is a (prolate) ellipsoid, a surface with a conic constant equal to -1 is a paraboloid, and a surface with a conic constant less than -1 is a hyperboloid. The conic constant can also be greater than zero, in which case the surface is an oblate ellipsoid.

Conic surfaces have the property of perfect point-to-point reflective imaging for a single pair of points along their axis. For instance, an ellipse provides perfect imaging between its two foci, while a parabola provides perfect imaging of a point at infinity. Thus, a parabolic surface can take a collimated on-axis input beam and focus the beam perfectly at its focal point. As soon as the beam moves off axis, however, there will no longer be perfect imagery.

Conics have long been used in the design of reflecting telescopes. The Ritchey-Chretien telescope form, which consists of hyperbolic primary and secondary mirrors, is an example of this. Conics' property of on-axis point-to-point imaging is useful as a test method in their production. A point source, or input collimated beam in the case of parabolic surfaces, can be used, and the resulting image evaluated. Assuming proper object and focus position, deviation from perfect imagery results from imperfect surface form. Interpretation of the point-source image allows correction of the surface to its proper form.

The second common form of description of aspheric surfaces is a polynomial addition to a sphere or conic, as is shown in Eq. (4.7). This is a fairly standard form used in most optical design software:

$$z = \frac{cr^2}{1 + \sqrt{1 - (1+k)c^2r^2}} + a_2r^4 + a_3r^6 + a_4r^8 + a_5r^{10} + \dots, \quad (4.7)$$

where the a_i terms are the coefficients of the aspheric terms, and the other variables are the same as in Eq. (4.6).

In Eq. (4.7), the asphere is described using even powers. Sometimes, the absolute value of odd power terms is also used (if the absolute value is not taken, the surface will not be rotationally symmetric). The naming convention of this type of aspheric surface is related to the highest power used. For instance, if the highest order coefficient used is that related to the eighth power, we would say that the surface is an eighth-order polynomial asphere. Obviously, there is an assumption that we are using the standard even-ordered form. To be more specific, we may say we have an eighth-order, even-polynomial asphere.

Polynomial aspheric surfaces do not exhibit the point-to-point imagery of conic surfaces. This means they cannot be tested in the same way. In the past, this often led to conics being selected over polynomial aspheres. Developments in machining and testing have made this much less of a concern, and polynomial aspheric surfaces are frequently used in systems where the surfaces will be produced by diamond turning and/or molding, such as in reflective, infrared, and plastic optical systems.

In addition to the conic and polynomial representation of aspheric surfaces, there are several other forms. Most optical design programs provide the choice of multiple aspheric surface representations as well as the option for user-defined surfaces. In some cases, the aspheric surface may not have a closed-form solution but may be represented by an array of points, with a suitable interpolation in between.

When required, nonrotationally symmetric aspheric surfaces are also used. In some cases, the surface is simply an off-axis section of a rotationally symmetric surface. Thus, the parent surface is rotationally symmetric, but the used section is not. In other cases, the surface is an extension of the surfaces already described. Instead of describing the surface by a single radial variable, two variables (often the Cartesian coordinates x and y) are used. For instance, a biconic surface may be utilized, which will be described by (possibly) two radii, and two conic constants, k_x and k_y . Similarly, we can change from the radial variable on the polynomial asphere to two Cartesian variables, resulting in an anamorphic asphere. Nonrotationally symmetric aspheres may be used in applications where some system characteristic is nonrotationally symmetric. For instance, laser diodes often emit a noncircular beam that has inherent astigmatism. Nonrotationally symmetric aspheric surfaces may be used to correct this output beam.

In classical optical design, designers are often taught to do as much as possible using spherical surfaces, with an asphere grudgingly used as a last resort in order to meet some performance requirement. As such, the general design form may be well established before the addition of an aspheric surface, with the asphere tweaking up the system performance but not necessarily altering the basic form of the system. This design philosophy is changing somewhat as the use of aspheric surfaces becomes less costly. In the design of plastic optical systems, the decision to use aspheric surfaces is often almost a foregone conclusion, changing the point at which they are inserted into the design process as compared to classical glass systems. This can result in design forms that are distinctly different from those created if the asphere were added near the end of the (local) design optimization. In theory, if global optimization is used (and works as desired), similar solutions should be identified regardless of what starting point design the aspheric surface is added to, provided the necessary variables are available. As the use of aspheric surfaces in optical systems has become more common, there have been increased publications on their use⁷¹ as well as courses on the subject.⁷² In addition to these resources, we refer the reader to the chapter by Shannon.⁷³

From a simple viewpoint, the location of an aspheric surface determines the effect it has on aberrations. By location, we are referring to the distance of the aspheric surface from the aperture stop (or pupils) of the system. An aspheric surface that is placed at the aperture stop will only affect the spherical aberration produced by the surface, while an aspheric surface that is placed away from the aperture stop will affect spherical aberration, coma, astigmatism, and distortion. One way to think of this is to consider the beam placements on the surface. At

the aperture stop, all the beams from the different field angles pass together through the stop. Thus, all fields will see essentially the same effect from the surface, affecting aberrations that do not depend on the field angle, such as spherical aberration. When the aspheric surface is away from the aperture stop (or its images, the pupils), the beams associated with different fields are separated on the surface and will each see a somewhat different effect from the aspheric surface. Thus, multiple aberrations can be affected.

There is no definitive rule for selection of which surface or surfaces on which to use aspheres. From the discussion above, it would make sense to place an asphere near the stop if we are attempting to control spherical aberration, and away from the stop if we are trying to control field-dependent aberrations, such as coma and astigmatism. Of course, in a multielement system, it is the combined aberration content that ultimately determines the image quality. If we place an aspheric surface on one element (for instance, near the stop), the ability to control spherical aberration through this asphere means that the other elements do not need to “worry” about spherical aberration as much, and we can change their shape to work on other aberrations. Thus, the addition of an aspheric surface on one element of the system can not only change that element but also potentially change the other elements. In general, placing an asphere on the surface with the largest beam extent will provide the maximum aspheric leverage.

In the design process of plastic optical systems, it may be useful to make one surface on each of the elements aspheric and allow the optimization of the lens design software to use them as appropriate. It is generally not good practice to make all of the surfaces aspheres from the start. In this case, the optimization algorithm may “beat” the aspheres against each other, introducing a large amount of aberration at one surface that is cancelled by a large negative amount of the same aberration at another surface. This is not a preferred design form because it will tend to be more sensitive to assembly tolerances of the system, such as decenter. If one aspheric surface per lens element is not providing adequate performance, additional aspheric surfaces may be added to determine any performance advantage.

When using aspheric surfaces in a design, a few changes in the setup used for spherical lens design are in order. First, the number of field angles defined should generally be increased. In a design using spherical lenses, it is common to have three or four defined field angles. In designs using multiple aspheric surfaces, it is possible for the lens to be well corrected at these three or four field angles and to perform poorly in between them. This can sometimes be seen by looking at the astigmatic field plots, which may curve repeatedly back and forth, crossing zero at each of the defined fields. By adding additional field angles, we can force the optimization algorithm to work on the performance across the entire image plane. We essentially do not provide the algorithm enough room in the field spacing to allow the performance to “wobble,” as it could with fewer defined fields.

In addition to increasing the number of defined fields, it is also useful to increase the ray density that the optimization algorithm uses. As in the case of too few field angles, if there are too few rays, they may be well corrected, but

other rays that are not evaluated may not be. This can result in a low merit-function value that suddenly increases when a denser ray grid is used. It is a good practice when using a large number of aspheres to increase the ray density and ensure that there is no significant change in performance.

A final adjustment when designing with aspheric surfaces is to either reduce the minimum improvement value that stops the optimization cycle or to repeatedly run the optimization algorithm. Sometimes, systems with multiple aspheric surfaces will stagnate during the optimization process. However, repeatedly running the optimization algorithm often will continue to drive the merit function value down.

Throughout the design process, or at least near its end, all aspheric surfaces should be evaluated against a “best-fit sphere.” A best-fit sphere is a spherical surface with the radius selected such that it has minimal departure from the aspheric surface. In the grinding and polishing of glass aspheres, the use of a best-fit sphere as a starting point minimizes the amount of material that must be removed to create the aspheric surface. In the case of molded parts, if a spherical base surface is used on the optical insert, it may determine how much nickel plating needs to be removed. An important reason to evaluate the asphere against a best-fit sphere is to determine whether the asphere is really necessary. Once aspheric terms are set as variables in the optimization algorithm within the lens design software, they will likely be used, regardless of how small the advantage gained.

If it is determined from comparison to the best-fit sphere that the asphere has a very small departure, then the aspheric surface should be made spherical and the design reoptimized to see if a performance difference results. This may seem at odds with our earlier statements about the cost of aspheres and spheres being the same. In actuality, we are not concerned about cost savings but in creating fewer potential problems by simplifying the system. To put it plainly, there is no need to make the system more complex than it needs to be, just because we can. A sphere, which is defined completely by its radius, is easier to describe than an asphere. This means that fewer terms need to be considered during the manufacture and testing of the surface. This results in less chance for errors due to interchanging digits during any entry of numbers, such as aspheric coefficients. Also, a sphere is easier to test than an asphere, giving more options for the vendor (or customer) to verify the surface.

In addition to comparing the aspheric surfaces to best-fit spheres, the designer should also evaluate the aberration produced by each surface. As mentioned earlier, having too many aspheric surfaces may simply result in them beating against each other. Most optical design codes will provide graphs and/or numerical values of the various aberration contributions of each surface. These can be used to determine if aspheric surfaces are working with or against each other.

While the optical design code can create any shape asphere allowed by the coefficients it can vary, the designer needs to ensure that the aspheric surface can be produced and tested. Most aspheres can be molded, but there are differing

levels of complexity. For instance, if the aspheric surface is so steeply curved that it rolls back past a line drawn perpendicular to its axis (that is, it goes beyond a hemisphere) it may not be able to be manufactured in a straight-draw mold. Also, steeply curved surfaces may be difficult to test, making it hard to verify that they have been correctly manufactured.

For the most part, the ability to use aspheric surfaces provides a positive impact on the design of plastic optical systems. However, there can be one particular drawback to using refractive aspheric surfaces (in plastic or glass), which is the chromatic variation of aberrations. Aspheric surfaces do not change the axial or lateral color introduced by a surface. If the asphere is being used to introduce spherical aberration with the intention of offsetting the contributions from other surfaces, it will also introduce spherochromatism, the variation of spherical aberration with wavelength. The amount of spherochromatism introduced will generally be the amount of spherical aberration created by the surface divided by the Abbe number of the lens material. In some cases, spherochromatism can be used to balance the effect of axial color. Other times, however, it may degrade an otherwise monochromatically well-corrected solution. The designer needs to keep in mind the purpose of each aspheric surface and what benefit it brings to the overall design. Aspheric surfaces that do not provide benefit to the design should be removed.

4.6 Diffractive Surfaces

Diffractive surfaces rely upon the wave nature of light to perform their function. They typically consist of some form of microstructure on the optical surface. As a beam of light passes through the diffractive surface, the various microstructure features impart phase delays (usually 2π) to the different portions of the incident beam. As the beam propagates, these different portions interfere with each other, similar to the way that waves in water can combine to form regions of larger or smaller waves. By correctly designing the diffractive microstructures, we can get a desired beam from the surface in a manner similar to designing an aspheric surface to create or eliminate specific aberrations.

Diffractive optical surfaces are used for a variety of purposes, such as fan-out gratings, athermalization, and beam shaping. In many plastic optics imaging systems employing diffractive optics, they are used for color correction, that is, to control chromatic aberration. In this section, a brief overview of their use for this purpose is given. Readers interested in a more thorough discussion of diffractive optics are referred to Ref. 74.

Probably the best-known diffractive surface is the diffraction grating. A diffraction grating normally consists of a substrate with a pattern of parallel-ruled lines on it. Diffraction gratings are often used to separate wavelengths of light, sending each wavelength away from the surface in a different direction. The diffractive surfaces used for color correction are somewhat similar to a diffraction grating, with the difference that instead of a series of parallel-ruled

lines, the diffractive-color-correction surface (from here forward called a “diffractive”) typically consists of a series of concentric rings. Unlike the equal spacing of lines often seen on diffraction gratings, the rings on a color-correcting diffractive surface are not normally equally spaced. The diffractive configuration most often seen has its rings becoming more closely spaced as we move out radially from the center of the surface.

Even though they rely on the wave nature of light, we are usually able to design and analyze diffractive surfaces using rays. This is helpful, as most optical design software specializes in ray tracing. When using a refractive or reflective surface, the path of the rays is determined by the use of Snell’s law. When using a diffractive surface, the path of a ray is determined instead by the grating equation, which is shown in Eq. (4.8) for the case of normal incidence:

$$m\lambda = d \sin \theta, \quad (4.8)$$

where m is the order of the diffractive, λ is the wavelength of light, θ is the output angle, and d is the grating spacing. Those familiar with diffraction gratings will recognize this equation. Figure 4.16 shows a ray passing through a diffractive surface with the various parameters labeled.

