Modern Optics Drawings

The ISO 10110 Companion

Modern Optics Drawings

The ISO 10110 Companion

Eric Herman
David M. Aikens
Richard N. Youngworth

SPIE PRESS
Bellingham, Washington USA

Library of Congress Control Number: 2021949589

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

Printed in the United States of America.
First Printing.
For updates to this book, visit http://spie.org and type "PM346" in the search field.

SPIE.

Table of Contents

Preface

The purpose of this book is to provide optical engineers, fabricators, and all parties in between a better understanding of the ISO 10110 drawing standard, and how to use the standard to create modern optical drawings. The authors presume that the reader has access to all parts of ISO 10110 and associated standards, and at least a basic familiarity with optics component technical drawings.

The world of standards is very small, and volunteer-based. Those who work on these committees are trying to do the right thing for our industry. Those few people are willing to take the time to draft, edit, and review these standards. Most of the time the result of their efforts is the result of compromise, and often it is far from perfect. Pursuit of an international standard that reflects the diversity of opinions, applications, and needs of the community comes at a price.

The first version of ISO 10110 was published in 1995 and has been the subject of multiple revisions, additions, and subtractions over the past 25 years. Today, there are twelve parts; 1, 5, 6, 7, 8, 9, 11, 12, 14, 17, 18, and 19. Because of this international effort over the past fifteen years, the ISO 10110 drawing standards have become a truly international standard reflecting the needs of the global optics community. It is vastly better today than in 1995 or 2005, and much more in harmony with US industry.

Even so, the standards can be difficult to read, more complex than we would like, and sometimes seem unfamiliar and unfriendly. We have written this book as a friendly guide to unfamiliar language, symbols, and a way of thinking about optics tolerances and specifications. However, the book is meant as a guide—not a replacement—for reading the standards themselves.

Additionally, there are a half-dozen other standards that are essential to using ISO 10110, and even more that can be used to make ISO 10110 more effective. As a result, this book is structured in chapters by subject, rather than by part of 10110, with related standards grouped within a chapter. Many practical examples are provided with a view toward a complete adoption of the methodology of standardized optics drawings including the drawing notation standards; and the metrology, environmental, and system performance test standards. It is the authors' hope that the book is readable enough to be read and understood by the uninitiated, and that the book serves as a useful reference or guide to users of the standard as they navigate the details of full implementation.

Acknowledgments

This book has been needed in the community for some time. It's incredibly important in the implementation of standards that people know how to use them. We'd first like to thank Bob Parks and Ron Kimmel for writing the previous iteration of this type of guide. Over the years, their book has been a great resource in learning the standards.

We want to thank the standards community, which has taken on the (often) thankless task of developing these standards and keeping them updated. The time and energy that has been put into these documents by a few hearty volunteers is a tribute to our community.

We want to thank the good people at ANSI, ISO, and DIN; without whom there would be no standards at all.

Throughout the development of this book, there were many people who assisted in discussions. We would like to thank Ray Williamson, Stan Schwartz, Michael Thomas, Dennis Leiner, and Daniel Gray; and the others we may have forgotten to mention.

We are grateful to Jeremy Govier, Cory Boone, Paul Smith, Mark Malburg, Eddie LaVilla, and Jim Wyant for assistance in providing necessary resources for us to create some of the figures in the book.

Mostly, we'd like to thank our partners for their patience and encouragement throughout the writing of this book. We wouldn't have been able to finish without their continuous encouragement.

Eric Herman
Dave Aikens
Richie Youngworth
September 2021

Chapter 1
Drawing Notation and Default Tolerances

ISO 10110: Optics and Optical Instruments – Preparation of drawings for optical elements and systems can seem complicated due to the coded tolerances and notation required. The barrier in understanding this standard is interpreting the coded notation. While initially challenging, this coded notation of a requirement becomes a significant and unambiguous benefit of the standard when analyzed further. The reduction of text description on a print increases the number of qualified fabricators a designer may use for their optics. The ISO system of normative callouts obviates the need for interpreting a specification and makes requirements clearer than without normative callouts regardless of communication factors. All that is required is an understanding and a proficient use of the standard.

1.1 Background

ISO Technical Committee 172—Optics and photonics (ISO/TC 172) was created in 1978 to address the need for worldwide optics and photonics standards. The Technical Committee is composed of seven subcommittees, each of which is tasked with developing standards for a different aspect of the optics and photonics community. Subcommittee SC 1, first convened in 1986, is responsible for fundamental standards for optics and photonics. The most important working group within ISO/TC 172/SC 1 for optics drawings is WG 2—which has been convened by various countries across Europe, the United States (US), and Asia. Within WG 2, the critical ISO 10110 series of standards was developed and is maintained.

The purpose of a technical drawing is to convey the requirements and tolerances of a component or system as clearly and unambiguously as possible. ISO 10110 strives to present optical components on drawings with a minimum number of notes and ambiguity. Clarity is achieved through use of a large array of symbols and indications. When the symbols and indications are combined with relevant dimensional data, ISO 10110 optical drawings initially appear enigmatic. However, the use of symbols instead of notes eliminates translation errors and misinterpretations that are common with notes-based drawings.[1]

The first edition of ISO 10110, published in the mid-1990s, had a separate internal standard (Part 10) for a tabular format of notation.[2] Since that time, the tabular format has become the preferred method of notation on an ISO 10110 drawing. In 2019, the tabular notation was added to Part 1 with new schema for tabular drawing layouts, and Part 10 was withdrawn. This tabular form of ISO 10110 is much easier to implement and is used for the vast majority of ISO 10110 drawings.

There is also a set of standard tolerances within ISO 10110, similar to block tolerances on mechanical drawings. In the ISO 10110 format, these standard tolerances are annotated in a different subsection.[3] This establishes an agreed-upon set of minimum tolerances for certain features of a lens element if the tolerances are inadvertently left off the element drawing. These default tolerances are looser than one may expect when used for an element.

1.2 Differences between ISO and US Standards

Standards for optics within the US were first published in the 1950s and 1960s. These standards (often called MIL standards) were originally written for military hardware procurement but were applied across the entire industry. They covered a wide variety of topics; many of them have since become inactive,[4–7] whereas others are still in use today.[8,9] Among the topics covered by the MIL standards were optics drawings, specifications, and notations; these standards were adequate at the time for their intended use. Since that time, optics drawings proceeded to convey more-complex optical elements as manufacturing technology improved, and applications grew and greatly diversified. Creating a modern optics drawing using these inactive standards leads to a large employment of notes and depends greatly on additional communication with fabricators.[10]

1.3 Overview of Coded Notation

The key to ISO 10110 is understanding the coded notation for tolerances. An overview of each tolerance, the coded notation used, and the attributable ISO 10110 subsection, or Part, is shown in Table 1.1.

1.4 Fundamental Information

On an ISO 10110 drawing, there are default units associated with each specification or dimension unless noted. All linear units are specified in millimeters (mm) if there is no unit callout within the title block or on the

Table 1.1 Coded notation table.

Type of Tolerance	Coded Notation	ISO 10110 Section
Material Stress Birefringence	0/	10110-18
Material Permissible Bubbles and Inclusions	1/	10110-18
Homogeneity and Striae Classes	2/	10110-18
Surface Form	3/	10110-5
Centering	4/	10110-6
Surface Imperfections	5/	10110-7
Laser Irradiation Damage Threshold	6/	10110-17
Transmitted Wavefront Deviation	13/	10110-14
Surface Texture	$\surd$	10110-8
Surface Treatments and Coatings	(λ)	10110-9

specific dimension. Each value can either be specified with a decimal point or decimal comma as the delimiter. The type of decimal used must be consistent across the entire drawing. To not confuse the decimal notation, a delimiter for every thousands should not be used.

Since some information become extremely temperature-dependent, the default temperature on the drawing is 20°C. It is acceptable to specify a dimension or tolerance at a different temperature, but this must be specified either at the exact specification or within the title block.

As this is an optical component drawing, the wavelength is used to specify material information. Additional tolerances may require a specified wavelength as well. Historically, a default wavelength of the Fraunhofer Mercury e wavelength was used,[11] but this has become updated to note that a wavelength may be specified on the drawing within either the title block or on the individual value. This is relevant when considering the surface form or transmitted wavefront. If the surface form or transmitted wavefront is specified in nanometers (nm) or micrometers (μm), then a default wavelength is not critical. A noted wavelength is necessary when the surface form or transmitted wavefront is specified in fringes or waves. The default units for wavelength are nanometers unless specified otherwise.

Lastly, to ensure that all interpretations are as per ISO 10110, a statement making this claim must be on the drawing. A note that the drawing follows ISO 10110 is typically found within the title block (Fig. 1.1).

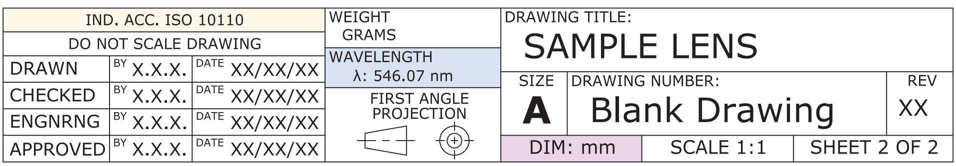


Figure 1.1 Title block.

1.5 Table and Drawing Field

Within an ISO 10110 drawing, it is possible to list the specifications for a single optical component, a cemented component, or an entire optical subassembly. Through most of the discussions within this book, a single optical component is discussed for ease of explanation. These can be expanded upon for optics drawings of various other types of optical assemblies and systems.

ISO 10110 specifies that a drawing is broken up into three fields (Fig. 1.2): the table field, the drawing field, and the title field (or title block).

For a drawing of a simple optical component with two optical surfaces, the table field is broken into three subfields. The leftmost subfield refers to the requirements of the left surface (or surface 1), and the rightmost field refers to the requirements of the right surface (or surface 2). The middle field refers to the requirements of the material.