We see from Eq. (4.8) that there is a parameter m , which defines the order of the diffractive. The fact that the equation can have multiple solutions for different values of m , the order number, means that we could have multiple beams coming

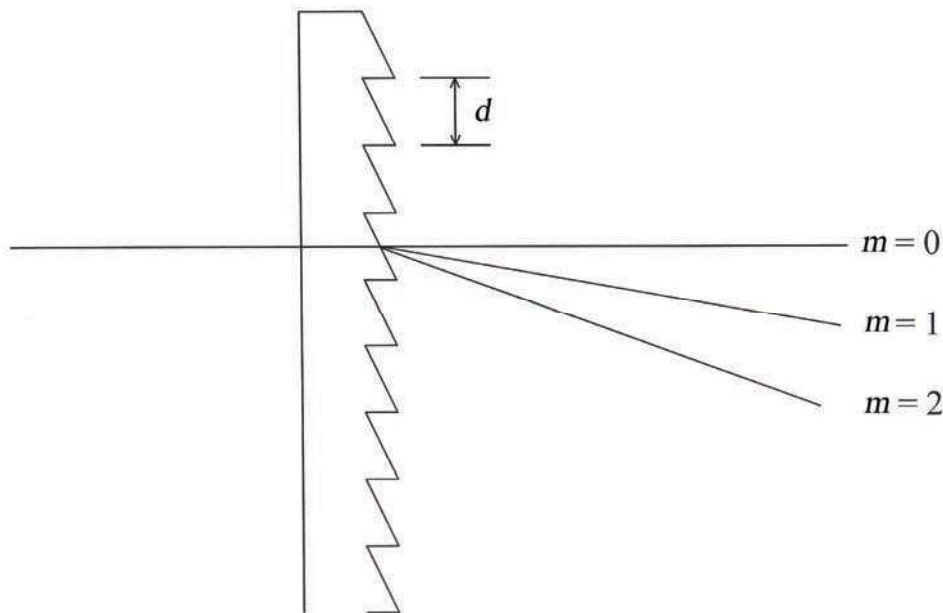


Figure 4.16 Ray passing through a diffractive surface.

from the surface. For instance, we may have one beam associated with the zero order, one beam for the first order, and another beam for the second order, as is shown in Fig. 4.16. The amount of light that is in each order is known as the diffraction efficiency of the order. For color correction in imaging systems, we typically want all of the light in a single order. The preferred diffraction order is known as the design order. In the optical design process, this order is usually taken to be the first order ($m = 1$). The amount of light in each order depends upon the shape and depth of the diffractive groove, as well as on the wavelength of light. The depth of the diffractive groove is usually selected for a wavelength somewhere near the center of the spectral band. For a diffractive surface in air, the depth of the diffractive groove is typically taken as

$$h = \frac{\lambda_0}{n-1}, \quad (4.9)$$

where h is the depth of the groove, λ_0 is the design wavelength, and n is the refractive index of the diffractive substrate for the design wavelength. Because the groove depth is set for one wavelength, the diffractive will not work perfectly for other wavelengths (except in special cases). That is, the diffraction efficiency in the design order will only be 100% (in theory) for a single wavelength. The diffraction efficiency for a given wavelength depends on the ratio of the design wavelength to the wavelength of interest. The efficiency of any wavelength, $\eta(\lambda)$, can be calculated from Eq. (4.10):

$$\eta(\lambda) = \left(\frac{\sin \left\{ \pi \left[\left(\lambda_0 / \lambda \right) - m \right] \right\}}{\pi \left[\left(\lambda_0 / \lambda \right) - m \right]} \right)^2, \quad (4.10)$$

where λ_0 is the design wavelength, and m is the design order of the diffractive. Figure 4.17 shows a plot of (normalized) diffraction efficiency using this equation for a design wavelength of 550 nm and a design order of one. Imperfect diffraction efficiency (less than 100%) for wavelengths other than the design wavelength results in the transfer of energy, as a function of wavelength, into other diffractive orders. The light in these other orders will continue to propagate through the remainder of the optical system and can end up as stray light or as ghost images.

As an optical element, diffractive surfaces have the interesting property of possessing a negative Abbe number, which we have not seen to this point. For the visible wavelength range, the Abbe number of a diffractive surface is -3.45 . We discussed earlier that the Abbe number describes how much the refractive index of a material changes with wavelength, and that lower Abbe numbers mean that the material is more dispersive. A more dispersive material results in a larger difference in direction between two colors, for instance red and blue. We also

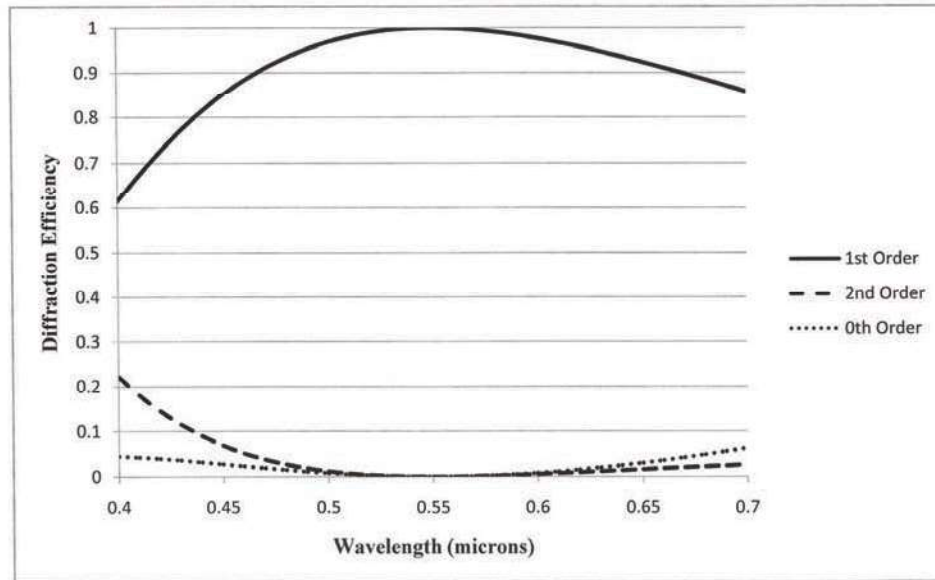


Figure 4.17 Plot of diffraction efficiency versus wavelength for a first-order diffractive design.

saw earlier that in a typical positive lens, the blue light would focus closer to the lens than the red light. For a (positive-powered) diffractive surface, the opposite is true: the red light will focus closer to the surface than the blue light. This is what is meant by the negative sign of the diffractive Abbe number. The red light focuses nearer to the surface than the blue because the light paths for a diffractive are determined by the grating equation instead of by Snell's law.

Knowing this, it is a logical step to consider combining a refractive and diffractive surface to get the red and blue light to focus together, eliminating axial color. Just as we can combine lenses of two different glasses to form an achromatic doublet, we can combine a diffractive surface with a refractive lens to form what is usually called a "hybrid lens." Also similar to a standard doublet, we do not necessarily need to fully correct the chromatic aberration of the combination. In this way, we are essentially creating a new glass, one that has the desired chromatic properties; Stone and George published a study on this subject.⁷⁵ Using a diffractive surface, we can do this without having to add an additional lens element. This is particularly useful for plastic optical systems because there are a limited number of materials to choose from, and constraints often limit the number of lens elements allowed in a system.

The design of diffractive surfaces using optical design software is fairly straightforward. The diffractive surface can be described in several ways; the most common of which is by using a polynomial description, similar to the one used for even-polynomial aspheric surfaces. It should be noted that the coefficients for the diffractive surfaces usually start at the second order, as opposed to the fourth order often seen on polynomial aspheres. This is an

important distinction, as the second-order coefficient is used to provide optical power, while the higher-order coefficients provide aberration correction. The use of the second-order diffractive term is typically all that is needed to control chromatic aberration. The use of the higher-order diffractive coefficients is generally not recommended unless they bring significant improvement to the design. Some vendors are unfamiliar with diffractives using higher-order coefficients, which can lead to confusion in their manufacture. If higher-order diffractive terms are used, extra attention to detail is warranted.

Similar to the effect of location of an aspheric surface, the position of a diffractive surface with the system will determine its effect on chromatic aberration. If the diffractive is positioned at the stop (or at a pupil of the system), it will only affect the axial color. If it is positioned away from the stop, it will affect both axial and lateral color. Because of this, engineers sometimes design systems with multiple diffractive surfaces, similar to using multiple aspheric surfaces. This is generally not a wise design practice. We have already discussed how using a diffractive element will result in losses due to diffraction efficiency. Using multiple diffractive surfaces will only compound this effect.

Also similar to designing with aspheric surfaces, it is easy to let the optimization algorithm run wild with the polynomial coefficients when using a diffractive in a design. Many a designer has had the sinking feeling that comes from seeing what looks like a great design solution turn out to be essentially worthless, due to the unrealistic groove spacings on the diffractive. In most design codes, minimum diffractive groove spacing can be set as a constraint on the optimization. The designer needs to keep an eye on the groove spacing as the design progresses to verify that it is not becoming too small. Most design codes have commands to plot the diffractive feature size as a function of radial location.

On the other hand, the designer also needs to ensure that the diffractive spacing does not become too large, at least in comparison with the beam size on the surface. Because the diffractive works by the interference of beams from different portions of the microstructured surface, it is important that the beam associated with each field angle covers several diffractive grooves. In most plastic optical systems this is not a concern; having too small a groove spacing is more common. However, unless told to do so, the design code will not be concerned with groove spacing. It relies on the polynomial coefficient representation of the surface and is not concerned with the effect of turning it into a surface relief pattern.

To convert from the polynomial description to an actual microstructured surface, the sag of the diffractive surface is first computed. Then, in a manner similar to producing a Fresnel lens, the surface is “collapsed” with a step occurring each time the sag reaches a multiple of the diffractive step height, which is typically the value shown in Eq. (4.9).

When calling out diffractive surfaces on a drawing, it is good practice to show a profile of the diffractive surface, often with the groove depth exaggerated,

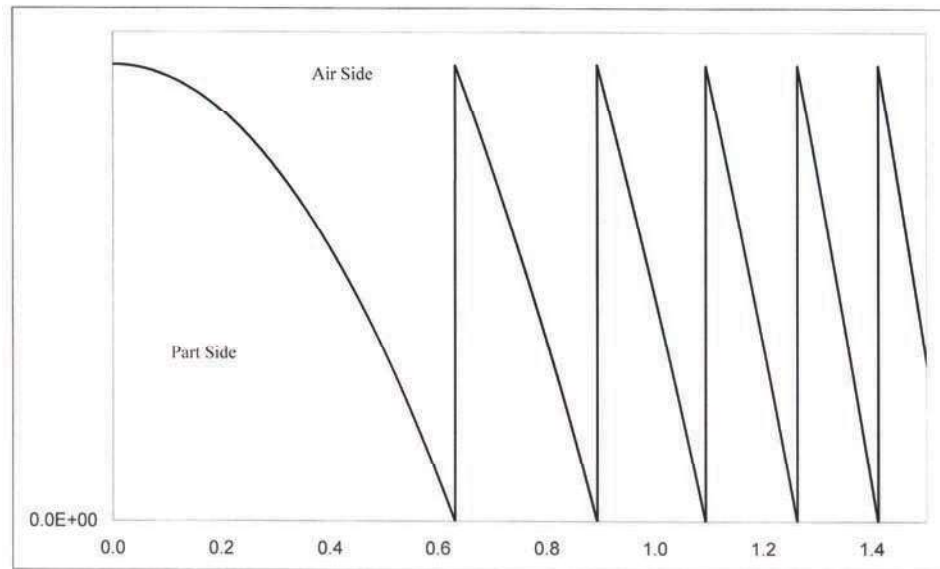


Figure 4.18 Example of diffractive profile showing step orientation.

and with labeling indicating the “part side” and the “air side.” This can prevent confusion as to the correct orientation of the diffractive steps. An example of this is shown in Fig. 4.18.

Diffractives, or the microstructures that form them, can be manufactured on planar, spherical, or aspheric base surfaces. The optical design codes support each of these surface types. The decision as to what type of base surface to use will depend in part on how the diffractive master is made. If the master is to be made by diamond turning, which is most often the case for the molded chromatic-controlling diffractives discussed here, the surface can be any form that can be diamond turned. For other types of diffractive functions, such as fan-out gratings, the master may be created through a lithographic process that typically works best with planar surfaces.

We mentioned earlier in this section that the first order is normally selected as the design order. However, there are cases where a larger design order is intentionally selected. These types of diffractives, known as multiorder diffractives (MODs) or harmonic diffractive lenses (HDLs), will be discussed in one of the design examples.

4.7 Athermalization

We discussed earlier that one potential downside to the use of plastic optics is that a focus shift may occur over temperature due to the thermal properties of the materials. Plastic optical materials have a coefficient of thermal expansion (CTE) and a change in refractive index with temperature (dn/dt) that are much larger

than that of glass optical materials. These properties make it more difficult to design a system that does not significantly change performance over temperature.

From the selfish viewpoint of the optical designer, the easiest way to deal with a potential focus shift due to temperature changes is to make it someone else's problem by requiring the use of a manual focus adjustment or an autofocus system. Of course, if the optical designer is also the optomechanical designer, the feeling of happiness at having shifted the problem may be short lived. While the use of manual or autofocus systems can greatly reduce the demands on the optical design over temperature, in many systems using plastic optics these are not viable solutions due to cost constraints, packaging issues, or customer desires. In this case, the designer needs to work towards developing an optical system that can work over temperature without relying on focus adjustment, that is, one that is athermalized.

Several methods are commonly used to athermalize systems that incorporate plastic optical elements. One popular method is to combine elements made of glass with elements made of plastic. This is sometimes referred to as a "hybrid" lens or as a plastic-glass design. This method can be thought of as trying to use "the best of both worlds." We know from our earlier discussions that glass lenses have a much smaller thermal focus shift than plastic lenses do. We also know that it is much easier and more cost effective to create aspheric surfaces on a plastic element. Thus, we can use a conventional (spherical) glass element to provide most of the optical power, which limits the focus shift, and use low-power aspheric plastic lenses to provide aberration correction.

Another method, which due to terminology can be confused with the method above, is the use of a diffractive surface to athermalize a design. Lenses with diffractive surfaces are also sometimes referred to as hybrid lenses. Unlike the previously discussed use of a diffractive to correct chromatic aberration, the diffractive in this case changes power over temperature to compensate the power change in the base refractive lens. Similar to equations that can be used in the first-order design of an achromatic doublet, a set of equations can be developed for the first-order design of a temperature-corrected refractive/diffractive lens, often referred to as an athermat.⁷⁶ Because of the chromatic dependence of the diffractive element, singlet athermats are typically used with narrow wavebands or in laser-based systems. It is possible to extend the waveband by using the diffractive surface in combination with a doublet. The reader is cautioned that several patents exist that utilize this athermalization method.^{77,78}

The use of plastic mirrors is another method of athermalization. It should be noted that an all-reflective system made of a single material is inherently athermal because a temperature change simply scales the entire system up or down for this type of design. It is as if the entire system was put into a 3D copy machine, and a slight scale factor applied. It is also possible to combine reflective and refractive elements, which is known as a catadioptric system. Since reflective elements do not rely on the refractive index of the base material, the large change in index associated with optical plastics does not affect the power of the mirror, only the coefficient of thermal expansion does this.

Another athermalization method, when using all plastic elements, is to balance the power changes amongst the elements such that they (partially or completely) compensate for each other over temperature. This method is similar to what is performed on many glass element systems. It is important to remember that when the temperature of the optical system changes, not only are the lenses affected, but also any associated spacers or mounting structures. This may cause the spacing of the lenses to change as well as the location of the detector or film that will capture the image. What is usually important is not that the image stay in the same location over temperature, but that the image remain focused on the detector over temperature.