The drawing field will contain a view of the optical component and its dimensional information, and any information not provided in the table field, such as additional notes and notations that apply to the entire component. Any datums necessary for interpretation of the centering requirement will be

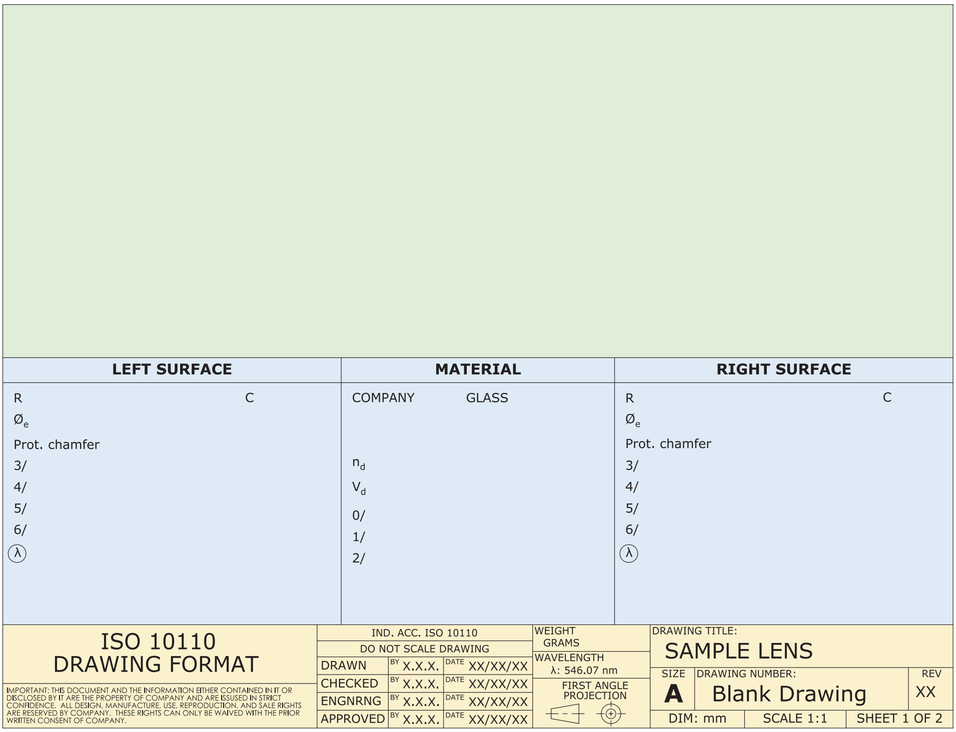


Figure 1.2 Table and drawing view.

shown in the drawing field. Note that coating information and the surface texture tolerances are often indicated in the drawing field as well.

The title block will contain the information needed to distinguish the part without any additional information. A title block will at least contain the name of the part, drawing projection, scale, units, revision number, and indication that the drawing is following ISO 10110. Drawing title blocks are described in depth in both ISO 128 and ASME Y14.1.[12,13]

The 2019 edition of ISO 10110 added several new schemas for the tabulated drawing layout. In the case of a single component with three or more optical surfaces, e.g., a prism, additional subfields are added to the table field for each additional surface. All surface subfields will be labeled to indicate the referenced surface. The subfields can be oriented vertically or horizontally. Wherever possible, the subfields should follow the path of the light. Additional schema are given in Part 1 for optical assemblies such as doublets or triplets.

While the tabular format is recommended (Fig. 1.3), it is not required. In an alternative drawing layout (Fig. 1.4), no table field is used, and the requirements and tolerances for the optical component are specified in the drawing field. The drawing field will contain a view of the optical component and the information to manufacture said component.

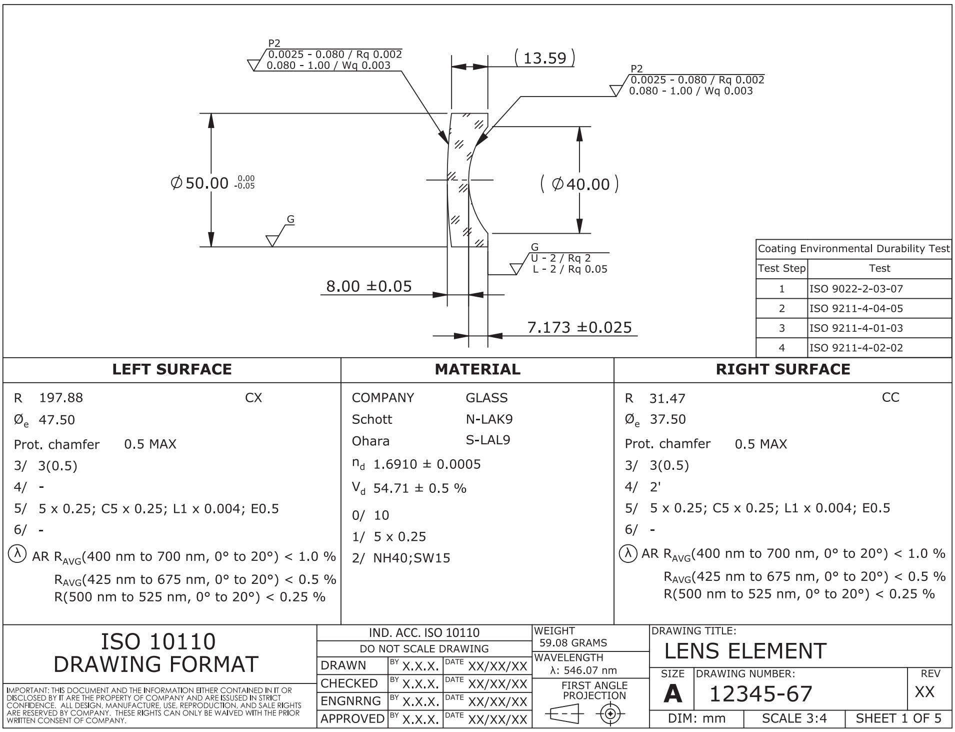


Figure 1.3 Tabular drawing format.

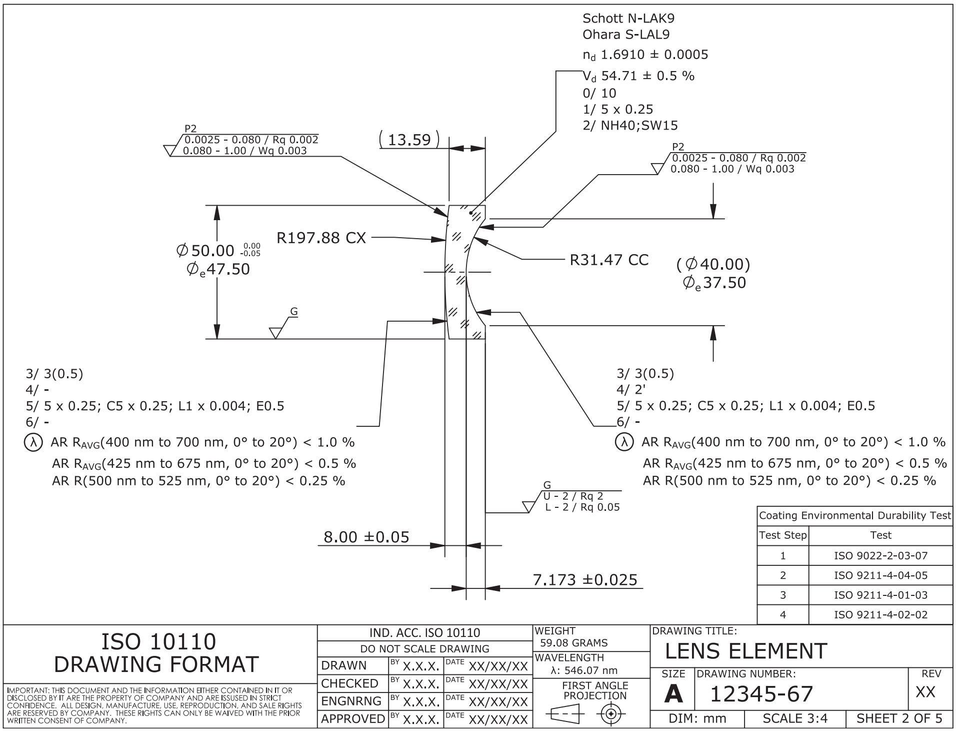


Figure 1.4 Nontabular drawing format.

1.6 Indications on Drawings

Views – Hatching

For a simple rotationally invariant optical component, either a side view or a cross-sectional view is necessary. If possible, the view should be shown with light going from left to right and the optical axis horizontal. Back edges and hidden lines can be omitted. This reduces the number of views and confusion in interpreting the shape of the component. If the optical component requires multiple orientations or views, then the orientations or views should be shown individually. When an optical component is shown in cross-section, any cross-hatch must be shown in a glass style (short–long–short) at 45° (Fig. 1.5). Cross-hatching can be omitted for improved drawing simplicity.

Coordinates

In general, it is assumed that the optical drawing is shown in the orientation with $+z$ going towards the right and $+y$ going up. It is not necessary to show the coordinate system used, but if the views of the drawing differ from the $+y$ / $+z$ orientation the coordinates must be specified to clarify how the views should be interpreted. If there are multiple views on a drawing, the coordinate system must be shown in each view.

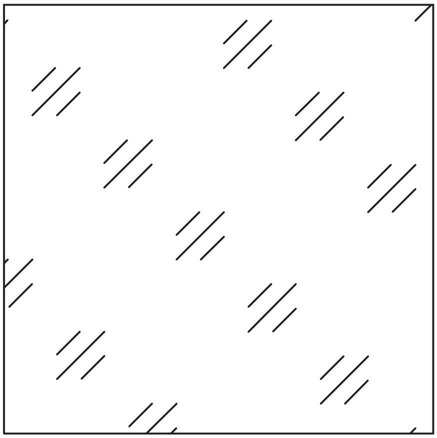


Figure 1.5 Hatching.