Understanding that the image may move, with a corresponding motion of the detector, allows for the athermalization of optical systems through the appropriate selection of materials. For instance, returning to the acrylic lens discussed in Chapter 2, we found that for a 50 °C temperature increase, the focus position of the lens shifted by 0.681 mm. If we were able to have the detector move by the same amount, the image would remain in focus over the temperature change. In that case, the image plane was originally 47.3 mm behind the rear surface of the lens. To achieve a 0.681-mm-length increase from a 47.3 mm length (over a 50 °C temperature change) would require a material with a CTE of about 288×10^{-6} per °C. This CTE is much higher than what is seen in most materials, so there may be no direct solution in this case. However, if we were to use multiple lenses and reduce the amount of focus shift, a suitable (and available) material could potentially be found.

If the desired CTE falls between that of two available materials, it is possible to combine different lengths of the two materials to achieve an effective CTE matching the desired value. Putting an optical spin on this, it is similar to combining two glasses in a doublet to create an effective glass with the desired chromatic characteristics. In addition to material selection and combination, mounting configuration can also be used to control thermal performance.⁷⁹ The use of proper optomechanical material selection and mount design, combined with the use of appropriate lens powers and optical materials, is the most frequently used method of athermalizing systems using plastic optical elements.

Regardless of which athermalization method is selected, the design and analysis of the system over temperature must be performed. This is typically conducted through a first-order analysis and/or by using optical design software. A first-order analysis is usually concerned with the lens power and image location and not with the image quality. In the case of singlet athermats, the first-order analysis would entail using the athermat equations to determine the necessary powers of the refractive and diffractive surfaces. For the method of balancing the power changes of the various refractive elements, similar equations can be developed. In some cases it may be useful to have a spreadsheet that calculates the power change of each element as well as the power change of the entire system and the image location, as a function of temperature. After the initial analysis, optical design software may be used to further develop the design.⁸⁰

For many systems, there is a requirement for both ambient image quality as well as (possibly reduced) image quality over temperature. We discussed some design techniques for controlling aberrations, which may limit image quality, earlier in this chapter. Having a requirement for performance over temperature can be thought of as an additional constraint on our image quality requirement. Thus, one way that we can attempt to fulfill both requirements during the design process is to add the thermal performance as a constraint in our optical design software model. This is typically done through the use of a multiconfiguration or “zoom” model. These models are normally used in the design of zoom lenses, which work over a range of focal lengths and object distances. In our thermal case, instead of having the system zoomed over focal length, we zoom it over temperature.

A fairly standard technique to do this is to create three zoom positions, also called a three-configuration model. The first zoom position models the lens system at ambient temperature, while the second position represents the lens system at an elevated (hot) temperature, and the third zoom position represents the lens system at a lower (cold) temperature. The various element parameters are linked across the zoom positions, such that their values are those associated with the position’s temperature. For instance, if we consider the radius on a surface in the first zoom position, the same surface in the second zoom position will have a slightly larger radius value. Similarly, the surface in the third zoom position will have a slightly smaller radius value, due to the change that would occur from the lens material’s CTE value in conjunction with the temperature difference between the respective zoom positions.

Linking the various parameters across the zoom positions in this way is known as “picking up” the parameter. As a linked parameter (for instance, a radius in zoom position one) is varied by the optimization algorithm, the radius values in zoom positions two and three will change accordingly, keeping the correct scale factor with respect to the value in zoom position one. The merit function of the lens should consider the image quality across all three zoom positions. If there is a performance relaxation at the temperature extremes, the zoom positions representing these temperatures can have a reduced weight in the merit function. While the radius scale factor depends on the material CTE and temperature, the scale factor for the refractive index, which also must be adjusted across the zoom positions, would depend on the dn/dt of the material and the temperature.

When using aspheric and diffractive surfaces, care must be taken with the scaling of the aspheric and diffractive coefficients. Aspheric coefficients typically scale with a power of one less than the term the coefficient is associated with, while diffractive coefficients (depending on the representation used in the program) may scale with a power one less than that of the associated term or with a power equal to that of the associated term. Simple tests should be run to verify that the appropriate scaling is taking place. It should also be noted that conic constants do not scale with temperature but remain at a constant value. As such,

if conic constants are used in the description of the surfaces, their values should be equal across the zoom positions.

The temperatures for the zoom positions are usually taken at the extremes of the temperature range the system must operate over, though they need not be. Some material properties can vary with temperature, so different CTEs, dn/dts , etc. may be associated with different zoom positions. In this case, depending on the complexity of the system, or the size of the temperature range, three zoom positions may not be adequate, requiring additional positions to be added. Similar to our discussion of the need to add extra field angles when using aspheric surfaces, before completing the design, the designer should evaluate temperatures in between the zoom positions to verify performance throughout the entire temperature range.

In addition to material properties, attention must be paid to the mounting features of the elements. Most optical design codes consider thickness based on the distance between the vertices of adjacent optical surfaces. However, this thickness may not be the appropriate distance to use in the calculation of lens position shift due to thermal expansion, as the lens is not normally attached to the barrel by its vertex. Consider the case of the lens shown in Fig. 4.19. The lens barrel length is different from the thickness between the lens vertex and the end of the barrel. In this case, the expansion of the lens barrel material, as well as that of the lens material, must be taken into account. This is often performed in the lens design program by adding “dummy” surfaces, which take the mounting features into account.

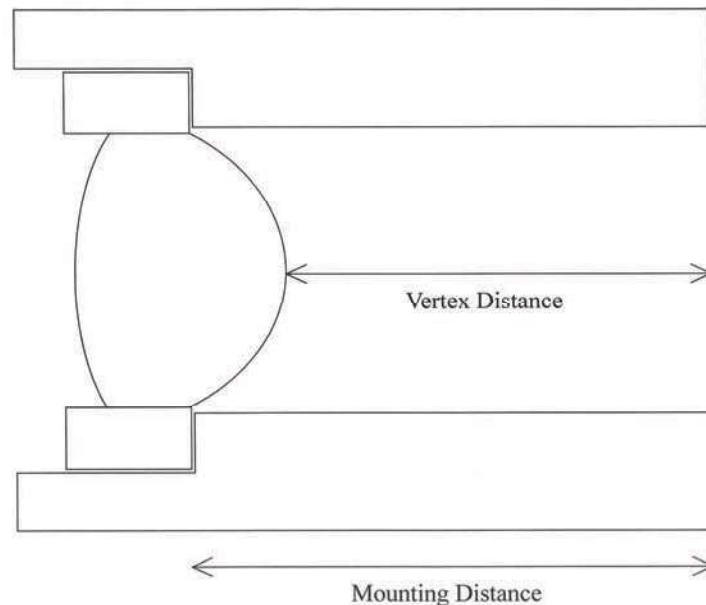


Figure 4.19 Difference in mounting distance versus vertex distance, which must be accounted for in athermalization optimization.

Once the multiconfiguration model has been created, optimization occurs in a way similar to that of a single-configuration system. Because of the additional thermal performance constraints, the design process may take more designer intervention than usual. The designer may need to change the lens or spacer materials or add additional elements to provide more variables for the optimization algorithm. Again, the software should not just be “left to its own devices;” the designer must remain involved in the design process.

Although not explicitly stated, the discussion on athermalization in this section has been based on the assumption of a thermal soak, which means that all of the system is at a uniform temperature. In reality, this may not be the case, as temperature gradients may exist throughout the system. This is most likely to be seen in systems subjected to rapid temperature changes, such as military hardware. In many cases, due to variations and permutations of operating conditions, no single thermal gradient model may cover the entire spectrum of thermal gradients that may be seen. The design and analysis of systems under the influence of temperature gradients is more complex than systems assumed to be under thermal soak, and we do not discuss them here. However, once an understanding of thermal soak design and analysis is obtained, the extension to thermal gradients, though they require more complex models, is a fairly straightforward and logical one.

4.8 Coatings

Coatings on plastic optics are typically used for one of two purposes: either to perform an optical function, such as reducing the surface reflectance, or to improve the mechanical or chemical properties of the surface. In some cases, multiple coatings are applied to achieve both purposes. Most of the development of coatings for plastic optics has occurred in the last 30 years, with increased improvements made in the last decade or so. The early development of coatings for plastic optics was driven by the ophthalmic market. The increased use of polycarbonate eyeglass lenses in the 1980s created demand for antireflection coatings that would work with that material. In addition, coatings to improve the scratch resistance of the lenses were also desired. As a result of this early demand, as well as the increased use of polycarbonate in the automotive market, coatings for polycarbonate are probably the most developed.

With the introduction of new materials, such as the cyclic olefins, and the increased use of plastic optics in general, there has been a renewed push for the development of coatings for plastic optics. The main issue with the development of optical coatings for plastic elements is the plastic materials themselves. As we mentioned previously, plastic materials have lower deflection and melting temperatures than optical glasses. As a result of this, standard processes used in the application of optical coatings to glass lenses cannot be employed with plastic lenses.

When applying optical coatings to glass elements, the process normally takes place at an elevated temperature. This temperature activates the surface the coating will be applied to, producing coatings that have stable optical and mechanical properties. The temperature required to properly activate the surface is higher than the deflection temperature of most of the optical plastics. This would result in deformation of the plastic optic if this standard coating process were applied.

There are three main coating deposition methods for use with plastic optics. These are physical vapor deposition, chemical vapor deposition, and wet processes (spin or dip coating). The first two methods are most commonly used. Since elevated temperatures cannot be used to activate the surfaces of optical plastics, alternate methods of surface activation have been developed, such as the use of ion guns and plasmas.

A commonly used coating for plastic optics is an antireflection (AR) coating. We saw in our earlier discussion of material properties that uncoated plastic optics have transmissions in the low 90% range. This is due to the reflection loss at each of the surfaces of the element. For an uncoated surface, the amount of reflection at an air/plastic interface, for normal incidence, is given by

$$R = \frac{(n-1)^2}{(n+1)^2}, \quad (4.11)$$

where n is the refractive index of the material. Using an index of 1.492, that of acrylic, results in a single surface reflectance loss of 0.039 or 3.9%. If we had a three-element lens system (six surfaces), the transmission after only taking the surface reflection losses into account would be 78.8%. This transmission value can be increased by using AR surfaces on some or all of the lenses.

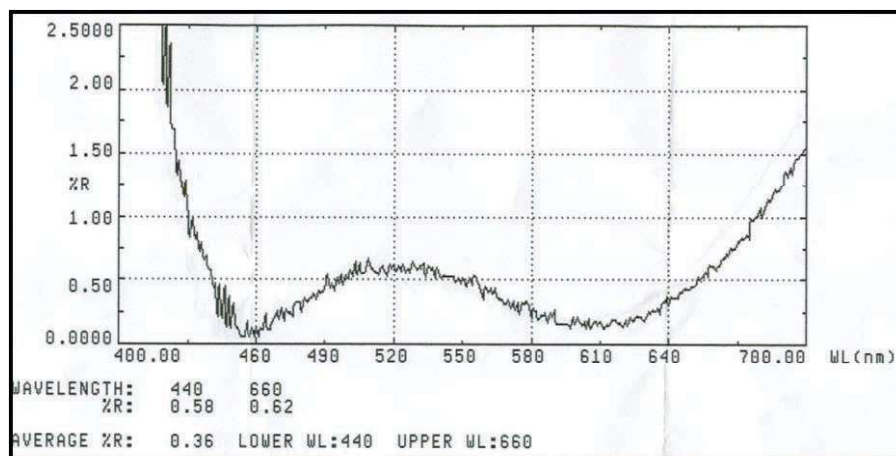


Figure 4.20 Measured performance of a typical multilayer AR coating. (Figure courtesy of Alan Symmons.)

Most coaters have a catalog of coatings to choose from, depending on the required performance.⁸¹ The rule of thumb for AR coatings (and most coatings in general) is that the wider the waveband desired, and the better the performance desired, the more complex the coating. Figure 4.20 shows the measured performance of a typical multilayer AR coating for the visible waveband. In addition to AR coatings, other optical coatings are also available, such as bandpass filters, infrared blocking filters, and reflective coatings.

The use of reflective coatings in the production of plastic mirrors is quite common in a number of applications, such as medical devices and the automotive industry. The material used for the reflective coating depends on the waveband of interest. Aluminum and silver are used for the visible waveband, while gold is often used for the infrared.

In addition to these optical functions, coatings on plastic optics can be used for other purposes. Because of the lower scratch resistance of optical plastics compared to optical glasses, hardcoats can be applied to the plastics to improve their handling performance. In cases where the optical components or housing surround sensitive electronics, conductive coatings can be applied to prevent the buildup of static as well as to reduce electromagnetic interference. Coatings can also be applied in order to reduce absorption of water by the plastic substrate.

Coatings for plastic optics can be specified in the same manner as coatings for optical glasses. The coatings are often called out to meet the adhesion, moderate abrasion, and humidity requirements of MIL-PRF-13830. One particular problem with coatings on plastic optical elements can be crazing due to thermal cycling, where crazing is the development of a network of fine cracks. As noted previously, plastic optical materials have a larger coefficient of thermal expansion than optical glasses. This results in a greater expansion over temperature, which the coating must be able to deal with. As part of the coating specification, a temperature range as well as the maximum temperature ramp (change rate of temperature) should be included. A thermal cycling test can also be included.

Proof of coating conformance is often obtained through the use of witness samples, small disks, or wedges that are coated together with the parts. The transmission or reflectance tests as well as the adhesion, abrasion, humidity, and temperature testing are typically performed on the witness samples. Of course, it is the performance of the coating on the optical parts that ultimately matters. In order to get the best correlation between part and witness sample, the witness samples should be made of the identical material as the parts and processed in the same manner as the parts.

Most producers of plastic optics either have coating equipment and capability themselves or work closely with a local company that has it. In general, it is preferred that the producer of the parts be held responsible for their coating, as opposed to the producer shipping the parts to the customer, who then sends them out for coating. This reduces the potential for contamination of the parts before coating and reduces finger pointing in the event of an unsuccessful coating run.

The primary downside to the use of coatings is their cost. When using plastic optics to reduce cost, the need for coatings can limit the cost savings. The application of coatings on plastic optics can increase their price anywhere from 10% to 100%. Coating runs, regardless of the number of parts coated, have a somewhat fixed price due to the need to pull a vacuum on the coating chamber. In order to reduce cost, as many pieces as possible should be coated in a single run. If there is considerable risk in the coating run, the parts may be divided into two runs in order to improve the odds of at least one batch of parts yielding.