It is sometimes also necessary to indicate a coordinate system for different surfaces in complex shapes. These can be shown in the drawing field for each surface or feature as needed. If the global coordinate system is not the vertex of the left surface, the vertex should also be indicated on the drawing. The coordinate transforms for each local coordinate system can be given in the table field with a series of three translational displacements and three angular displacements with respect to the global coordinate system; e.g., the notation (0, 0, 50) 180°, 0°, 0° would indicate a coordinate system that is translated 50 mm along the z-axis and then rotated 180 degrees about the x-axis. Figure 1.6 shows an example of the notation for a global coordinate system for an optical component where the vertex is not the origin point.

Standardized Coordinate Systems for Aspheric and General Surfaces
Aspheric surfaces and general surfaces, including freeforms, have standardized coordinate systems—as will be discussed in Chapters 8 and 9. If drawings specify such surfaces, it is not necessary to include the coordinate system unless there is a deviation from the standardized default.

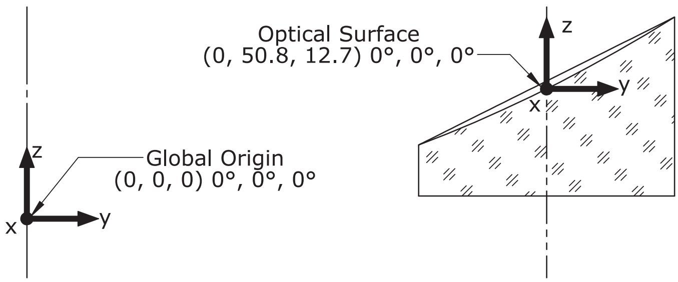


Figure 1.6 Global coordinate systems.

Figure 1.7 (a) Optical axis and (b) centerline.

Note that if a surface is on the right side, the default coordinate system used in ISO 10110-12 may cause confusion because some countries (such as the US) use the $+z$-axis into the part on both sides of the element. Care must be taken in the drawing to be sure the z-axis used is understood by all parties.

Axes

On the drawing, the optical axis is assumed to be the centerline. If the optical axis and the centerline are not coincident, then each should be specified individually. The optical axis notation is shown by a dash–dot–dot–dash line [Fig. 1.7(a)]. The centerline notation is shown by a dash–dot–dash line [Fig. 1.7(b)]. The optical axis notation takes priority if it is being used coincident with the centerline.

Leader Lines

It is often necessary to show a leader line for a dimension or annotation to the optical component. These are specified with an arrow and a subsequent line if the information is directly applied to the surface or edge. If the information is being conveyed to the material itself, the leader should be a dot with a subsequent line to the dimension or information.

Test Regions

For certain optical components, there may be different regions of the part that require different requirements. With these different regions, specifying the aperture at which the test conditions are required is important. An example of this is a lens with different clear aperture diameters or multiple clear apertures with different requirements on a single surface (Fig. 1.8).

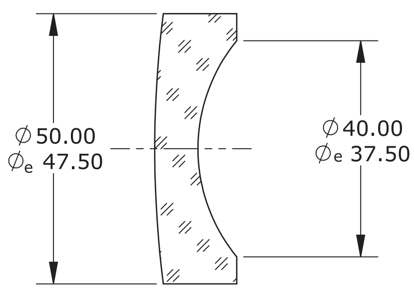


Figure 1.8 Test regions on a singlet.

Dimensional

There is certain dimensional information necessary for an optical component that is not included in the coded notation of ISO 10110. These are mechanical dimensions associated with the optical component such as radius, thickness, and diameter. Each of these should be listed in either the drawing view or on the tabular column.

For the radius value, the dimension should be preceded by an R to show that this is in fact the radius value. If the radius is listed in the tabular field, a specification of the orientation of curvature (convex [CX] or concave [CC]) should be listed along with the value. If there is only an interferometric tolerance used on the radius value, then no tolerance should be listed along with the value. If the surface is planar (i.e., plano), then the radius should be shown as R∞ or R PL. As there are more types of optical surfaces than either plano or spherical, the drawing where an R would be present is replaced by the appropriate designator. A list of these designations is shown in Table 1.2. Surface types that are not listed in Table 1.2 must be written out in full and will not have a shorthand designator.

ISO 10110-1 is suitable for defining spherical, plano, and cylindrical surfaces. ISO 10110-12 and ISO 10110-19 provide a framework for designating low-order quadrics, aspheric surfaces, and general surfaces that can include freeforms. The latter two Parts are discussed further in chapter 9.

For a thickness value with no tolerance indicated, a default tolerance is used. These can either be a sheet tolerance or the default tolerance from ISO 10110-11.[3]

A diameter should be listed with a preceding ∅ symbol. If there is a difference between the physical diameter and the clear aperture diameter, the clear aperture diameter (or optically effective diameter) should be listed as $\varnothing_e$ (Fig. 1.9).

On an optical component, there will be sharp, uneven edges due to the manufacturing process. To replace these sharp edges, a chamfer or bevel is applied. A chamfer is nonfunctional and defined by a small, angled cut that is specified only by the linear dimensional range. A bevel is functional and defined by a small, angled cut that requires not only the linear dimensional range but also the angular dimensional range.

Table 1.2 List of surface types and their respective designators.

Surface Type	Surface Designator
Spherical	R
Plano	R∞ or R PL
Cylinder	CYL
Rotationally Invariant Aspheric	ASPH
Rotationally Variant Surface	GS

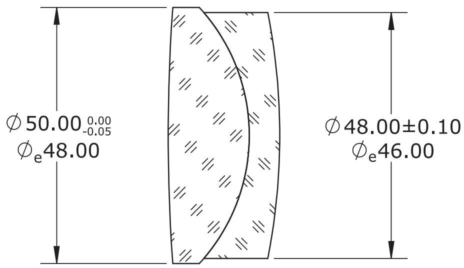


Figure 1.9 Indication of diameter and clear aperture.

Materials

The definition of the optical material is critical to the specification of the optical component. The optical material type may be defined in multiple ways. An optical material should be defined by the manufacturer and the manufacturer's trade name (e.g., Schott N-BK7 or Ohara S-BSL7), the index of refraction and Abbe (or dispersion) number (e.g., n_d = 1.5168, V_d = 64.2; or n_d = 1.51633, V_d = 64.1), or the international glass code (e.g., 517642 or 516641). It is possible to define the optical material on a drawing by multiple terms, such as the manufacturer, glass type, index of refraction, and the Abbe number.

If the material is crystalline, the material chemical formula may be used (e.g., MgF_2 or Si). In this case, it may be necessary to reference the crystallographic orientation for the component. This can be indicated using Miller indices as described in IEC 60758:2016.[14]

Once the optical material is listed, tolerance information for the index of refraction and Abbe number must be defined. Further material tolerance information will be discussed in chapter 2.

1.7 Subassemblies

As previously mentioned, it is possible to create an ISO 10110 drawing for more than a single optical component. Creating drawings for doublets or even full subassemblies is acceptable.

When discussing the layout of a doublet drawing, the number of columns should grow from three up to six for the entire doublet to be specified on a single print (Fig. 1.10). The additional columns represent each of the additional surfaces and materials. It is also possible to consider a doublet as a subassembly and annotate it accordingly, with separate drawings for each of the elements and then an assembly drawing whose tolerances only apply to the finished assembly.

With doublets, there are a few things of note to ensure the part is properly defined. A note is added specifying where the surface is to be cemented and type of cement to be used. The surfaces that will be cemented together may

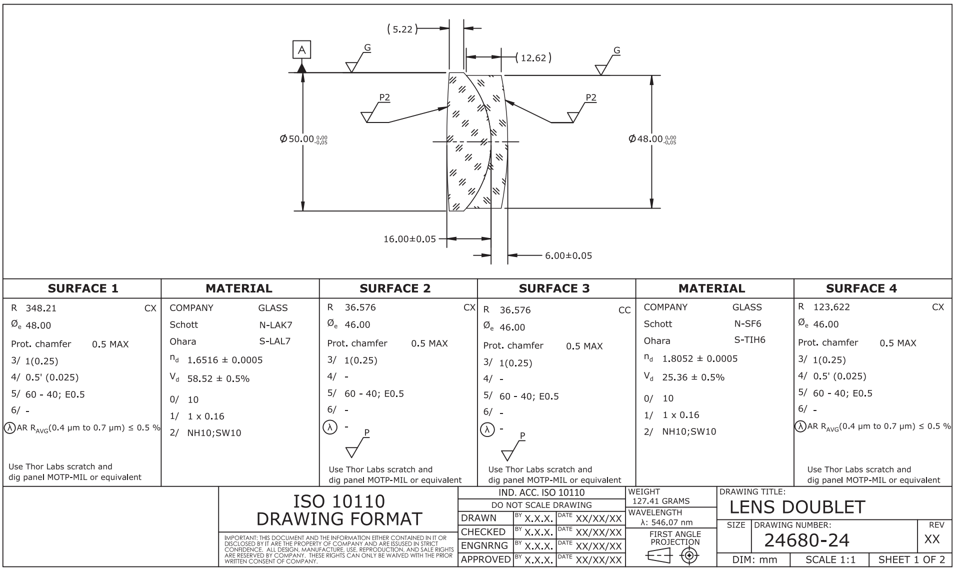


Figure 1.10 Doublet drawing.

also have a note that they are "to be cemented," ensuring that the fabricator understands the polishing requirements for said surfaces.

If the thickness tolerance for the entire doublet is specified, independent of single element thicknesses, the letter `M` is added (e.g., `9 M ± 0.2 M`) to the thickness dimension tolerance. This ensures that the thickness is matched for both parts.

In a drawing of a subassembly more complex than a cemented doublet, it is desirable to simplify the drawing. Instead of having all of the relevant surface and material data for each component in tabular format on the assembly drawing, the component requirements should be listed on additional sheets or independent drawings. Each of the parts must be specified by its reference part number. All dimensions and datums must be specified and toleranced accordingly.