For the reader interested in additional information on coatings for plastic optics, a thorough review, with extensive references, can be found in the chapter by Schulz.⁸² More recent articles can also be found in journals and in conference proceedings.^{83,84}

In addition to surface coatings, changes in spectral performance of plastic optics can also be achieved through the use of dyes, which can be mixed in with the plastic pellets during the injection molding process. A well-known example of this is the purplish windows seen on TV remote controls. The purplish color results from the absorbing dye in the material. In this case, the dye is used to block the visible band while allowing the near infrared (NIR) band to be transmitted. The complementary dye is sometimes used in optical systems employing complementary metal-oxide semiconductor (CMOS) sensors, such as Webcams, which are sensitive out to the NIR band. In this case, the near infrared is blocked and the visible spectrum is passed.

The main concern with these dyes is the sharpness of their spectral characteristics. The dyes tend not to have sharp cutoffs and cut-ons. Instead, these features tend to be more of a rolling change. This may be a problem in systems that require a well-defined transition between transmission and absorption/reflection. The dyes can also sometimes be sensitive to the thermal environment of the molding machine and may change property if the residence time is too long. However, the use of these dyes, in combination with the surface coatings above, may result in a cost savings by reducing the complexity of the coating required.

4.9 Optomechanical Design

The goal of the optomechanical designer is to adequately maintain the position (and shape) of the optical elements in a system over the various environments the system is subjected to. By “adequately,” we mean that the positions and shapes of the elements can change but only by an amount that ensures the system performance requirements are still met. The characteristics of plastic optics, such as molded-in flanges and integral mounting features, can provide the optomechanical designer with freedom and room for creativity in his optomechanical design.

The optomechanical designer should be involved early in the design process. Too many times, the optical engineer performs their design work, then “throws

the design over the wall” to the optomechanical designer. Working in this fashion often produces a nonoptimal design. The best overall system designs typically result from having the optical, optomechanical, and manufacturing engineers (potentially the molder) all working collaboratively throughout the design process. Other disciplines, such as system or electrical engineering, should be included as well.

In some areas, the optomechanical design and analysis of plastic optical systems is similar to that of glass optical systems. However, there are some significant differences that must be taken into account. The main differences between optomechanical design for glass and plastic optical elements are the material properties of the elements themselves and the manufacturing methods used to produce them. Plastics tend to have higher fracture resistance than glass but a lower hardness. Plastics have a higher coefficient of thermal expansion than glass, which must be taken into account when considering the temperature range the system must operate over. Plastics also have lower service temperatures than glass, which can lead to permanent deformation of the optics under certain conditions. Additionally, plastic materials will usually have some amount of out-gassing (that new car smell) that must be considered.

Most glass optical elements are still produced by grinding and polishing, although the use of glass molding is increasing. These manufacturing methods tend to limit the optomechanical features that can be incorporated in the glass elements. Plastic optics, on the other hand, can be produced by several methods, with molding being the most popular. This allows the integration of a variety of optomechanical features into the parts.

We begin by discussing typical optomechanical features and concerns with plastic optic elements. If the parts are to be injection molded, there are two main considerations. First, there must be a gate to allow the plastic to be injected into the mold cavity forming the part. For most plastic optical parts, the gate is positioned on or near the diameter of the part. From the optomechanical designer’s viewpoint, the gate on a plastic optic part usually results either in a protrusion from the diameter or necessitates a cutout on the diameter of the part. As discussed earlier, as the part comes out of the mold, it is usually attached to a runner system by the gate. The part is then degated, with residual material (known as gate vestige) remaining. This gate vestige is the reason that gate flat cutouts are often used on molded parts. The flat should be set low enough to provide room for the expected gate vestige. The amount of vestige may depend upon the degating method used. Some of the different methods are the use of a pair of side-cut dikes, a hot knife, ultrasonics, or laser systems. If a gate cutout is not desired, the gate protrusion from the part can be used as an alignment or antirotation feature for the element. If neither flat nor protrusion is desired, the gate can be milled off, leaving a fairly clean diameter. It should be kept in mind that the preferred location for a gate is directly above the optical element, as opposed to being positioned out on the end of a flange being used as a spacer. The molder will normally want to fill the optical areas first and then fill any other features.

The second consideration for molded parts is the potential need for draft on the part to enable it to be ejected from the mold. Draft, as discussed earlier, is the angling of surfaces oriented in the direction of the mold-opening motion. The amount of draft required can depend on the size of the part, the length of the surfaces parallel to the mold draw direction, and the plastic material that is used. In some cases, the optomechanical designer may not want draft on the outer diameter of the part, especially if the outer diameter is being used as a centration feature. However, the molder may feel he cannot produce the part without some draft. One way to achieve a compromise in this area is to use a relatively small no-draft zone on the outer diameter. The no-draft zone, which can be used as the diametral locating feature, can be situated on the part between two drafted zones. In this way, most of the diameter of the part is drafted, making the molder happy, and a defined diameter is available to the optomechanical designer. Of course, a small no-draft zone typically won't be able to control tilt of the optical element, which must be taken into consideration in the overall design.

Another guideline for optomechanical design is to avoid sharp corners. For instance, if an eyepiece lens for a microdisplay is being designed, the lens may be made rectangular, instead of circular, to reduce its volume and weight. The corners on the rectangular piece should be designed as rounded as opposed to sharp. This will help with stress and shrinkage in the part.

Flanges are often incorporated into plastic optical elements. These flanges may be used to set the spacing between adjacent elements or between elements and apertures or other features. It is good practice, when possible, to have the flange on a plastic optical element extend past the highest point, which may or may not be the vertex, on an optical surface. By doing so, the flange can protect the surface from damage when set in a tray or during assembly. In some cases, it may not be possible to extend the flange beyond both optical surfaces. If the flange on at least one side of the element can be extended past the highest point on the surface, it will increase the ability to safely handle the part.

The use of flanging to set spacing between elements is common, but it must be tempered by the realities of the production process. Long, thin flanges can be difficult to fill out during the molding process and can affect the quality of the optical surfaces. In cases where the spacing between elements becomes significant, alternate mounting methods (such as the use of opaque molded spacers) should be considered.

The design of the individual components should be conducted with an understanding of the overall system approach as well as the required positional tolerances of the elements. There are several methods commonly used for assembling multielement plastic optical systems. One method is to use a clam-shell design. In this design, the barrel is made of two halves that are often identical. The two halves are formed by dividing the cylindrical barrel by a plane containing the axis of the barrel. The optical elements can be loaded into the lower half of the barrel, and the top half is brought down to seal the assembly. The two halves can be connected in a number of ways, such as screws, bonding,

or with an external sleeve. Features can be integrated into the barrel halves to locate the optical elements, or separate spacers may be used.

Another common method of assembling a multielement system is the use of a single-piece stepped barrel. An example of this is shown in Figs. 4.21 and 4.22, with the first figure showing an exploded view, and the second showing a cross-sectional view. A single-piece barrel assembly has several attractive features. By using a rotationally symmetric barrel, the concentricity of the steps that the lenses mount on can be held quite tightly. This is because the inner barrel features can be formed using a single pin, which is fabricated by diamond turning or precision grinding. Both of these methods, performed on high-precision spindles, produce highly concentric features. The limitation of the concentricity of the features in the molded part will be due to the molding process (such as its ejection), not by the fabrication of the tool.

There is often a debate as to whether the outer diameter of the elements, the inner diameter of the part that holds it (in this case the barrel), or both need to be a specific size. In some instances, the lens mold is machined and the diameter of the lenses that are produced from it is measured, and the complementary part (such as a barrel) is adjusted to the lens. In many cases, this may be the preferred method. It may be easier, in the precision grinding of the holes that form the lens diameters for a multicavity tool, to get repeatable size as opposed to a specific size. With multiple-element systems, multiple molds are typically built, whereas a single barrel tool will often suffice. The manufacturing method of the barrel insert also lends itself to precise finishing.

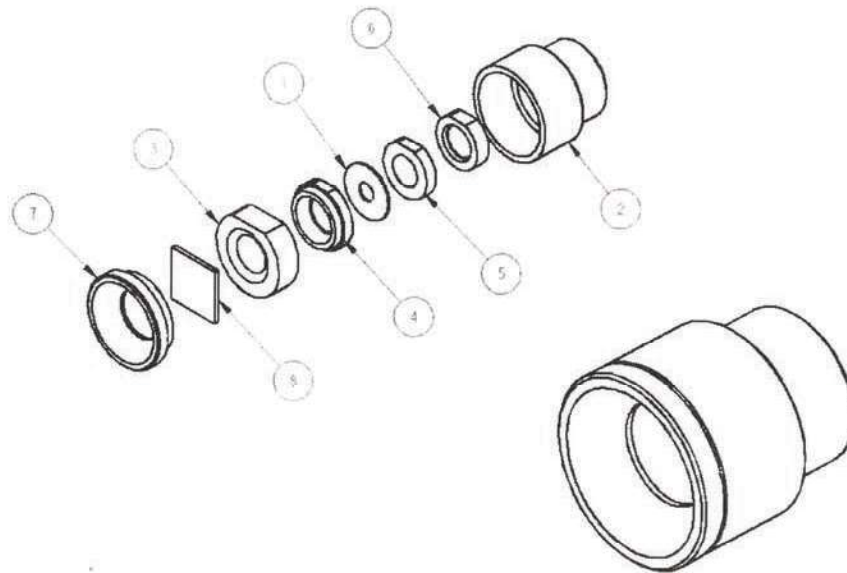


Figure 4.21 Exploded view of a single-piece barrel lens assembly. (Figure courtesy of Paul Merems.)

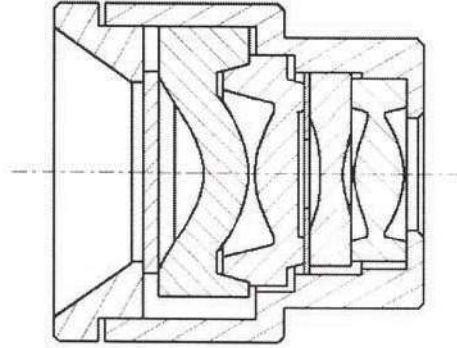


Figure 4.22 Cross-sectional view of a single-piece barrel lens assembly. (Figure courtesy of Paul Merems.)

The use of a single-piece barrel also allows for easy automated or manual assembly, allowing the lenses to be stacked upon each other, along with any apertures, spacers, or stray light baffles. As can be seen in Fig. 4.22, the flanges of the lenses can be designed to interface with each other (for example, through the use of tapers). Using tapers to provide centration between two elements can be a useful technique. It must be remembered that tapers, while providing centration, do not provide accurate spacing at the same time.

In the design of a part that requires features to be accurately aligned to one another, it is best to keep all of the features on one side of the mold parting line. For instance, in the case of the single-piece barrel discussed above, all of the lens steps, where the lenses are mounted, would be in a single side of the tool. Similar conditions exist for lenses. If a taper is to precisely align the lens relative to a specific optical surface, it will be best to have the taper and surface on the same side of the mold.

The optomechanical analyses that are performed on plastic optical systems are similar to those performed on glass optical systems. Evaluation of the distortion of the elements under temperature is a typical analysis. The distortion due to thermal expansion (which again is larger than glass), with the constraints of the mounting structure, must be considered. Distortion plots or coefficients can be provided to the optical designer, who can assess the impact on the optical performance. Deflection under vibration is another standard analysis. Bending of the structure can result in changes in the line of sight of the system. In certain cases, stress-induced birefringence must also be considered.

The final considerations for optomechanical design are configuration control and agreement between models. In developing the optomechanical model, the designer typically receives a listing of the optical model (the prescription) as well as an Initial Graphics Exchange Specification (IGES) file of the system. For plastic optical systems, particularly those using aspheric surfaces, importing the

IGES file into the CAD model will not adequately represent the optical surfaces. These should be modeled in the CAD program using the prescription provided.

Optical design codes usually carry many more significant digits than CAD programs. Often, the optomechanical designer will round off values to three or four digits, which may be adequate for mechanical parts, but not for optical ones. The optomechanical designer should provide his prescription of the system back to the optical engineer, who should modify (a copy of) his optical model using the specific values called out in the CAD model or drawings. This can be particularly important when using higher-order aspheric surfaces. If a performance change is seen due to the reduced number of digits, the optical and optomechanical designer must agree upon the number of significant digits required.

Once the final optomechanical design has been completed, the optical designer should create a final optical model from it, verify the performance, and archive the model. It is good practice for the optomechanical model to reference this corresponding optical model. If any changes occur in either the optical or optomechanical model, similar changes need to be made in the corresponding model, and the new models archived.

4.10 Stray Light

Stray light is unwanted light that degrades the performance of a system. In general, there are two types of stray light: that due to in-field sources and that due to out-of-field sources. An example of out-of-field stray light would be the flares often seen in pictures taken with consumer imaging systems, such as digital cameras, due to the sun being just outside the field of view. An example of in-field stray light, again considering digital cameras, would be the halos sometimes seen around street lamps in a picture taken at night. In addition to imaging systems, stray light can also degrade the performance of nonimaging or illumination systems. For instance, an automotive headlamp reflector that projects light in the wrong direction may blind other drivers, creating a safety hazard.

Plastic optical systems have at least the same susceptibility to stray light as do glass optical systems. In some cases, due to their special characteristics, plastic optics may be more susceptible to stray light than glass optics. Consider, for instance, the use of integrated mounting flanges on lenses. Because the flange is made of the same material as the lens, it will also transmit light. In contrast, a glass lens will generally not have an integrated flange and will most likely rest on an opaque spacer or lens cell, which will not allow light to pass. This is not to say that we should not take advantage of integrated flanges in plastic optics; we just need to be mindful about doing so.

Consideration and analysis of stray light issues should be conducted during the optical and optomechanical design phase of system development. Frequently, stray light and stray-light analysis are an afterthought, left as something to be

dealt with once the rest of the design is complete. This mindset can be severely detrimental to a project. Finding a stray light issue after the rest of the design work has been completed, with prototypes and/or injection molds built or being built, can result in a significant cost in terms of both money and schedule. In many cases, simple up-front design changes could have eliminated the stray light problem.

Stray light can result from a number of mechanisms, alone or in combination with each other. There are several techniques to reduce or eliminate stray light from each of the mechanisms. Depending on the particular application, design form, cost, and other constraints, it may not be possible to utilize a particular technique. In this case, alternate techniques, redesign, or simply living with the stray light may be required. We now discuss several stray light mechanisms and mitigation techniques.

The first stray-light mechanism is straightforward reflection from optical surfaces, an example of which is shown in Fig. 4.23. We can see in this simple example that light reflects from the rear surface of the lens, heads back towards the front lens surface, reflects again, and then passes through the second surface, arriving at the image plane. This type of stray light is referred to as a “ghost.” For a point object, there will be an illuminated spot of stray light on the image plane in addition to the direct image of the point object. In this case, the size of the ghost image will depend on how far from the image plane the stray light focuses. In this example, the stray light is focused near the primary image plane. The irradiance of the ghost image (how bright it is, usually in W/cm^2) will also depend on how far from focus it is as well as the amount of energy reflected from the surfaces. Depending on the light source and sensor geometry, the ghost may be laterally offset from the primary image, in which case it could appear that there are two objects of different intensity. Alternately, the ghost may fall on top of the primary image, most likely resulting in the appearance of a halo.

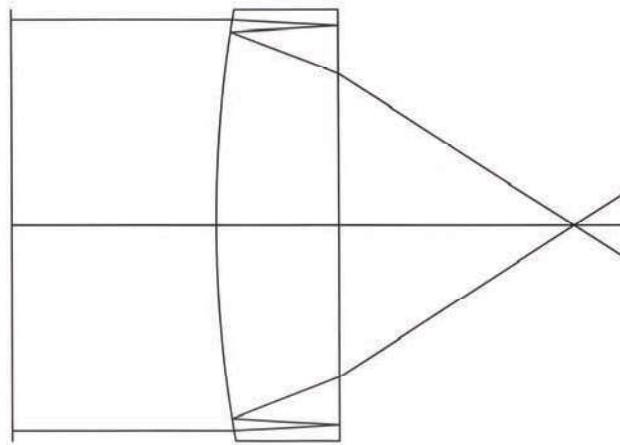


Figure 4.23 Stray light due to ghost reflection.

There are two main techniques for reducing or eliminating ghost images. The first technique is to design the optical surfaces such that any ghost images are well out of focus. In this way, the energy from the ghost will be spread out over a large area and appear only as a small background change. The first-order analysis of ghost images can be performed by the designer during the optical design process. Most optical design programs have built in features (or at least macroprogramming capability) to analyze ghosts due to two reflections, as was shown in Fig 4.23. The analysis is performed by considering the image location of each pair of potentially reflecting surfaces within the system, with the appropriate refraction at each of the intervening and subsequent surfaces. For simple systems, this analysis can be performed manually by converting refracting surfaces to reflecting surfaces and performing the appropriate ray tracing. The number of possible surface pairs rises rapidly as the complexity of the system grows, which is why the automated method of the programs is preferred.

In addition to reflection off the surfaces of lenses, reflections can also come from any filters between (in multielement systems), in front of, or behind the lenses, or from the surfaces of the window covering the detector, and/or the detector surface itself. Because of this, it is generally good practice to avoid having the final optical surface concentric to the detector (concentric to the detector meaning that the detector is at the center of curvature of the surface). If the surface is concentric to the detector, or close to being concentric, reflections from the detector surface will be directly reimaged, resulting in a sharply focused ghost.

Besides controlling the focus position of the ghost image, we can also try to control the amount of energy it contains. This is performed through the use of AR coatings, which were discussed previously. Antireflection coatings become more important as the number of elements in the system increases, not just for ghost image reasons but for overall system transmission as well. In addition to coating the optical elements, AR coatings on the detector window and detector surface can help to control the amount of energy in a ghost image.

In addition to reflection from optical surfaces, stray light can also result due to reflections from nonoptical surfaces. In most systems, there are a number of nonoptical surfaces that stray light reflections can come from. These include any housings, baffles or sun shields, spacers, lens flanges, barrel walls or openings, and even the aperture stop of the system. Each surface must be evaluated for possible reflection paths that end on or near the detector surface. Most of these surfaces are generated during the optomechanical design, so it is useful for the stray light analyst and the optomechanical designer (if they are not the same person) to be in close collaboration.

Stray light due to reflection from a nonoptical surface is usually handled using one or more of three methods. The first method is to prevent the reflection from reaching the image plane through the use of baffling. The second method of handling reflections is to alter the shape, location, or orientation of the surface that is causing the stray-light reflection. For example, consider a two-element system with an opaque spacer between the lenses. In this system, it is possible for

a reflection off of the inner diameter of the lens spacer to go directly to the image plane. By altering the spacer geometry from a cylindrical section to a conical section, a given reflection can be shifted from going to the detector to going into the rear flanging of the barrel where it is blocked. Of course, the entire range of possible input angles must be evaluated. It may be that due to the range of input angles, there is not a suitable angle for the spacer such that reflections for all input angles are dealt with. If so, it may require multiple techniques to prevent unwanted reflections from reaching the image plane. For instance, a “sunshade” may be necessary at the front of the system to limit the angular range of input rays. The sunshade, in combination with selection of the appropriate spacer angle, may solve the stray light problem.

The third method of dealing with reflections from nonoptical surfaces is to alter the reflective nature of the offending surface, which can be accomplished in several ways. One way would be to apply an absorptive coating to the surface, such as paint or black ink. Ink can sometimes be applied using a fairly straightforward pad printing process. In this process, the surface is pressed against an ink pad, similar to those used with rubber stamps, transferring ink onto it. While absorptive coating techniques can be effective, they are generally not the preferred approach. The reason they are not preferred is because they require an additional process to the part, which will tend to increase its cost. If we have elected to use plastic optics because of their relatively low cost, we do not want to add processes that will increase their price.

An alternate technique is to change the part reflectance by texturing the surface. Texturing or roughening of plastic parts is quite common. It is possible to mold parts with a wide range of surface textures, as can be seen from observing plastic parts such as keyboards, televisions, or other electronic equipment. In nonoptical applications, texturing is often used for styling purposes, to improve the ability to grip an item, or to hide fingerprints or other contamination. In our optical application, the amount and type of texturing will determine how the light is spread after striking the surface. The spread or scattering of the light will depend on the type of plastic material, the roughness of the surface, the angle of incidence of the light, and to a certain extent, the wavelength of light used.

For purposes of stray-light analysis and modeling, we would like to know the distribution of light reflected from the surface as a function of the angle of incidence of light on the surface. This distribution is known as the Bidirectional scatter distribution function (BSDF). Knowledge of the BSDF is typically acquired by directly measuring parts using equipment known as a scatterometer. This type of equipment is commercially available,⁸⁵ although many designers will not have sufficient need to justify its purchase. BSDF information may also be obtained from various scatter databases, though in many cases they are considered proprietary. If scatter measurements are desired but the designer does not own a scatterometer or have access to a scatter database, the measurement can be performed on a contract basis, often from the equipment manufacturers. However, many times the parts are not available upfront or there is limited

funding to allow directly obtaining the BSDF. In these cases, it is possible to use approximations or analytic models to perform the stray-light analysis. For example, the assumption may be used that the surface will have a Lambertian scatter distribution, where the radiance of light scattered at a given angle from the normal to the surface is constant.

In many cases of roughening parts, a scatter model is assumed for the analysis, the parts are made as rough as reasonably possible, and the stray light performance is tested when the system becomes available. Most molders have a three-ring binder full of plaques that show a range of possible surface textures. The customer, often in consultation with the vendor, can then select the appropriate surface texture. It can be important to include the vendor in the selection, as the texture may introduce difficulty in molding the part. If the texturing causes the part to stick in the mold, this can cause trouble during its ejection, which may result in warping the part or potential mold damage.

One example of a troublesome reflection from a nonoptical surface is that from the rear aperture of a lens barrel. Consider the following example, where a molded plastic barrel was used to house a multielement lens system. The rear of the barrel had an opening smaller than its diameter. The size of the opening was set sufficiently large to pass the cones of light going to the image plane, but small enough to provide a ledge on the barrel on which the final element of the system rested, as is illustrated in Fig. 4.24. The barrel was molded using a single through pin to allow precise control of the inner-stepped diameters where the lenses were mounted. The barrel material was Acrylonitrile Butadiene Styrene (ABS), which can be molded to achieve a high-gloss finish, inadvertently in this case. This resulted in a shiny cylindrical aperture in the rear of the barrel. When mated to a detector, a brightly illuminated ring was seen in the image. The ring resulted from reflection off the cylindrical surface by rays that would otherwise fall outside the detector.

Resolution of this stray light issue was achieved by altering the offending reflecting surface in two ways. First, the length of the surface parallel to the optical axis was shortened by adding a chamfer, and second, the surface was roughened to diffuse the light striking it. Because the roughening of the surface created problems molding and ejecting the part, the surface eventually was roughened using a secondary manual process. This secondary roughening could have been eliminated had a more thorough stray-light analysis been performed early in the design process. One solution would have been to shorten the length of the reflecting surface (as was done) as well as to move it to a larger diameter where the reflections from it would not have reached the used portion of the detector. Regrettably, by the time the problem was found, the lens mold was already being manufactured, and the lens-mounting flange structure could not easily be changed. In hindsight, this type of design error is readily apparent; however, it is easily overlooked during the rush to get a design completed and systems built. This example shows the potential benefits of making the effort, and taking the time, to perform stray-light analysis early in the design cycle.

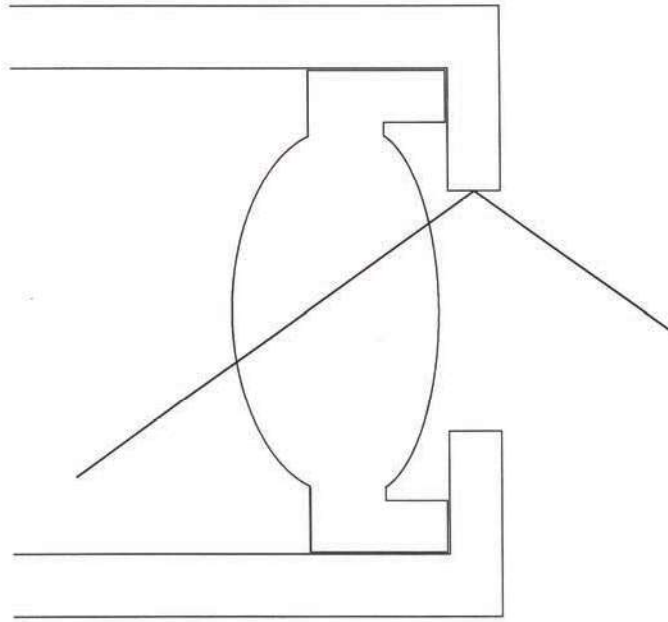


Figure 4.24 Stray light due to reflection off of the rear-lens-barrel aperture.

Another type of reflection that can cause stray light issues is total internal reflection (TIR), which occurs when light is incident on the boundary of two media, going from a higher to a lower-index material, at an angle larger than the critical angle. In this case, the light is totally internally reflected: totally, in that the amount of light reflected is (theoretically) 100%, and internally, since it was attempting to go from a higher-index material to a lower-index material. Total internal reflection can occur on lens surfaces, typically from out-of-field rays. The total internal reflection of any in-field rays should be seen during the design process if the range of field angles is sufficiently represented. Total internal reflection of out-of-field rays is most likely found during the stray-light analysis. In addition to the lens surfaces, TIR can also occur in lens flanges or in the transition region between the optical surface and the flange of the lens.

Like other stray light due to reflections, TIR reflections can be controlled by surface position, orientation, and texture. Since total internal reflection occurs when rays strike an interface beyond the critical angle, it should be a goal to limit the ray angles to less than this angle. Achieving this may require adjusting the slope of a flange surface or limiting the input ray angles with a baffle or sunshade.

Stray light due to TIR can be particularly severe due to the amount of energy in the reflected beam. Unlike reflection from an AR-coated lens surface, which may be 1% or less, a beam that totally internally reflects will carry the entire amount of energy incident on the interface. While a detriment in certain situations, total internal reflection, with its ability to reflect all the incident

energy, is often used in the design of illumination systems. For example, LEDs are often encapsulated with or inserted into a part designed to totally internally reflect the light emitted at large angles, thus directing it in forward angles.

In addition to being caused by reflections, stray-light issues can also arise from scattering. Probably the most familiar example of this is looking through a dirty car windshield. The light from the sun, or from cars' headlights at night, is scattered by the contamination on the windshield, which creates a large amount of glare. The amount of scattered light can be enough to essentially blind the driver, washing out all useful visual information. A similar effect can happen from rough surfaces within an optical system. The scattered light that reaches the detector can increase the background light level, reducing the contrast of the image.

We previously mentioned roughening surfaces as a stray-light mitigation technique in order to diffuse light striking them. In some instances, a smooth (rather than rough) surface is desired. This can be the case when a surface is oriented such that all reflections from it are blocked from reaching the detector. On the other hand, if the surface is rough, the scatter would spread the light over a larger angular range, possibly resulting in stray-light paths to the detector. The effect of surface finish of all elements in and around the imaging path should be considered. In some instances, such as a long cylindrical section that is being injection molded, it may not be possible to mold the part with a rough surface, as this will cause it to stick in the mold.

In addition to reflections and scattered stray light, there can also be stray light that results from direct paths. While not typically seen with inline refractive glass optical systems, it is sometimes seen with inline refractive plastic optical systems. An example of this is shown in Fig. 4.25, where the direct path travels

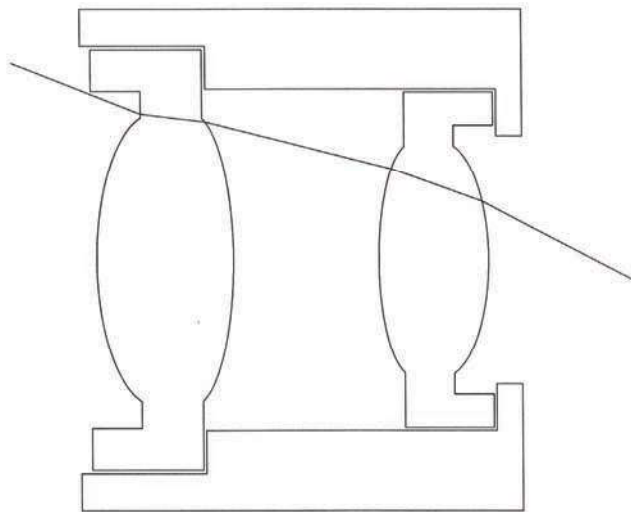


Figure 4.25 Stray-light path through lens transition zone on first element.

first through the transition zone between the optical surface and the lens flange on the first element, then through the optical surfaces of the second lens. This direct path could be dealt with in several ways. One solution would be to place a baffle or blocking structure in front of the first element in order to prevent the light from reaching the transition region of the lens. If the lenses are being stacked into a barrel and are loaded from the front, a retaining ring used to hold the lenses in place could serve the dual purpose of preventing the direct stray-light path.