With a subassembly drawing, it is possible to include optical information relevant for the system itself. ISO 10110-1:2019 Clause 6 includes a list of additional system information that might be added to the assembly drawing. For example, the inclusion of a pupil position or an image plane is acceptable with the proper notation (Fig. 1.11). To include the image location for a subassembly, an $\times$ is indicated along the optical axis. Listing a pupil position is done similarly except with a $|$. See Clause 6 for additional information.

(a) (b)

Figure 1.11 (a) Image plane and (b) pupil plane.

1.8 Non-Toleranced Data (Default Tolerances)

For any optical component, there are tolerances that are applied to all specifications. In ISO 10110, for many of the common requirements, if a specification does not have an associated tolerance with it a set of default tolerances are applied. For linear or angular values, there are typically sheet tolerances listed on the drawing that are applied given the number of digits that follow the decimal places. For optical component drawings, a list of sheet tolerances may not be listed in the title block.

These dimensional values and optically required values have their own set of default tolerances that are tabulated in ISO 10110-11. These tolerances are listed based on the diameter ranges of the component. The diameter ranges are broken up into four groups: between 2 mm and 10 mm, between 10 mm and 30 mm, between 30 mm and 100 mm, and between 100 mm and 300 mm. The tolerances associated for each range are very loose relative to manufacturing capabilities and are meant to be a safeguard against missing data. These tolerances are considered either commercial-grade or looser.[15] As shown by DeGroote et al.,[16] for tolerances of a diameter between Ø 30 mm and Ø 100 mm, these tolerances may not be acceptable for diffraction-limited or high-performance designs. It is possible to use tolerances that are looser than default tolerances, but looser tolerances must be explicitly listed.

If a dimension is not listed but a dash is listed instead with the associated code, no requirement should be applied to the given tolerance. There is no default tolerance for laser damage threshold or for some other less-common specifications. It is necessary to list the surface texture on a print, even if it only specifies that a surface is polished or ground.

1.9 Discussion of Other Standards in Use

1.9.1 MIL-STD-34[4]

MIL-STD-150 defines terms used in photographic lenses and how to analyze an optical system.[17] Although many critical aspects of optical fabrication are defined, and metrology methods discussed within MIL-STD-150, no single drawing standard is provided. MIL-STD-34 was first introduced in 1960 to build on the mechanical drawing standards, MIL-STD-8 and MIL-STD-150; MIL-STD-34 added optics-specific properties and their notations.[4,17,18]

MIL-STD-34 created a singular method for how optics drawings should be implemented. As with ISO 10110, there is a commonality throughout MIL-STD-34. The key difference between MIL-STD-34 (and the additional inactive optical drawing standard ASME/ANSI Y14.18M[19]) and ISO 10110 is that MIL-STD-34 conveys most information in written terms, either through dimensions or in notes. Because of the large amount of information that must be discussed, drawings following MIL STD-34 look significantly cluttered.

Information conveyed through notes includes focal lengths, clear aperture, surface quality, and centering. Material definitions of the optical element are unexpectedly within the title block of the drawing. This definition of properties, such as units and material information, follows standard mechanical drawings.

As MIL-STD-34 was based on MIL-STD-8, dimensional properties are defined on the view of the optical element. This reduces reliance on notes for defining an optical surface. Some tolerances for these dimensions are stated as words, such as the power and irregularity in fringes. These dimensions are more commonly stated in waves or nanometers on modern drawings.

There are several issues with using a MIL-STD-34 format drawing. First, this type of drawing requires the fabricator to find additional key information. A fabricator must break down the drawing to understand the dimensions or tolerances pertinent to the requirements. Second, the standard uses shorthand for tolerances based on a presumed manufacturing method that may be incorrect or obsolete. Finally, the standard was written in 1960, cancelled in 1995, and has not been updated in more than 50 years.

1.9.2 ASME/ANSI Y14.18M

ASME/ANSI Y14.18M[19] was created to update and supersede MIL-STD-34. The goal for this standard was to move control of optics drawing standards away from the US military to a voluntary engineering organization. ASME/ANSI Y14.18M relies heavily on ASME Y14.5 except when noting an optical dimension or tolerance.[20] At the time of the implementation of ASME/ANSI Y14.18M in 1986, the ISO standards for optics drawings were not in place.

The authors of ASME/ANSI Y14.18M tried to expand MIL-STD-34 to include further necessary definitions within an optics drawing. The former standard increases the overall number of documents to reference. These include ANSI PH3.617 and ANSI Y14.36, in addition to ASME/ANSI Y14.18M.[21,22] Breaking up the complexity within a single standard encourages more-widespread implementation due to ease of execution.

This optics drawing standard was last revised in 1986 and was deactivated in 2006. Since that time optical manufacturing methods have grown to include a wider variety of optical elements, including more prevalent aspheres and the emergence of freeform optical surfaces.

One of the key benefits of ASME/ANSI Y14.18M is the open-endedness of the drawing requirements. Users can create a custom drawing template and respective standard implementation. This allows companies and users to follow the standard and still create a drawing based on how optical designers communicate with their fabricators.

Each dimension within ASME/ANSI Y14.18M is defined, but the corresponding explanation is ambiguous. The standard requires significant designer and fabricator communication for effective implementation. All dimensional values and tolerances follow ASME Y14.5. Values that have a tolerance outside of what can classically be defined within ASME Y14.5 are outlined in ASME/ANSI Y14.18M. Similarly to MIL-STD-34, the definition and explanation for how to handle a nonspherical surface requires use of an equation and a method for verifying said equation; e.g., a sag table.

Example drawings within the ASME/ANSI Y14.18M standard are included to assist in the creation of a template. The standard contains an explicit statement that sample drawings are for reference. Users may express their drawing in any way they wish. As a result, the reliance on notes subsequently is commonplace within US optical drawings.

Many critical values or callouts of an optical element are defined within a note or table. The reliance of notes is both a major benefit and a major restriction of the standard. Unlike MIL-STD-34, the ASME/ANSI Y14.18M standard allows the designer to communicate through their own methods. Because of the reliance on notes, ASME/ANSI Y14.18M limits which fabricators a designer may use. A designer typically must communicate back and forth with a new fabricator to ensure that the designer's intention comes across with the verbiage chosen for a dimension, tolerance, or method of test. The inevitable need for notes results in a potentially confusing optical drawing. The freedom of both standards induces a need for continual communication between the optical engineer and fabrication staff.

1.10 Drawing Example

A discussion of basic dimensioning for an ISO 10110 drawing is best illuminated by an example. Being able to associate the terms with a sample drawing allows users to move forward with creating their own drawings in the future.

Each of the optical terms defined for the drawing in Fig. 1.12 is broken down by the annotated letters in Table 1.3 below.

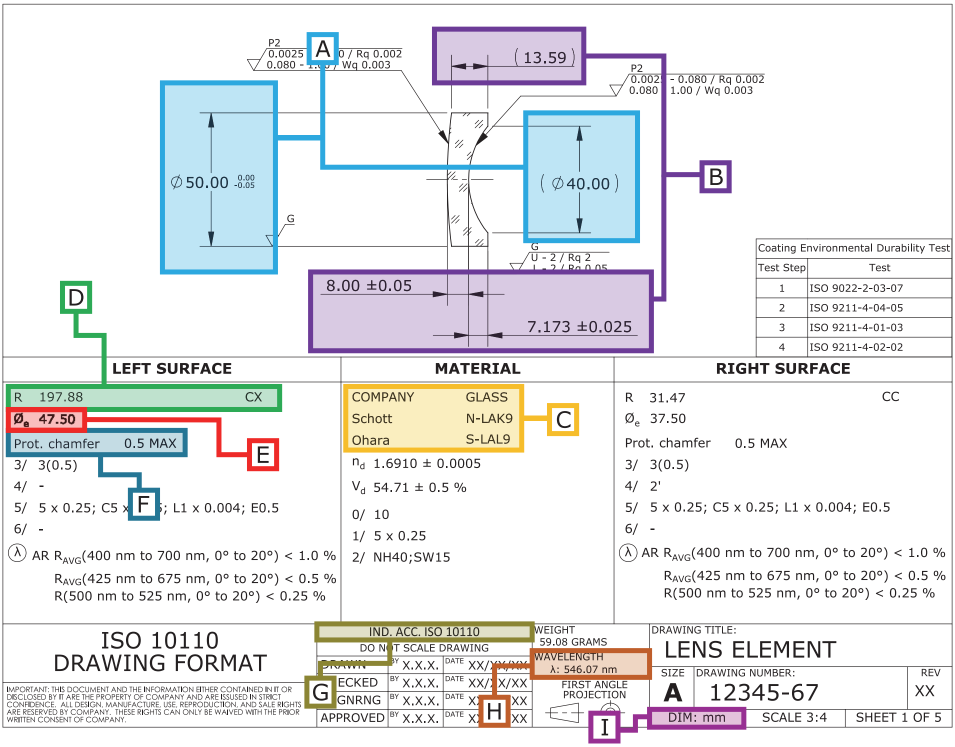


Figure 1.12 Example drawing.

Table 1.3 Example drawing annotation.

Annotated Letter	Description	Explanation
A	Element Diameter	The drawing lists two diameters: the outer diameter Ø 50 mm and the smaller diameter Ø 40 mm that leads to the annulus.
B	Thicknesses	There are three thicknesses listed on the drawing: center thickness (8 mm), sag distance (7.173 mm), and reference edge thickness (13.59 mm).
C	Material Information	On this print, the material information is defined by both the company and glass type (either N-LAK9 or S-LAL9), along with the d-line index (1.6910) and Abbe tolerance (54.71).
D	Radius and Curvature Orientation	There are two radii listed on the drawing. The radius annotated by letter C is for the left surface of 197.88 mm and specifies the curvature orientation of convex (CX).
E	Clear Aperture	The clear aperture is defined by the diameter (47.50 mm on the left surface) and subscript e ($Ø_e$). This diameter is where the optical aperture will be tested over.

(*continued*)

Table 1.3 (*Continued*) Example drawing annotation.