If a blocking structure cannot be placed in front of the element, the direct path could also be blocked behind the lens. There are several options for the blocking element, which again may serve a dual purpose. One option would be to use a molded spacer, which sets the distance between the two lenses as well as blocks the direct path. However, it must be considered that the axial length of the spacer may create an additional stray-light path. A way of limiting the axial length of the blocking element is to use a thin disk of material, such as a Mylar or metal washer. These washers can be stamped from sheets of material of various thicknesses. The thickness should be thin enough to limit potential reflections from the inner diameter of the washer but thick enough to allow for handling. The thickness tolerance of the washer material selected must be kept in mind, as the addition of a spacer or washer may increase the overall lens spacing variation. This additional tolerance stack-up must be accounted for in the system tolerance analysis.

Another potential solution to eliminate the direct path would be to reduce or move the transition zone. Provided there is suitable edge thickness, it may be possible to carry the optical surface out to a larger diameter. Of course, the effect of the increased aperture on the system must be considered as well as any implications on the lens mold or other components that interface with the altered lens.

One other potential source of stray light, as discussed earlier, is diffractive surfaces. The efficiency of standard (as opposed to MOD) diffractive surfaces is optimum for one wavelength—the design wavelength. As we move away from this optimum value, the diffraction efficiency decreases, and light appears in other orders. These other orders may reach the image plane, resulting in stray light images. Any stray-light analysis on a system containing diffractive surfaces should consider diffraction-efficiency effects.

Stray-light analysis, like optical design, is a specialty that is learned partially through study but mostly through practice and experience. There are a number of consultants and companies that specialize in stray-light analysis. Several of these companies also have commercially available software programs, which are often used for illumination design as well as for stray-light work. The most well known of these programs (listed alphabetically) are ASAP,^{xiv} FRED,^{xv} and TracePro.^{xvi}

^{xiv} ASAP is a registered trademark of Breault Research Organization.

^{xv} FRED is a registered trademark of Photon Engineering LLC.

^{xvi} TracePro is a registered trademark of Lambda Research Corporation.

Some level of illumination and stray-light analysis can also be performed in the optical design programs mentioned earlier, but it can be argued that they are not as well suited as the programs listed here, which are written for these particular purposes.

One of the most important aspects of performing a successful stray-light analysis is to have the optical system accurately represented in the analysis software. As with many other software programs, the output will only be as good as the input it is created from. This representation is referred to as a “stray-light model.” The model can be entered into the software in a number of ways and most often requires a combination of methods.

The most basic time-consuming way is to create the model from scratch, using the geometric elements available in the software. This is sometimes done when there is a need for an independent model to verify previous work or when no computer-aided design (CAD) or electronic models exist. Today, the optical and optomechanical design work will most likely be performed using software programs that have the ability to export various file types. In this case, optical and CAD files are generated, which are then imported into the stray-light analysis program. Typical CAD file types that the stray-light programs can handle are IGES and standard for the exchange of product (STEP) model data.

Once the files have been imported, properties must be applied to all elements and surfaces. These properties include the reflectance and transmittance of surfaces, their scatter distributions (BSDFs), any polarization effects, and the refractive indices and dispersions of the various materials. Antireflection or other optical coatings (such as band-pass filters) can also be applied. Some of these characteristics may be automatically assigned when importing the optical files. In general, however, some amount of time must be spent manually assigning and verifying the correct surface and element properties.

One problem that frequently arises during the construction of a stray-light model is that the optical and optomechanical files it is constructed from do not themselves accurately represent the actual parts produced. For instance, in many optical programs, the designer sets a clear aperture radius or diameter value that is used to determine which rays pass through a surface and which rays are blocked. In molded plastic parts, as discussed above, it is desirable to provide a region outside the optical clear aperture to allow for edge break. This additional area outside the clear aperture needs to be properly represented in the stray-light model. If the stray-light analyst simply enters the aperture sizes in the optical model, regardless of the optical surface dimensions of the actual part, the model will be incorrect. This may result in an existing stray-light path that is not shown in the stray-light model.

Another frequent flaw in stray-light-model development of plastic optical systems is missing transition radii. Sometimes, either in order to simplify the optomechanical model or because of a lack of understanding of most molded parts, the optomechanical designer will leave cusps in their model instead of including transition radii. An example of this is shown in Fig. 4.26, which shows

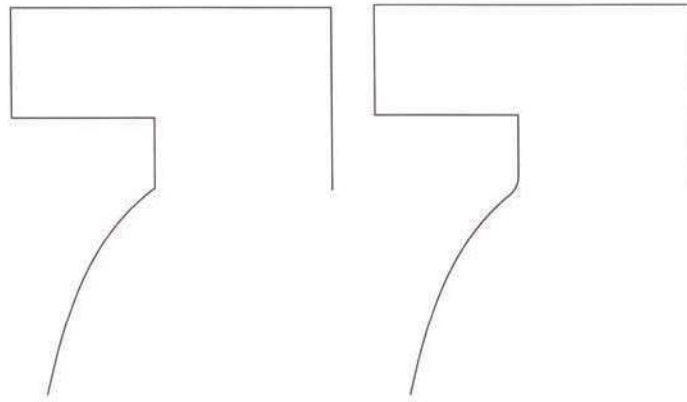


Figure 4.26 Unrealistic versus realistic transition radius.

the difference between a theoretical part (on the left-hand side) and the more realistic molded part (on the right-hand side). The transition radius in the molded part can act as a small negatively powered annulus, which may spread out light that would, in the incorrect model, refract as through a plane surface. All corners, which are shown in the figure as sharp, would in reality also have some minimum radius value.

Another often-overlooked feature in molded parts is the gate flat, which (as discussed earlier) is often used to provide space for any gate vestige during the degating process. With gate flats, the part is no longer circularly symmetric, which may have been the way it was modeled. In multielement systems, the gate flats are sometimes aligned during assembly by choice, and other times they are randomly oriented. This can complicate the stray-light analysis if the gate flats are creating stray-light paths, for instance, if rays totally internally reflect off them.

One difficulty with getting an accurate stray-light model of the system is the tolerances on the actual parts. In the case of a transition radius, there will normally be some allowed range of radius value. As an example, the radius may be called out with a maximum radius of 0.5 mm, which means the transition radius can vary from 0 to 500 μm . Because of the large number of rays that are traced during a stray-light analysis, and the time it can take to trace them, tolerances and perturbations to the system are usually not extensively evaluated. In this case, it is probably best to select a representative value for the transition radius and determine if it plays a role in significant stray-light paths. If it is found to play a large role, further perturbation analysis may be necessary.

Another problem with the stray-light model can occur due to uncommunicated changes to the system components. What appears to be a small change to one engineer may have a considerable effect on the stray-light analyst. For instance, consider the effect of a small change in the draft on a lens flange. To the mechanical engineer or lens designer, this may have no apparent effect, since it did not significantly change the structural stability of the system or alter

the aberrations of the lens. To the stray-light analyst, however, this change in angle could move a reflection that was previously not reaching the detector into the image region. To prevent this kind of mishap, there must be ongoing discussion between all parties involved with changes to the baseline design. This includes both internal and external parties, such as vendors.

We now consider the process of performing a basic stray-light analysis. As with some other processes described in the text, we do not have space to provide a complete description; however, most providers of stray-light software offer training on this type of analysis. Once the model has been entered and all appropriate properties applied, the first task is to perform a simple in-field ray trace. This is similar to generating a spot diagram in an optical design program. The purpose of this initial ray trace is to compare the output of the stray-light model with that of the optical design model in order to verify that the optical portion of the model has been entered correctly. This is particularly important when the system contains diffractive or highly aspheric surfaces.

Once the optical performance of the model has been verified, the next step is to run a “backwards ray trace.” This is performed by considering the detector surface as a light source that is illuminating back through the system. It may at first seem counterintuitive to perform a backwards ray trace, but there is a simple reason to do so. By considering the image plane as a source, emitting into (at least) a full hemisphere, we are tracing all possible ray paths to the detector. From the backwards ray trace, a list of all surfaces that rays reach is generated. The surfaces on this list are known as “critical surfaces.”

A forward ray trace is next performed, where rays are sent into the front of the system from all angles. A list of surfaces that the rays hit is again generated, with these surfaces known as “illuminated surfaces.”

The critical and illuminated surfaces are compared, with the surfaces that appear on both lists being determined. Surfaces on both lists are the ones that can generate a stray-light path because they can be illuminated by external sources, and the detector can see them.

With the various stray-light paths identified, forward ray tracing can be performed on any particular path to determine its effect on system performance as well as to evaluate the mechanisms involved in its creation. With this knowledge, appropriate mitigation strategies can be developed. The proposed solution may need to be a compromise between the optical, optomechanical, and production teams. Changes that eliminate one stray-light path may possibly create other unintended paths. As such, it may take an iterative design approach to eliminate or reduce stray-light effects.

4.11 Special Considerations for Small and Large Parts

Certain applications (such as mobile phone cameras, telecommunication components, and fiber optic couplers) often have severe packaging constraints,

resulting in the use of small-diameter lenses. It is not unusual for these applications to use lenses that have diameters ranging from 2 to 5 mm. These small-diameter lenses require special consideration during both their design and fabrication.

From the optical design standpoint, adequate performance of these systems often requires tight tolerances on the tilt and decentration between elements as well as between the surfaces on each individual lens. This results in a need to precisely register the elements during assembly, which can be achieved through interlocking features or through control of the element diameter or mounting features. In general, the tilt of the individual optical surfaces should be well controlled by the mold build process. For example, the holes for the optical inserts may be machined through both mold half plates at the same time, which all but guarantees they are parallel. Even if not machined this way, the angle of the insert hole with respect to the mold face will be tightly controlled. The inserts themselves should also run true, if proper set up on the diamond-turning machine is performed.

Proper control of the decentration of the two surfaces of a lens with respect to each other can be a challenging task. In these types of designs, the optical designer will often ask that the centers of the two surfaces be aligned to each other within 5 μm . For those who prefer units of inches, this would be two ten-thousandths of an inch (0.0002 in.). Achieving this tight decenter tolerance requires precision in all steps of the mold build process. To the credit of various plastic optic molders, these values can and are being achieved.

Even on a small lens, some room around the clear aperture surface is required for edge break in the injection molding process. This can use a significant portion of the lens area, possibly resulting in the need for a larger diameter. If this is not acceptable, trades should be discussed with the molder as to the minimum extra surface area required. Alternately, the designer may be able to put up with a small amount of error around the edge of the element or stop the system down slightly to eliminate rays from this portion of the lens.

Because of the small clear apertures of these lenses, digs or contamination on the optical surface can block a large portion of the beams passing through them. Digs are usually controlled by caring for and inspecting the optical inserts that form the surfaces. Contamination can be controlled through good housekeeping practices, proper packaging, and the use of air guns and deionizers. Care must be taken not to generate large amounts of debris during the degating process.

We previously mentioned that diamond turning of parts or optic inserts leaves a small mark at the center of the optical surface. If the diamond-turning setup is not performed well, this mark can become an appreciable area on a small lens. In general, centration requirements of the optical surface will force the diamond-turning mark to be small.

We discussed earlier the ejection of parts from the mold. Small parts are often optically ejected because of the limited area around the optic surface. Optical ejection, however, requires that the optic pin on the moving half of the mold slide back and forth in its receiver. This necessitates additional diametrical

clearance from what is required for simply inserting the pin. If this additional clearance is considered too large, ejector pins or sleeves may be utilized. Of course, the small available area for ejection requires thin ejection sleeves or pins, which are more fragile and susceptible to bending. The ejection method is often selected by the molder, based on his or her experience of what works best for a particular part type.

Another consideration for small parts, particularly for those with convex optical surfaces, is having a small annular ring around the optic surface. This annulus results from the desire not to have a “knife edge” on the concave optic pin that forms the surface. A concave optic pin with a sharp edge is more likely to be damaged during handling or insertion into the mold base. This annular region may lead to stray-light problems, as mentioned earlier, but also may be useful in testing, as we shall discuss later. Convex optic pins, which form concave part surfaces, do not require this annular flat.

Another concern for small optical parts is the size of the gate region relative to the size of the part. In order to adequately fill the part during molding, the gate needs to be some minimum size. This size can often be a significant portion of the edge of the element. Sufficient room must be left for degating the part and for any gate vestige that remains. Handling small parts during degating and other processes can be difficult. Element-specific tooling can be useful to assist in handling the elements, as can using the runner system as a support structure. In addition, special racks or trays may be needed for coating as well as for storing or shipping the elements.

We discussed stray light in the previous section; however, we mention the topic again as it applies to small elements. Systems using small elements can be particularly susceptible to stray light. This is because the packaging constraints that drive their use often do not allow length for stray-light suppression features, such as sunshades. Performance requirements often impose the need for multiple elements, which, when constrained to a short overall length, results in closely spaced lenses. Many times, this does not allow the use of molded opaque spacers. Given this, the use of inked surfaces or stamped apertures may become more important. In the case of stamped apertures, they may interfere with lens centering features, such as molded-in tapers. The use of multiple apertures with overlapping blocking regions and cutouts for alignment features has been attempted, as has stamped apertures with conforming tapers. The additional tolerance stack-up due to using stamped apertures must not be overlooked. Figure 4.27 shows a schematic of a small lens and optical insert, reiterating some of the special considerations for small elements.

At the other end of the range, large parts also need certain considerations. The definition of what constitutes a large part is somewhat arbitrary. For our purposes, we will consider large parts to be those over 50 mm in diameter. Some vendors specialize in very large parts, such as Fresnel lenses, while others will only work up to a given size, since manufacturing larger parts normally requires larger molding machines.

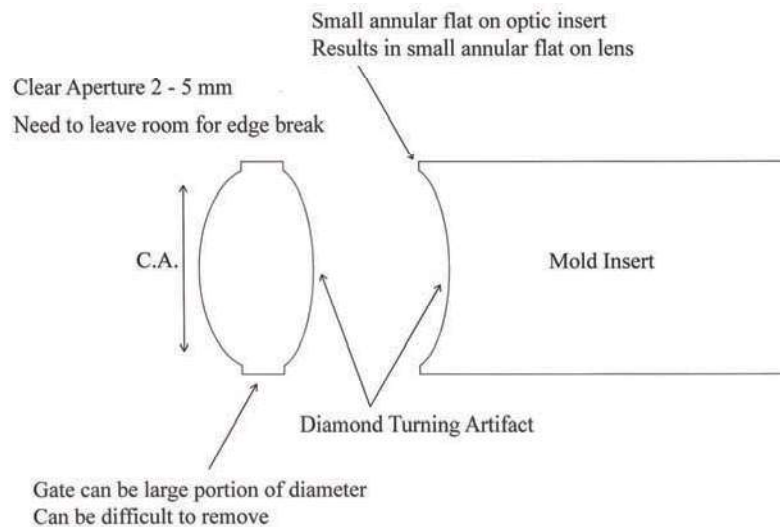


Figure 4.27 Considerations for small elements.

The primary consideration for large parts, when they are molded, is a realistic view of the tolerances that can be achieved. As parts grow larger, the accuracy that the surfaces can be held to decreases. In our discussion of tolerances, we stated that irregularity can be held to about one fringe on up to a 25-mm-diameter part. Beich⁶⁹ states that one fringe of irregularity per 25 mm can be held on up to 75-mm-diameter parts, which would mean that a 75-mm-diameter part has three fringes of irregularity. Ning,⁵⁵ on the other hand, suggests one fringe for the first 8 mm of diameter, and three fringes per each additional 10 mm, which would give about 20 fringes of irregularity on a 75-mm-diameter part. Obviously, there is a significant difference between these two values. This may be explained in part by differences in the level of effort and cost associated with larger parts and what is considered the baseline level. The use of standard injection molding or injection compression molding can also make a difference. Requirements for large parts should be discussed with potential vendors to ensure that the tolerances required can be met at a reasonable cost. Overly optimistic expectations can quickly leave the designer looking for alternate solutions.

As we discussed earlier, cost is typically proportional to cycle time. Higher irregularity tolerances are obviously easier to achieve and will require fewer mold compensations and shorter cycle times. Larger parts, because of their larger mass, will require a longer cycle time than smaller parts to achieve a given quality level. This will result in an increased cost compared to smaller parts. In addition, as parts become larger, the impact of material selection may also have a larger influence on their cost. Depending on the size and shape of the large part, alternate manufacturing techniques to injection molding (such as injection-compression, straight compression molding, or machining) may need to be considered.

4.12 Drawings

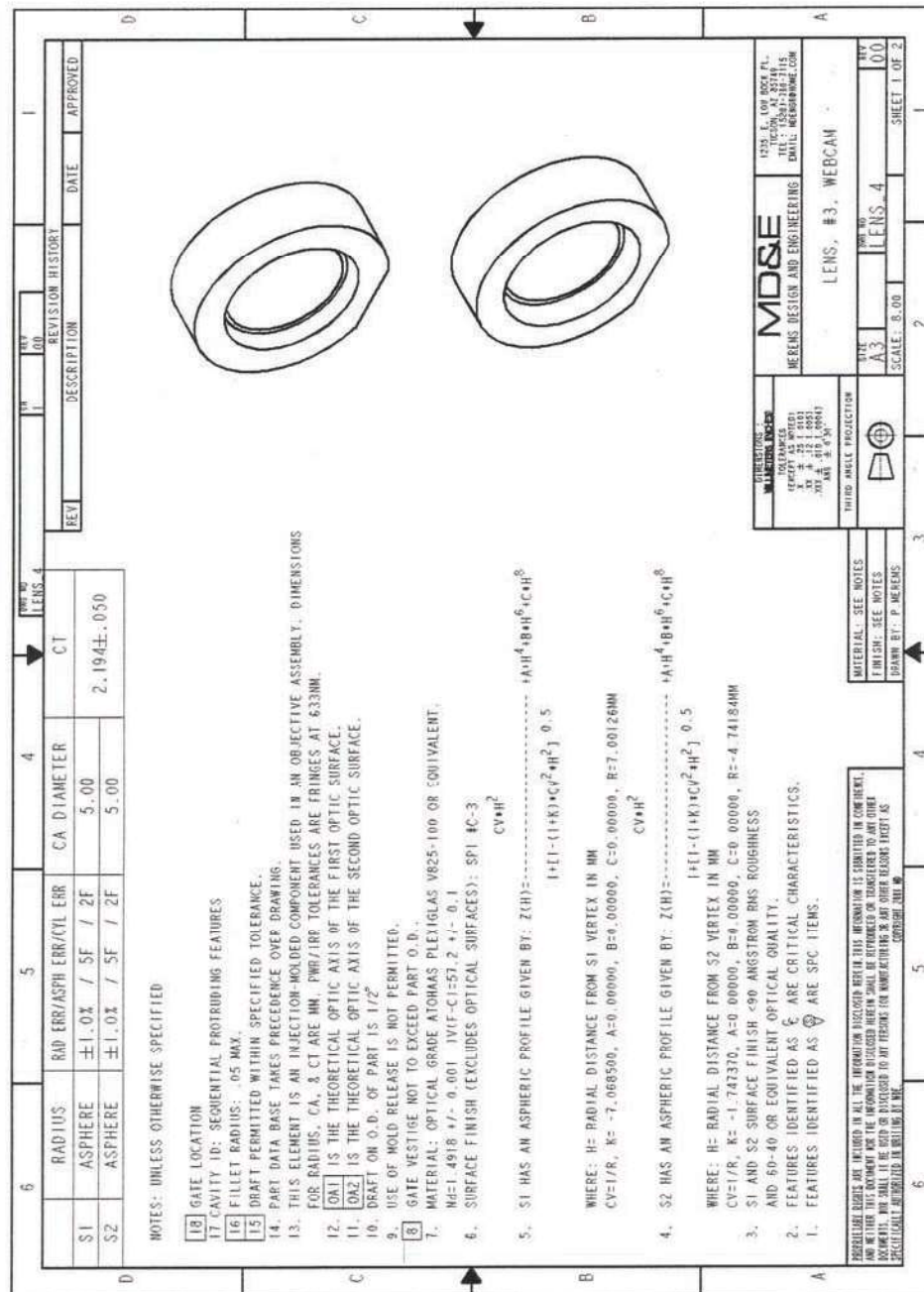
Drawings for plastic optical elements are similar in many ways to drawings for glass elements. The drawing should specify all the necessary information to correctly produce the part. Because of the special characteristics and manufacturing methods of plastic optics, the drawings may require additional callouts not seen on glass optical elements. It should be noted that several drawings may need to be produced during the product development. For instance, an initial drawing for diamond-turned prototype parts may be used, followed by a production drawing, which includes features for a molded part, such as draft and a gate flat. Figure 4.28 and Fig. 4.29 show the two sheets of a production plastic lens drawing. We do not claim this drawing to be error free or that it should be used as a template. It does, however, contain most of the features that a plastic optic drawing should contain. We discuss this drawing, working our way through the notes and pointing out relevant features.

Notes 1 and 2 indicate that statistical process control (SPC) features and critical characteristics (CC) have been called out on the drawing. Because of the repeatability of the injection-molding process, as well as the desire for lower cost, most plastic optic elements are not individually inspected. Instead, samples are selected at some frequency, and certain characteristics are measured. These measurements are tracked and analyzed using statistical process methods, which are well known in high-volume production environments and are used more frequently in general manufacturing. We do not elaborate on these SPC methods, except to say that they can be used to track a process and predict when it is going out of control, so that adjustments can be made before nonconforming parts are produced. This is done through the use of control limits, which are not the same as, and are normally smaller than, the tolerance on the SPC feature.

For this part, the overall flange thickness has been selected as the SPC feature. Selection of the individual SPC features is often done in consultation with the vendor, based on an understanding of their molding process. The critical characteristics are features of the part that are considered to be critical to the fit or function of the part or to the system it goes into. For this part, the flange thickness, center thickness, the stack of surface S1 (the distance from the vertex of the surface to the flange), and the diameter have been called out as critical characteristics. Critical characteristics are usually determined by the part and system designers, often the optical and optomechanical designers, who understand the effect of each dimension on the part or system performance.

Note 3 on the drawing calls out the requirements for surface roughness and scratch-dig on the optical surfaces. For surface roughness, it is best to supply a frequency range that the roughness should be measured over. This, as well as the subject of scratch-dig, will be discussed in the section on testing.

Notes 4 and 5 specify the forms of the two optical surfaces. For this part, both of the surfaces are aspheric; in particular, they are conics. On this drawing, the surfaces are specified using the standard polynomial aspheric form, with the



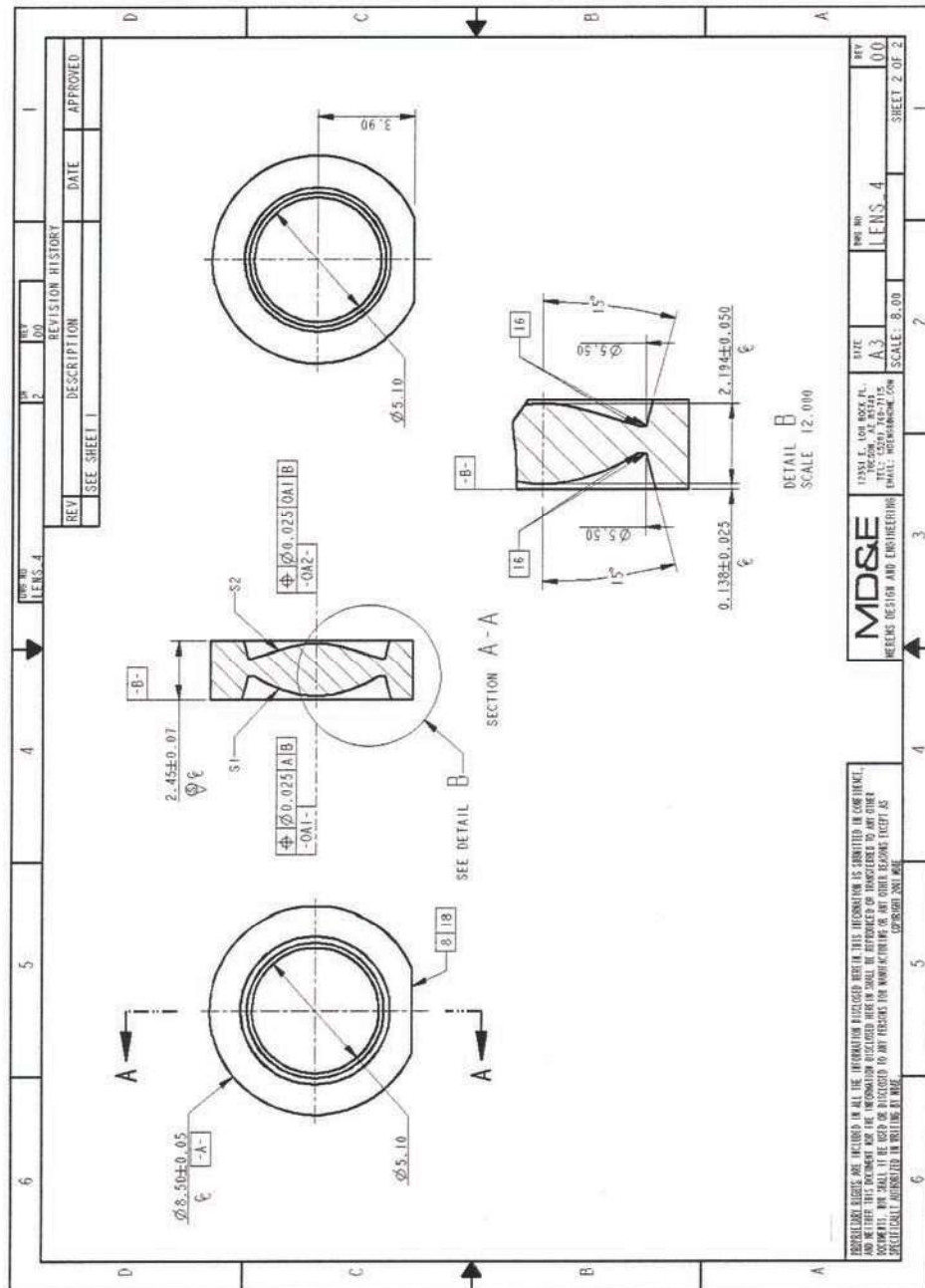


Figure 4.29 Sheet 2 of a plastic lens drawing. (Figure courtesy of Paul Merems.)

radii and conic constant values supplied, and the polynomial terms all set to zero. This form was used, even though the polynomial terms are zero, because the drawing is created from a template in the CAD software. Using standard drawing forms can be a good practice if the template is properly set up to ensure that all necessary data will be included on the drawing.

Note 6 specifies the surface finish on the nonoptical surfaces of the part. It is useful to specify a surface finish for these areas so that the parts will match the input to the stray-light model. It is especially important when a desired amount of texturing is required.

Note 7 calls out the optical plastic material of the part. In this case, the material specified is acrylic, one brand name of which is Plexiglas. A specific supplier and grade of material have been listed, with the possible use of an equivalent material. As noted earlier, it is wise to call out the specific grade of material to be used. There are usually multiple grades of each plastic optical material, with each grade having different additives and/or characteristics. If the designer is going to allow or consider alternate or equivalent materials to be used, it is best to require notification by the vendor of this material change before it is implemented. This can be dictated by a statement such as, "Substitution of the primary material called out by an alternate or equivalent material must be approved in writing by the customer prior to substitution taking place." Such a statement reduces the possibility of errors due to doubt about the definition of "equivalent." This statement may be placed on the drawing itself, or it may be included in the purchase order, depending on the drawing philosophy of the customer.

It is also generally wise to include a note that precludes the use of reground or recycled material. Most optical parts are not manufactured using regrind material, and most vendors themselves believe this to be appropriate. However, this is a case where it is better to be safe than sorry. A note dictating no regrind may be written along the lines of "Virgin material only—no recycle, remelt, or regrind material is allowed."

Note 8 on the drawing specifies that the gate vestige must be below the outer diameter of the part. As we recall, gate vestige is material remaining on the gate flat after the part has been separated from the runner system (degated). This is a fairly standard callout, particularly for lenses that will be inserted into a cylindrical barrel.

Note 9 prohibits the use of mold release. Mold release is often used in standard plastic molding. The mold release may be contained within the plastic itself (another reason to specify the material grade), or it may be added to the plastic or to the mold. Mold release is a lubricant that allows the parts to come out of the mold more easily. Most optical elements are produced without the use of mold release, which may contaminate the optics, outgas, or make them difficult to coat.

Note 10 calls out the amount of draft allowed on the diameter of the part. We discussed earlier that draft is the angling of surfaces that would be otherwise parallel to the draw direction of the mold. Draft helps parts release from the

cavities, both when the parts are pulled with the B mold half as well as upon ejection. In this case, the designer is allowing up to one-half degree of draft. Many times, when the diameter of the lens is used for centration control, a small no-draft section will be included on the part, as was discussed in the optomechanical design section.