Annotated Letter	Description	Explanation
F	Protective Chamfer	Notating that a protective chamfer is acceptable on the optical surface is noted with a maximum of 0.5 mm. This allows the fabricator to put on a nonfunctional chamfer without worrying about the functionality.
G	Drawing Interpretation	To ensure that the drawing notations are interpreted properly, the indication standard (ISO 10110) is clearly defined for the fabricator.
H	Reference Wavelength	In order to interpret the surface form tolerances, the reference wavelength that should be used on a drawing is required (546.04 nm).
I	Reference Dimension and Projection Angle	Even though it is formally defined in ISO 10110, clearly defining the units for the drawing may be useful as well as the projection used for presenting the view, especially for drawings within the US.

References

1. International Organization for Standardization (ISO), *Preparation of Drawings for Optical Elements and Systems—Part 1: General*, ISO 10110-1:2019 (Geneva: 2019).
2. ISO, *Preparation of Drawings for Optical Elements and Systems—Part 10: Table Representing Data of Optical Elements and Cemented Assemblies*, ISO 10110-10:2004 (Geneva: 2004).
3. ISO, *Preparation of Drawings for Optical Elements and Systems—Part 11: Non-Toleranced Data*, ISO 10110-11:2016(en) (Geneva: 2016).
4. US Department of Defense (USDOD), *Preparation of Drawings for Optical Elements and Optical Systems; General Requirements for*, MIL-STD-34 (Arlington, Virginia: 1960).
5. USDOD, *Military Standard: Surface Roughness, Waviness and Lay*, MIL-STD-10A (Arlington, Virginia: 1963).
6. USDOD, *Military Specification: Glass, Optical*, MIL-G-174B (Arlington, Virginia: 1986).
7. USDOD, *Military Specification: Mirror, Front Surfaced Aluminized; For Optical Elements*, MIL-M-13508C (Arlington, Virginia: 1973).
8. USDOD, *Military Specification: Coating of Glass Optical Elements (Anti-Reflection)*, MIL-C-675C (Arlington, Virginia: 1980).
9. USDOD, *Performance Specification: Optical Components for Fire Control Instruments; General Specification Governing the Manufacture, Assembly, and Inspection of*, MIL-PRF-13830B (Arlington, Virginia: 1997).
10. R. N. Youngworth, S. R. Kiontke, and D. M. Aikens, "Standards: a key element of optical design, engineering productivity, and time to market," *Proc. SPIE* **9293**, 929317 (2014) [doi: 10.1117/12.2075347].
11. ISO, *Reference Wavelength*, ISO 7944:1998 (Geneva: 1998).

12. ISO, *Technical Drawings—General Principles of Presentation*, ISO 128-1:2020(en) (Geneva: 2020).
13. American Society of Mechanical Engineers (ASME), *Decimal Inch Drawing Sheet Size and Format*, ASME Y14.1-2005 (New York: 2005).
14. International Electrotechnical Commission (IEC), *Synthetic Quartz Crystal—Specifications and Guidelines for Use*, IEC 60758 (Geneva: 2016).
15. Optimax Systems, s.v. "Manufacturing Tolerance Chart," *last accessed September* 16, 2021, http://www.optimaxsi.com/innovation/optical-manufacturing-tolerance-chart/.
16. J. DeGroote Nelson, R. N. Youngworth, and D. M. Aikens, "The cost of tolerancing," *Proc. SPIE* **7433**, 74330E (2009) [doi: 10.1117/12.828433].
17. USDOD, *Military Standard: Photographic Lenses*, MIL-STD-150A (Arlington, Virginia: 1959).
18. USDOD, *Military Standard: Dimensioning and Tolerancing*, MIL-STD-8C (Arlington, Virginia: 1963).
19. ASME, *Optical Parts*, ASME/ANSI Y14.18M-1986 (New York: 1986).
20. ASME, *Dimensioning and Tolerancing*, ASME Y14.5-2009 (New York: 2009).
21. American National Standards Institute (ANSI), *Definitions, Methods of Testing, and Specifications for Appearance Imperfections of Optical Elemends and Assemblies*, ANSI-PH3.617 (New York: 1993).
22. ASME, *Surface Texture Symbols*, ASME Y14.36M (New York: 1996).

Chapter 2
Optical Materials

In addition to the material properties (e.g., index, dispersion, manufacturer, and material type) discussed in Chapter 1, there are material tolerances that are key to specifying optical component substrates. These additional essential material tolerances are stress birefringence, bubbles and inclusions, homogeneity, and striae. While these material tolerances are usually thought of in terms of optical glass, they can also be present in components made from other materials. Each of these tolerances can impact the overall optical performance, albeit they require an understanding of the material fabrication process.[1,2]

Stress within optical materials is a mechanical stress that is typically a function of the type of material and the annealing process, creating a birefringent effect in transmissive optical elements. This stress in the refractive element leads to an anisotropic optical retardation between the electric field component directions. Stress birefringence variation with temperature can cause thermal instability in both the refractive and reflective optical elements.[3] This is especially true for optics made from plastics and some crystal materials.

During the melting and annealing processes multiple types of imperfections can arise. Pockets of air and small contaminants, typically toleranced together and referred to as bubbles and inclusions, are produced from the raw materials in the melting and thermal cycling processes. These imperfections in the optical material lead to scattering effects within the optical material.[4] In the melting and annealing cycles, both occurring over a thermal cycling process, variations of refractive index continuity are created. The variations are known as striae or homogeneity. The distinction between these two types of refractive index variations is attributed to the range length through the optical material.[5,6]

2.1 Background

Preparation of an optical material may lead to imperfections in the blank optic and subsequently the final component. Each of the stress birefringence, bubbles and inclusions, homogeneity, and striae imperfections have varying impacts. In recent years, the ISO 10110-18 standard simplified understanding these material tolerances by combining the previous iterations of the standards.[7] The previous standards (ISO 10110-2, -3, and -4) were independent regarding each of these material tolerances. By combining the standards, a clearer understanding of how to define a material has become evident.[8] Concurrent with the combination of ISO 10110-2, -3, and -4, the standard for raw optical glass specifications, ISO 12123 was updated and harmonized with the drawing notation standard. The updates to ISO 12123 and the creation of ISO 10110-18 allows a uniform type of specification for both raw optical glass and the finished optical component.[9,10]

2.2 Finished Part versus Raw Material

Defining an optical component may require specifying the raw material as well as the finished optical component. With the commonality in standards, specifying the raw optical material is similar to the final optical component in the notation. Although the notation for each material tolerance works for both cases, tolerances can vary in the different steps of the fabrication process. Some tolerances are difficult to evaluate at one stage of production and easier to validate in another. For example, measuring the homogeneity of a finished lens is very difficult, but measuring the homogeneity of the blank can be relatively straightforward.

The parameters discussed for raw material and finished optical components are those described in ISO 10110-18 as well as the refractive index, Abbe number or dispersion. Different possible specifications are implemented with the ISO 10110 notation to ensure that a tolerance is kept on the optical material for the raw material and/or the finished component. When specifying a tolerance at the raw optical material level in ISO 10110 the coded notation is preceded by the number `0`. As shown in Figure 2.1, the same optical component may have the material tolerances specified for both the raw material and/or the final component. Users of ISO 10110-18 can also choose to specify the tolerances solely for the finished component or the raw material on their print.

The notation for the raw optical glass, as described in ISO 12123, is only applicable to the glass blank. This is useful for a fabricator communicating specifications during initial glass procurement. As some users may specify material properties for both the raw optical material and the finished optical component together, use of ISO 10110-18 notation is acceptable for both the raw material and finished optical component tolerances. While it is acceptable to use the ISO 10110-18 notation for raw optical glass, the ISO 12123 notation is preferred and should be used when appropriate. Note also that while ISO 12123 was written specifically in the context of optical glass, the same notation may be applicable to specifying optical blanks of other materials as well.

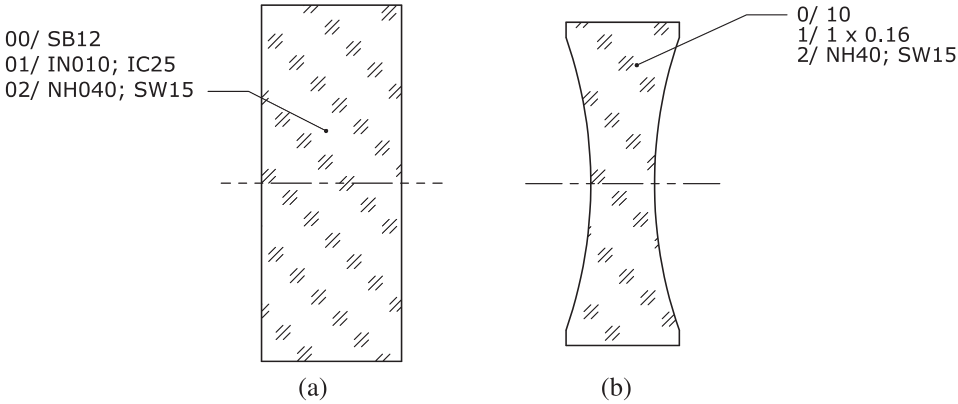


Figure 2.1 Differing specifications: (a) raw material and (b) finished optical component.

2.3 Overview of Material Properties

Each of the material properties are derived from the melting and fabrication process. Note that specifications for raw optical materials do have an impact on the final optical component.

2.3.1 Refractive index

Both refractive index and Abbe number were discussed previously in Chapter 1.[11] These parameters drive a given glass selection in the optical design process and specify the effective speed light will travel through a medium.