Notes 11 and 12 reference the optic axes of the two optical surfaces. The optic axes are geometrically toleranced to control the position of the optic surfaces to the relevant part features.

Note 13 states that the part is an injection-molded element and that the dimensions are in units of millimeters. This dimensional statement may also be called out in the block near the drawing name. This, again, is fundamentally a drawing philosophy issue. More importantly, the note also calls out that power and irregularity tolerances are specified in fringes, at a wavelength of 633 nm. This is a fairly standard callout on optical drawings. Having it explicitly stated prevents any confusion as to whether the tolerances are in fringes or waves, which would result in a factor of two differences in tolerance size. The wavelength of 633 nm is that of a helium-neon (HeNe) laser, which is the red laser commonly seen in laboratory interferometers. Sometimes optics are specified using the mercury green line of 546.1 nm. This applies more to parts that are to be checked using test plates (under a mercury lamp). Since most plastic optics will not be tested using test plates, we recommend using the HeNe wavelength.

Note 14 sets the precedence of the various descriptions of the part. In this case, the note specifies that the electronic CAD model of the part takes precedence over the drawing. This may seem a bit unusual given that we are reviewing the part drawing; however, the fact is that, as a result of the advances in CAD software and computer-controlled machining, many vendors prefer to work from the electronic model of the part rather than the drawing. It is much easier for the molder to design the injection mold starting with the CAD model of the part, supplied by the part designer, than to receive a drawing and have to create a CAD model from it. Electronic models can also be used to program automated test equipment, such as coordinate measuring machines. The CAD models of the parts can (and should) be placed under configuration control, just as the drawing should be.

Another advantage of using an electronic model is that the molder can suggest and implement (if permitted) changes to the model that will make the part more producible. These changes may include adjustment of the gate region or the addition of desired draft. This type of concurrent engineering with the molder can improve the design from a cost and producibility standpoint.

Note 15 reiterates the allowed draft on the part, though no callout block associated with the note is shown on the drawing.

Note 16 calls out the allowed fillet radii on the transition between the optical surface and the flat annular region surrounding it, and the transition between the flat annular region and the flange of the part. We have mentioned previously that sharp corners or cusps are generally not preferred, and they are not typically

achieved without extra effort. Calling out the maximum fillet radius sets an upper bound on the part geometry in the transition regions. Since sharp corners are not preferred, a minimum radius callout may be added as well. As we discussed in the section on stray light, the effect of these fillet radii, which can act as negatively powered annular zones, must be considered.

Note 17 calls out that the parts will have a cavity ID molded on them. In multicavity molds, this is a common practice. Molding a cavity ID on or into the part will allow the determination of which mold cavity it was produced from. This can be useful in linking system performance to component performance. For instance, if it is found that low-performing systems all contain a lens molded from cavity 5, the cavity 5 lenses can be pulled from the assembly line and stock area and be segregated and analyzed to determine the root cause of the low system performance. In some cases, it may be determined that system yield is maximized when certain cavity combinations are assembled together. For example, consider a two-element lens system, with each lens molded in its own four-cavity mold. It may be that the lens 1 from mold cavities 1 and 3, when combined with lens 2 from cavities 1 and 2 (of the lens 2 mold), provide high performance systems, but not when combined with the lens 2 from cavities 3 and 4. This is obviously not the preferred situation; we would like all lens and cavity combinations to work equally well together. However, we may find that we need to assemble the systems according to certain lens cavity combinations. Having cavity IDs on the parts will allow for easy separation and verification of which cavities the lenses are from and an orderly flow of the appropriate parts to the assembly floor.

Cavity IDs can be produced in a number of forms. Most commonly, numbers (1, 2, 3,...) or letters (A, B, C,...) represent the cavities. In other cases, symbols (such as dots) may be used. For larger parts, some molds also have an adjustable date stamp, which marks the part with the date it was produced. The cavity ID feature can either be proud of or below the surface. It is best to call out where on the part the cavity ID should be located, a maximum and minimum size, and if the ID is to be proud of or below the surface. Simply calling out the need for a cavity ID may result in a feature in an undesirable location. If possible, it is best not to put the cavity ID on an area that will be used as a mounting surface.

The final note on the drawing, Note 18, specifies the location of the gate. Appropriately, this is indicated on the gate flat feature. Sometimes a callout is made that the part must be single gated or that multiple gates are allowed, provided certain restrictions (such as “no knit lines within the clear aperture”) are met. Gate size and location issues should preferably be discussed with the molder before the final part design is set.

In addition to the notes, there are a number of dimensions and tolerances on the part. The blocks on the upper left of sheet 1 call out the tolerances on radius error, aspheric error, and cylindrical error as well as the minimum clear aperture diameter, the center thickness, and its tolerance. Radius tolerances are often called out as either a percentage or as a number of fringes of power. For glass optics, test plates are often used, which provide a measurement of radius error in

fringes. For plastic optics, radius measurement techniques other than test plates are most often used. In these cases, a percentage tolerance is easier for the molder to work with than a fringe tolerance, due to the measurement technique. If the drawing calls out the radius tolerance in fringes, it may just be converted to a percentage for the vendor's internal measurement. On this drawing, the tolerance on both radii is given as $\pm 1.0\%$.

There are several ways to specify the tolerances on an aspheric surface. Providing a bounding zone for the aspheric surface is one way; a bounding zone combined with a slope error specification is another. The surface can also be specified not directly by its form but by the parameters of a particular wavefront after passing through it (typically the wavefront after passing through the entire element). Yet another method is to specify the surface required with respect to a defined null configuration. However the surface is specified, the designer should verify that the vendor has the capability to measure it in the desired manner. If the vendor does not have certain types of equipment, an alternate or equivalent tolerance specification may be required.

On this drawing, the aspheric form is called out using an aspheric error term, specified as five fringes (F). This tolerance callout sets a bounding zone around the desired aspheric form. That is, the asphere may depart by up to five fringes from the desired form. There is some ambiguity in this tolerance callout, as it is not completely clear if the tolerance is five fringes departure peak to valley or a departure of plus or minus five fringes, which would potentially double the allowed tolerance. This ambiguity can be removed by stating the appropriate interpretation either in the drawing notes or in the tolerance box itself. There is also an unstated assumption that the radius may be varied within its allowed tolerance range to minimize the aspheric error, before the aspheric error is calculated.

The cylinder error called out allows variation in the surface form in two orthogonal directions. From the molding standpoint, cylindrical error may arise from asymmetric shrinkage, often in directions with and opposite the gate, or it may result from imperfect ejection.

The clear aperture diameter callouts indicate the diameters over which the surface specifications must be met. It is assumed that the clear aperture is centered on the theoretical optical axis of the surface. We emphasize once more the need for the optical surface to be sufficiently larger than the clear aperture diameter in order to account for "edge break" or shrinkage at the edge of the optical surface on molded parts. On this part, the clear aperture diameters are called out as 5.00 mm, while the optical surface itself is 5.10 mm. This is less room for edge break than most molders usually would want. When specifying the clear aperture and the optic surface diameter, it is important to take the transition zone fillet radius into account. If the clear aperture, optic surface, and fillet radii are not considered together, it can result in conflicting dimensions on the drawing.

On this drawing, center thickness and its tolerance are included in the tolerance block as well as being listed on sheet 2. This again comes from the use

of a standard drawing template. The specification for this lens is a center thickness tolerance of ± 0.050 mm, which is ± 50 μm , a tolerance that should easily be met.

Finally, on sheet 1 in the lower left corner is a proprietary statement block. This type of statement has become fairly standard on drawings. Most vendors will sign nondisclosure agreements in order to handle proprietary drawings and will control the information accordingly. Nevertheless, it is good practice to declare any and all information that is considered to be proprietary.

Moving to sheet 2, there are a number of different dimensional and tolerancing callouts. The diameter of the lens is specified as the primary datum, with a given value and tolerance, while the front face of the lens flange is called out as the secondary datum. In addition to the center thickness (the “stack” of S1), the distance from the vertex of S1 to the front face of the flange is called out. Typically, one surface stack is called out, and the other surface stack floats within the given tolerances for the stack (CT) and overall flange thickness. From an optical design standpoint in a multielement lens system, this may put a tighter surface-to-surface spacing (between adjacent elements) on one side of the lens than on the other. If one side’s spacing is more critical than the other, putting the stack callout on the critical side may help with the tolerancing. On the other hand, if the center thickness is less sensitive, it may be more appropriate to call out both surface stacks and let the center thickness float. Whatever choice is made, it is important that the tolerance analysis accurately reflect the way the parts will be built and assembled. The drawing callouts should reflect the desires of the designer, as opposed to choosing which stack to tolerance by chance or convention.

With regard to centration of the two lens surfaces, on this drawing the first lens surface S1 is referenced to the diameter of the lens, while the second surface S2 is referenced to the first surface. In some cases, both surfaces are referenced to the diameter. The choice of how to dimension the two surfaces should again depend on the sensitivities of the design, in this case the sensitivity of the two surfaces to each other versus the sensitivity of the two surfaces to the diameter.

This drawing does not call out any coatings for the lens. If any coatings [such as antireflection (AR) coatings] are required, they should be called out in the drawing notes. In some cases, the element may have multiple or different coatings on the different surfaces. For instance, the rear surface may have an AR coating, while the front surface may have an AR coating, a scratch-resistant coating, or both. Any coatings should be fully specified without any ambiguity. For instance, if a transmission is specified, it should be clear if the specification is for the surface or for the entire element. The note should include the required coating properties, such as allowed reflectance over spectral band and angular range as well as any adhesion, abrasion, and environmental requirements. In addition, if coating properties are to be tested using witness samples, appropriate control of the witness samples should be specified.

The drawing shown and discussed in this section details most characteristics of a typical plastic optical element. Plastic optics of different forms and for

different applications may require alternate or additional callouts and/or dimensioning schemes. The ultimate goal of the drawing is to document the part that has been designed and to allow it to be produced and compared against measureable characteristics. If there are doubts about the correct way to specify a particular part characteristic, discussion with the vendor is recommended.

4.13 Vendors and Vendor Interaction

The best time to begin discussions with plastic optic vendors is at the beginning of the project. This is particularly true if the designers are unfamiliar or relatively inexperienced in the design of plastic optical systems. Discussions with the vendor can help avoid potentially costly missteps in the early stages of the design. The best overall system designs result from collaborative engineering between the designers and the vendor. The designers should understand the product they are designing, which requirements are firm, and which can be traded for cost, performance, or producibility. The vendor should be able to provide input on the trades that can be made, such as those that will help them produce the parts more easily, or what the effect of not making certain changes to the part design will be. In addition, based on their experience, they may be able to offer alternate suggestions with regard to mounting, assembly, or test procedures. In the article by Beich, there is an excellent case study on concurrent engineering and the use of multiple prototype and mold methods to transition from the prototype design to production. In it, he provides a guide for use upon starting a program using plastic optics.⁶⁹

Most vendors are willing to sign nondisclosure agreements (NDAs), which allow and protect the transfer of customer proprietary information to the vendor (and potentially vice versa). The designer should clearly specify the proprietary nature of any information that is passed to the vendor, while the vendor should also specify any information they consider proprietary, such as tooling techniques. Careful labeling of documentation and the secure transfer of data, through methods such as encryption, should be considered.

Most companies have a procurement and/or supply chain person or group who deal with the contractual details associated with working with a vendor. These individuals should be contacted when beginning vendor discussions. In some cases, the vendor may not be on an “approved supplier list,” which may require the need for a site inspection before any orders can be placed. The designer should verify if their company has any history with a vendor, or if any special requirements (such as signed NDAs) need to be fulfilled before they can speak to them.

The selection of a plastic optics vendor can be driven by several considerations. In some cases, geography or nationality may be a factor. For instance, in the defense sector, selection of a manufacturer may be limited to domestic suppliers or those owned by allied countries. For medical devices, the associated need for certification of facilities may limit the supplier base. In the

case of consumer electronics, it may be required that subassemblies (such as optical systems) be manufactured in close proximity to the location of the final assembly facility. This can also sometimes be true in the automotive industry.

Cost can be a significant factor in the selection of a vendor. We mentioned earlier the movement of manufacturing towards lower-cost regions, a trend which will likely continue. When considering vendor selection based on cost, it is important to include not just piece price but also the cost of any required engineering support, travel, and on-site supervision.

Reputation and relationships can be another factor in selecting vendors. Having worked with the designer to develop a product through the prototype phase, the vendor that was used should have the best knowledge of both the elements and product. This may give them an advantage in understanding the needs and cost of making the parts. Basic trust can also be an issue, particularly with companies not known for the highest standards of business ethics. One of my colleagues experienced the interesting situation of walking into the “wrong” stockroom, only to find additional assemblies of our product labeled as the manufacturer’s own. While rare, these types of situations do exist and are best mitigated through supervision and vendor inspection.

Quality control and testing can be additional vendor selection criteria. Different industries and companies have various quality requirements, such as International Organization Standardization (ISO) certification. Most vendors have some kind of established quality control system, with appropriate procedures and documentation. Modern molding machines, computers, and e-mail allow the relatively easy collection, analysis, and transfer of quality control or SPC data, and most vendors have adequate storage for any necessary part retention. Available test equipment can sometimes play a role in selection of a particular vendor. For instance, a vendor may have an interferometer but not an MTF test station. If the specification on the system is a required MTF value, it must be determined if the use of an interferometer (instead of an MTF test) is adequate.

When placing an order, either for a prototype or for production quantities, clearly defined goals, schedules, and costs should be in place. For instance, if purchasing an injection mold, any nonrecurring engineering (NRE) costs, expedite fees, process development costs, mold compensation limits, required secondary process tooling, or guaranteed production should be understood by both parties. The process of qualifying the mold should also be known, whether based on a full first-article inspection (FAI) or some other method. If a pull-ahead mold is used, with a limited number of cavities initially used, the method of qualifying the additional later cavities should be agreed upon.

There are a fairly limited number of high-quality plastic optics vendors in any country or region of the world. Designers who will be developing plastic optics should make an effort to contact these potential suppliers and begin building a relationship with them. Most suppliers have one or more individuals that spend much of their time providing customer support and product development. These individuals are more than happy to be contacted and begin

discussions on a project. They can often be contacted at the exhibitions at various conferences or through the information on their websites.

We cannot overemphasize the advantage of working with vendors at the beginning of a project; it is in the best interest of both the designer and the vendor. Vendors themselves often state that they wish they would be contacted earlier. If nothing else is gained from reading this tutorial text, remember this: begin discussions with vendors at the start of a project.