The Abbe number is the ratio of the refractivity divided by the principal dispersion. In general, the Abbe number is specified for either V_d (dFC spectral lines) or V_e ($eF'C'$ spectral lines). For visible light, specifying the V_d or V_e is acceptable as long as the notation is properly conveyed. The differences between these two parameters are where the center wavelength (d or e) lies, and how blue and red are defined. Use of the e-line specification puts a heavier emphasis on blue correction. Specifically, these Abbe numbers are defined by

$$V_d = \frac{n_d - 1}{n_F - n_C} \tag{2.1}$$

or

$$V_e = \frac{n_e - 1}{n_{F'} - n_{C'}} \tag{2.2}$$

where:

n_d = refractive index at 587.56 nm (helium d-line),
n_F = refractive index at 486.13 nm (hydrogen F-line),

n_C = refractive index at 656.27 nm (hydrogen C-line),
n_e = refractive index at 546.07 nm (mercury e-line),
$n_{F'}$ = refractive index at 479.99 nm (cadmium F'-line), and
$n_{C'}$ = refractive index at 643.85 nm (cadmium C'- line).

These Abbe numbers are defined in ISO 7944.[12] ISO 7944 also includes other reference wavelengths that may be used to specify dispersion for wavebands other than dFC and $eF'C'$ wavelengths. Care should be taken in specification with different wavebands to ensure that the requirements are understood by everyone in the procurement and fabrication process chain.

2.3.2 Stress birefringence

During the melting process, if heating and cooling cycles occur too rapidly, mechanical stress develops inside the material. This stress results in a birefringence that causes an optical path difference between electric field components in orthogonal axes.[3] Thus, stress birefringence may result in different focal lengths and other aberration effects between electric field components in the orthogonal directions.

The amount of stress birefringence may vary greatly by the type of optical material and manufacturing method. For example, molded or plastic optics may exhibit higher amounts of stress birefringence than found in traditional glass optical components.[13]

2.3.3 Bubbles and inclusions

As glass is refined in the melting process gaseous bubbles are created and are essential in minimizing inhomogeneity.[4,14,15] Although these bubbles are important in the formation of the glass from a melt, they are problematic for a finished optical component. The glass manufacturer removes these bubbles with temperature adjustment (causing the bubbles to come to the surface) and chemical reactions (fining agents in the chemistry). Nevertheless, some residual bubbles can exist. Additionally, contaminants or inclusions may remain in the glass from the fabrication process. This is due to particles from either previous melts or fabrication equipment containing residual contaminates.

Minimizing both bubbles and inclusions is important because they scatter light within the optical path. In certain planes of an optical system, bubbles or inclusions may be imaged through the optical system and create unwanted structure in an image, degrading the optical performance. Such errors can be critical in certain types of optics such as reticles.

2.3.4 Homogeneity

Homogeneity is the uniformity in the index of refraction throughout a material blank.[5] Inhomogeneity can be caused by a thermal gradient in the melting process of the glass. This results in a gradual variation of the consistency of the refractive index over the optical path length. Inhomogeneity leads to slowly varying optical path variations (i.e., gradient index effects) over the entire optical component rather than just a small region. Increases in the size of the optical component may increase the possibility of appreciable inhomogeneity across an optical component.

2.3.5 Striae

Striae are variations in refractive index over a short scale.[6] Unlike homogeneity, striae may be found in different regions across the optic and vary with propagation direction. Striae are caused by not allowing the material to become homogenous during the melting process. Striae usually appear as swirl patterns in the optic. Most common optical glasses are extremely free of striae, but more exotic or rarely melted glasses such as filter glass or plastics can exhibit significant striae in some directions.

2.4 Indications on Drawings

There are a few parameters between ISO 12123 and ISO 10110-18 that do not overlap with their overall notation. When following ISO 12123, the notation required for a raw material may not require an optical drawing. In contrast, the purpose of the ISO 10110 series is to create optical drawings. A larger description of the differences between ISO 10110-18 and ISO 12123 is discussed in Sec. 2.4.5.

2.4.1 Material property notation

As discussed in Chapter 1, basic material information such as manufacturer and glass type, international glass code, index and Abbe number, or material chemical composition (e.g., MgF_2) is listed within the material section of the print. There are no specified drawing symbols or qualifiers for these terms. The tolerance of index and Abbe number are typically listed explicitly as values without an associated notation code.

2.4.2 Stress birefringence

The drawing symbol for stress birefringence is `0/`. The quantifier added after the `0/` is the amount of stress birefringence, in retardation versus distance traveled, allowed in a part in units of nm/cm. Calculation of stress birefringence is found by analyzing the optical path difference between the two orthogonal axes. Additionally, determining the amount of stress within the optic leads to a calculation of the stress birefringence

$$A = \frac{\Delta s}{a} = K\sigma \tag{2.3}$$

where:

A = stress birefringence value,
Δs = optical path difference in nm,
a = sample thickness in cm,
K = stress optic coefficient in $\frac{\text{nm/cm}}{\text{kg/mm}^2}$, and
σ = mechanical stress in kg/mm^2.

The amount of stress birefringence in an optic can be minimized by the rate of speed for the annealing or thermal cooling. Stress birefringence may be seen in an optic by viewing the optic through a polariscope, two crossed polarizers, or a dark-field Schlieren system. The stress areas are the differences between the dark and light regions in the observed patterns.

As stress birefringence degrades the image quality in a standard optical system, the application determines how much stress birefringence is allowed. As stress birefringence changes the optical path in the orthogonal axes, controlling the impact of this error on polarization-based systems is critical. Table 2.1 provides examples of types of optical applications and the suggested tolerance for stress birefringence.

2.4.3 Bubbles and inclusions

The drawing symbol for bubbles and inclusions is `1/`. The quantifier after the `1/` is $\mathtt{N} \times \mathtt{A}$, where `N` is the number of bubbles or inclusions at the largest permissible size, and `A` is the largest permissible size bubble or inclusion. The size of the largest permissible bubble or inclusion is a grade system that follows a Renard series. The grades are nonarbitrary and are the square root of the area of the bubble or inclusion in mm.

Table 2.1 Guidance on allowed stress birefringence for example applications.[3]

Optical Path Difference Allowed	Application
<2 nm/cm	Polarization Instrument Interferometric Instrument
5 nm/cm	Precision Optics Astronomical Optics
10 nm/cm	Photographic Optics Microscope Optics
20 nm/cm	Magnifying Glasses Viewfinders
No Value (indicated with a dash)	Illumination Optics

The allowed grades are therefore 4, 2.5, 1.6, 1.0, 0.63, 0.04, 0.025, 0.016, etc. The smallest grade value listed for bubbles or inclusions in ISO 10110-18 is 0.006. Bubbles and inclusions at such a small size are generally only relevant for components in systems with a high laser-damage threshold specification or components near an image conjugate (field lenses and reticles). For most imaging systems, the smallest bubble or inclusion size typically used will be 0.040.

Accumulation Rule

Bubbles and inclusions in glass are typically found with a range of sizes. Therefore, specifying the maximum size bubble is not sufficient, and an analysis of the quantity of bubbles over a range of sizes is also necessary. There is an explicit accumulation rule in ISO 10110-18 that requires the sum of multiple bubbles of various sizes to meet an overall specification. The size of the bubbles that can be summed are from `0.16A` grade up to the maximum allowed grade `A`. The total cross-sectional area of all bubbles and inclusions greater than `0.16A` must be less than $N \times A^2$. An example of this accumulation rule is shown in Sec. 2.6.3 of this chapter.

Concentration

To minimize grouping of bubbles or inclusions a concentration rule is included in ISO 10110-18. The standard states that you cannot have `0.2N` of the maximum grade bubbles in a 5% subregion if the value of `N` is ≥10. As quantity specifications may be expressed with `N`<10, the standard states that you cannot have 2+ maximum allowed grade bubbles in a 5% subregion of the optic. Bubbles that are smaller than the maximum allowed grade are summed by area to an equivalent area of maximum size bubbles for purposes of determining a concentration.

2.4.4 Homogeneity and striae

The drawing symbol for homogeneity and striae is `2/`. Homogeneity and striae are different quantifiers and shall be considered independent of each other. The quantifier for these two tolerances uses the `2/` notation with `A;B`, where `A` is the homogeneity grade, and the `B` is the striae grade.

Homogeneity

As stated, the `A` quantifier after the `2/` drawing symbol is the homogeneity grade. This quantifier is represented by a class indicator of `NH` and numeric indicator. The indicator value is the maximum variation of the refractive index across the part in units of 10^{-6}. In practice, this means that a quantifier for homogeneity being `NH10` has a refractive index variation over the part of 10×10^{-6}.

The legacy indicator from past ISO 10110 parts for the homogeneity is still acceptable but not recommended. This indicator referred to a homogeneity grade shown as a symmetric tolerance, e.g., $\pm 10 \times 10^{-6}$. Since the actual value of the index is not known to the same order of magnitude, a symmetric tolerance makes no sense, and its interpretation can cause confusion.

Certain glass manufacturers may have their own notation used for homogeneity. As ISO 10110-18 tries to minimize confusion regarding the glass supplier, it is best to verify the fabricator's notation against the homogeneity value.

Striae

The `B` quantifier after the `2/` drawing symbol is the striae grade. Striae can be expressed in two different methods. The primary method of striae is expressed with a class indicator of `SW` followed by a numeric indicator. The indicator value is the maximum wavefront deviation over a 50 mm path length in nm. This indicator is broken down into four different categories that have a direct relationship to classic shadowgraph classes. This method of characterizing striae is widely accepted in the industry and traces its origins to the US military specification for optical glass, MIL-G-174.[16]

The legacy indicator from ISO 10110-3 for striae is accepted as well but not recommended. This system classified striae in five different classes relating to the percentage of test region obscured when the density of striae causes at least 30 nm of optical path difference. The legacy indicator was a class system that did not have a direct relationship to the value and its evaluation is not supported by most glass manufacturers. The change to the current striae indicator relates directly to the deviation length and is preferred.

2.4.5 Material tolerance notation on ISO 10110 drawings

In addition to specifying the birefringence, bubbles and inclusions, homogeneity, and striae for the finished component described above, it is also possible to specify the tolerances of the raw material or blank. The tolerances described in both ISO 10110-18 and ISO 12123 can be used. Even though there are many similarities, the quantifiers for raw materials shown in the two standards have differences. Each of these quantifiers are summarized in Table 2.2. The raw material quantifiers are listed by a preceding zero the drawing symbol, e.g., `01/` for the raw bubble and inclusion specification. We recommend using the ISO 12123 system for raw material rather than the ISO 10110-18 quantifiers to reduce confusion.

Using ISO 12123, the index tolerance is a coded notation with a quantifier of NP (for principal refractive index) preceding a code that is the value $\times 10^{-5}$. Assuming an ISO 10110 drawing lists the refractive index tolerances as ± 0.0005 the raw material could be specified as `NP050`. Additionally, the raw glass Abbe tolerance is a coded notation with a quantifier of `AN` (for Abbe number) preceding a code that is the value following the decimal in an Abbe tolerance. On an ISO 10110 drawing, the Abbe tolerance is listed solely as the numeric tolerance. Assume that the material has an Abbe tolerance as $\pm 0.5\%$. The raw optical material notation for this glass tolerance is defined by `AN5`.

Table 2.2 Drawing symbols and quantifiers for material properties. Notations in bold are recommended.

Tolerance	ISO 10110-18 Preceding Drawing Symbol	ISO 10110-18 Quantifier	ISO 12123 Quantifier
Index	None	None	`NP`**xx** **where the xx is ± xx × 10^{-5}**
Abbe Number	None	None	`AN`**x** **where the x is ± 0.x %**
Finished Component Stress Birefringence	`0/`	**`A`** **the max optical path difference in nm/cm**	n/a
Finished Component Bubbles and Inclusions	`1/`	**`N` × `A`** **`N` = the max number of bubbles** **`A` = the grade number of bubbles**	n/a
Finished Component Homogeneity	`2/`	**`NHxx`** **where the xx is xx × 10^{-6}**	n/a
Finished Component Striae	`2/`	**`SWxx`** **where the xx is ≤xx nm per 50 mm path length**	n/a
Raw Material Stress Birefringence	`00/`	`A` the max optical path difference in nm/cm	**`SBxx`** **where xx is ≤xx nm/cm**
Raw Material Bubbles and Inclusions	`01/`	`N` × `A` *N* = the max number of bubbles `A` = the grade number of bubbles	**`INxx`** **where xx is max number of bubbles per 100 cm^3** **`ICxx`** **where xx is the max cross-section of bubbles in a volume of mm^2 per 100 cm^3**
Raw Material Homogeneity	`02/`	`NHxx` where the xx is xx × 10^{-6}	**`NHxxx`** **where the xxx is xxx × 10^{-6}**
Raw Material Striae	`02/`	`SWxx` where the xx is ≤xx nm per 50 mm path length	**`SWxx`** **where the xx is ≤xx nm per 50 mm path length**

Stress birefringence tolerance in ISO 10110-18 is the numeric value for the maximum optical path difference in units of nm/cm. This specification is similar to that of ISO 12123 for raw optical glass. The stress birefringence specification for raw materials has a quantifier `SB` (for stress birefringence) and is followed by two numeric digits—the stress birefringence must be less than this value (reported in units of nm/cm. Since both raw optical materials and a completed optical component may be listed on an ISO 10110 drawing both of these two aforementioned specifications may be indicated. If the specification for the raw optical material is stress birefringence less than 12 nm/cm, the indication would be listed as `01/ SB12`. If this same specification is for the finished optical component, the indication would be `1/ 12`.

Specifying bubble and inclusion tolerances using ISO 10110-18 requires the quantity and maximum grade of bubbles or inclusions allowed in the optical component. This coded notation leaves this tolerance as a grade system that can be calculated for the area of each bubble or inclusion. In contrast, the ISO 12123 specification for bubbles and inclusions has independent preceding notations for both the number of bubbles or inclusions and their size. The grade notation for the number of bubbles is `IN` and the grade notation for the size of the bubbles is `IC`. If, for example, the bubble and inclusion indication for a finished component was `1/ 10` × `0.1`, the same specification for the raw material could also be written as `01/ IN010;IC10`.

Lastly, specifying homogeneity and striae is consistent between the finished optical component and the raw material. Both specifications use the same preceding code of `NH` and `SW` for homogeneity and striae, respectively.

2.5 Infrared versus Visible Materials

Infrared optical components may differ from visible-light-based systems because of the material properties required for the different wavebands.[17,18] Tolerances that might be considered loose for a visible optical component may be considered normal or tight for an infrared component.

There are certain additional specifications that should be considered when making a drawing for an infrared optical component. As many infrared materials do not transmit through the visible wavelength region, specifying the refractive index a n_e or n_d wavelength is impractical. Considering the various wavebands that infrared components may be used (short-wavelength infrared, long-wavelength infrared, etc.), specifying the reference wavelengths for the refractive index and dispersion calculation is crucial.[19]

2.6 Drawing Example

An example print is shown below in Figure 2.2 to clarify effective use of ISO 10110-18. Along with explaining the notation on this drawing, the specifications given can also be applied to the raw material as well.

2.6.1 Material property notation

The glass information on this print is listed as both the glass vendor and type (light blue). This print allows two different vendors to provide the substrate material. In addition to the glass vendors and types, the refractive index (purple) and Abbe number (green) are listed. Additionally, tolerances on the refractive index and Abbe number are listed as well. The tolerances for the refractive index and dispersion for the raw material could instead be given per ISO 12123 as `NP50` and `AN5`, respectively.

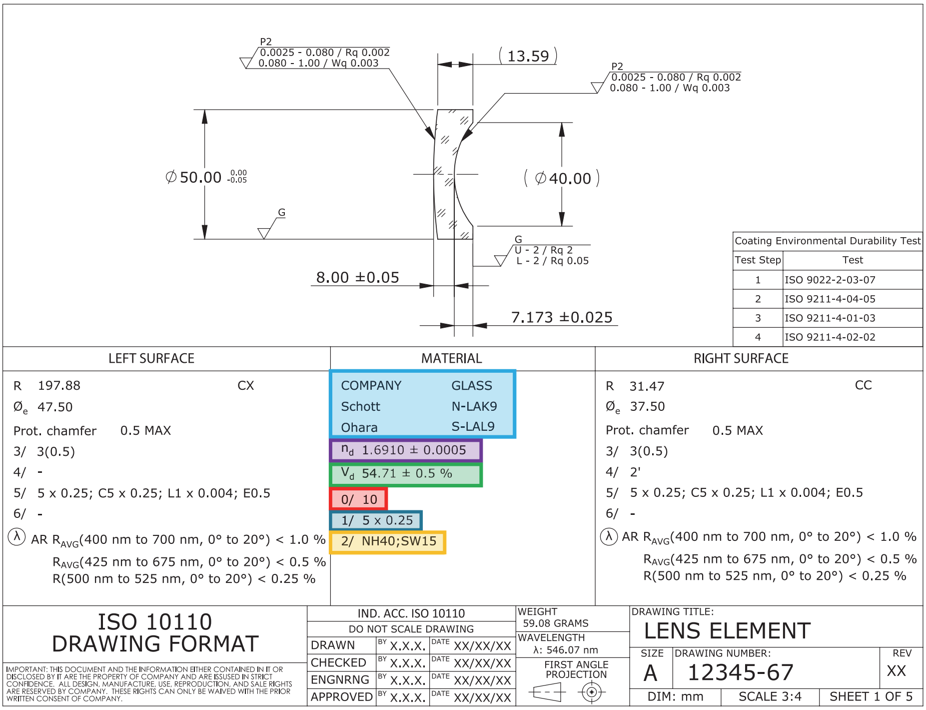


Figure 2.2 Example of lens ISO 10110 drawing, highlighting material information.

2.6.2 Stress birefringence

The stress birefringence for this example component is listed as a grade of 10. This means that the maximum amount of optical path difference is 10 nm/cm across the material. As listed in Table 2.1, a tolerance of 10 nm/cm would suggest that this component has the quality for a photographic or microscope system, but certainly can have other applications.

2.6.3 Bubbles and inclusions

The specification for bubbles and inclusions is listed as `1/5` $\times$ `0.25`. As a single value, this specification allows five bubbles or inclusions (`N`) of 0.25 grade (`A`). There may be bubbles that are smaller than the maximum permissible value (0.25). Based on the specification, the smallest bubble size that must be verified is greater than $0.16A$ or a bubble or inclusion grade of 0.04.

To provide an example of accumulation, we must look at the summation of areas ($N \times A^2$). Considering this, our nominal area is 0.0625 based on the drawing specification. Summing the number of those areas, we are allowed an area of 0.3125 (5×0.0625). If we solely look at a bubble and inclusion size of 0.16, we are allowed a quantity of 12:

$$Allowed = 0.3125 \geq 12 \times 0.16^2 = 0.3072 \tag{2.4}$$

Since it is not likely that a single bubble size will be found, a sample breakdown is as follows with three different bubble sizes: 0.25, 0.16, and 0.1:

$$Allowed = 0.3125 \geq (3 \times 0.25^2) + (3 \times 0.16^2) + (4 \times 0.1^2) = 0.3043 \quad (2.5)$$

2.6.4 Homogeneity and striae

This drawing example lists the homogeneity and striae as `NH40` and `SW15`, respectively. The preceding qualifier informs the manufacturer of the type of tolerance between the homogeneity or the striae and the quantifier value. The striae tolerance historically was known as a shadowgraph B tolerance for a frame of reference.

The second example drawing (Figure 2.3) differs from the first as this one is for an infrared component. The same notation is in place for both types of optical components.

In this example the material is listed without a single vendor. Unlike Figure 2.2, the specification for the refractive index is given at different wavelengths since this infrared component would be used in a different wavelength region than in Fig. 2.2. Since infrared optical systems are more sensitive to thermal changes than visible-light-based optical systems, the refractive index for this infrared material is also listed with the testing temperature values or range.

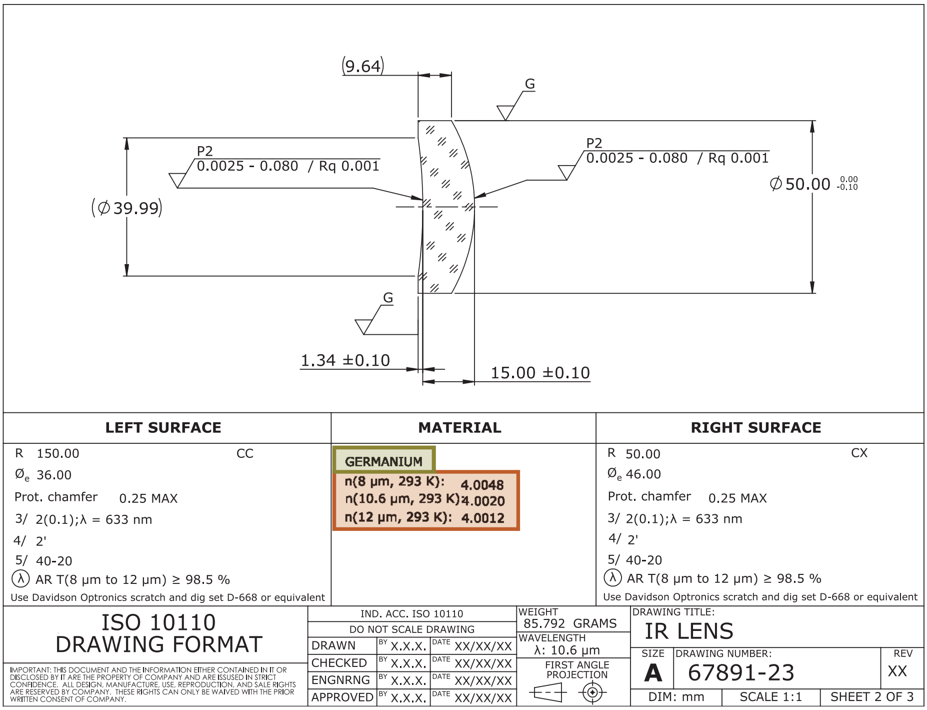


Figure 2.3 Example of infrared lens ISO 10110 drawing, highlighting material information.

References

1. P. Hartmann, *Optical Glass*, 1st ed., SPIE (Bellingham, Washington: 2014).
2. H. H. Karow, *Fabrication Methods for Precision Optics*, 1st ed., Wiley (Hoboken, New Jersey: 2004).
3. Schott Advanced Optics (SAO), *Stress in Optical Glass*, TIE-27 (Duryea, Pennsylvania: 2004).
4. SAO, *Bubbles and Inclusions in Optical Glass*, TIE-28 (Duryea, Pennsylvania: 2016).
5. SAO, *Homogeneity of Optical Glass*, TIE-26 (Duryea, Pennsylvania: 2016).
6. SAO, *Striae in Optical Glass*, TIE-25 (Duryea, Pennsylvania: 2006).
7. International Organization for Standardization (ISO), *Preparation of Drawings for Optical Elements and Systems—Part 18: Stress Birefringence, Bubbles and Inclusions, Homogeneity, and Striae*, ISO 10110-18 (Geneva: 2018).
8. A. Krisiloff, "Specification and tolerancing of bulk glass material imperfections with ISO standards," *Proc. SPIE* **10377**, 1037709 (2017) [doi: 10.1117/12.2276647].
9. ISO, *Specification of Raw Optical Glass*, ISO 12123 (Geneva: 2018).
10. P. Hartmann and R. Jedamzik, "Optical glasses and optical elements: comparison of specification standards ISO DIS 12123 and ISO 10110," *Proc. SPIE* **7102**, 71020L (2008) [doi: 10.1117/12.797568].
11. SAO, *Refractive Index and Dispersion*, TIE-29 (Duryea, Pennsylvania: 2016).
12. ISO, *Reference Wavelength*, ISO 7944 (Geneva: 1998).
13. A. Symmons and M. P. Schaub, *Field Guide to Molded Optics*, 1st ed., SPIE (Bellingham, Washington: 2016).
14. W. Vogel, *Glass Chemistry*, 2nd ed., Springer–Verlag (Berlin: 1994).
15. T. Izumitani, *Optical Glass*, 1st ed., American Institute of Physics (New York: 1986).
16. US Department of Defense, *Military Specification: Glass, Optical*, MIL-G-174B (Arlington, Virginia: 1986).
17. ISO, *Characterization of Optical Materials Used in the Infrared Spectral Range from 0.78 μm to 25 μm*, ISO 11382 (Geneva: 2010).
18. A. M. Phenis, "An international effort to standardize infrared material properties and their characterization techniques," *Proc. SPIE* **10590**, 105900I (2017) [doi: 10.1117/12.2309505].
19. American National Standards Institute, *Optical Elements and Assemblies—Infrared Spectral Bands*, OEOSC OP1.007-0000 (New York: 2019).

Chapter 3
Surface Figure and Form

Surface figure and form tolerances are low spatial frequency errors in an optical surface that directly contribute to the wavefront error on an optical surface. For surfaces sufficiently near the stop or a pupil conjugate, these errors are scaled and applied to the wavefront error of the system. Surfaces that are nearer than other surfaces to the object or an image conjugate will have the illuminated region affecting the wavefront. Such field lens and flattener types of optics can also have slope requirements depending on the required field flatness, distortion, and piston. These types of form errors are critical in defining the expected tolerances on an optical surface and may drive the tolerance budget for the entire optical system. Surface figure errors can be measured using numerous methods such as profilometry and interferometry. Data is assessed using different performance metrics such as power, irregularity, wavefront error [root mean square (RMS), peak-to-valley (PV), or robust peak-to-valley (PVr)]. Data can be analyzed in other various ways including decomposition into Zernike polynomials. In general, the measurement of precision optics' surface figure errors is often performed with an interferometer. Because there are multiple ways to describe the surface figure error for an optical surface there are many ways to describe the surface figure tolerances with ISO 10110.

3.1 Background

Much has been written about surface figure and how it can be measured.[1–3] For an ISO 10110 drawing, the notation for surface figure is specified in ISO 10110-5.[4] Along with measuring an optical surface, ISO 10110-14 is in place to specify an entire optical element or system figure wavefront deformation.[5] In conjunction with ISO 10110-5 and ISO 10110-14, the ISO 14999 series can be used to interpret the descriptions of each quantifier, as well as the methods for measuring said deformations.[6]

Since it is possible to measure the surface figure error in multiple ways, there are many different types of quantifiers available in ISO 10110-5.

A widely used method for measuring the surface figure (interferometry) uses the wave nature of light to interfere two light waves. There are numerous interferometers used for figure error assessment, using fringes formed by interfering a test beam from the surface or system under test with a reference beam. Optical interferometers for metrology are well-calibrated, a mature technology, and ubiquitous. Most of the notations discussed in ISO 10110-5 utilize the language of interferometric methods of measurement. However, using ISO 10110-5 does not restrict users to only interferometric testing.

3.2 Units

When specifying optical surface or optical system figure, there are multiple units that can be used to convey the same information. ISO 10110-5 and ISO 10110-14 both use similar units for classifying an optical surface. The recommended system of units is linear SI units. If the figure is characterized in linear SI units, the notation must use either nanometers (nm) or microns (μm). Many optical surfaces or systems, however, are described in terms of waves or fringes. As the characterization of waves may be confusing depending on the measurement method, waves are not allowed in ISO 10110-5 but the units of waves are the default notation when using ISO 10110-14. If using the units of fringes or waves, the wavelength must be explicitly specified. In previous versions of the ISO 10110 series there was a default wavelength, mercury e-line (546.07 nm), assumed for optical drawings. To avoid confusion the default wavelength is no longer allowed and the wavelength must be specified. The conversion between fringes and waves is as follows: 1.0 fringe is equivalent to 0.5 waves, as a double-pass interferometer is assumed. When dealing with surface slope, the units are angular: radians, degrees, arc minutes, or arc seconds.

3.3 Indications on Drawings

The drawing symbol for surface figure is `3/`. The quantifiers after the drawing symbol vary depending on which type of surface or system deviation is being used.

3.3.1 General quantifiers

Assuming the surface being analyzed is rotationally invariant, the basic form for figure quantifier is

`3/ A(B/C)`

where:
`A` = PV power of the optical surface or system,
`B` = PV irregularity of the optical surface or system, and
`C` = PV rotationally invariant irregularity of the optical surface or system.

Solely with the surface figure power term, an `A` that is a dash (–) does not indicate that a tolerance on the radius is completely ignored. By using the dash on the power term, the tolerance of the surface radius should be applied directly to the radius, listed in either the table field or on the drawing field. For all other tolerances following the `3/`, inclusion of the dash indicates that no tolerance is given. If no dash is given the non-toleranced data in ISO 10110-11 applies for the power and irregularity of the optical surface.

Use of a dash is critical when describing certain types of surface figure. If using any quantifier other than power and standard irregularity, the other quantifiers need not be included in the `3/` drawing symbol. Power and irregularity require a dash if there is no tolerance or if the power or irregularity quantifier is ignored. When not including the rotationally invariant irregularity a dash is not necessary and can be left off the drawing.

If the units are not explicitly specified for any of the quantifiers they are assumed to be in units of fringes. In this case, the wavelength must be listed either on the drawing or explicitly with the surface quantifiers. In that instance, the basic form for figure quantifiers becomes

`3/ A(B/C);λ=E`

where:
`A` = PV power of the optical surface or system,
`B` = PV irregularity of the optical surface or system,
`C` = PV rotationally invariant irregularity of the optical surface or system, and
`E` = wavelength used for the units of fringe spacings.

3.3.2 RMS quantifiers

Extending the amount of information conveyed, the general form for figure quantifiers is shown below. This quantifier form is intended if an RMS deviation from a surface is being specified

`3/ A(B/C) RMSx < D`

where:
`A` = PV power of the optical surface or system,
`B` = PV irregularity of the optical surface or system,
`C` = PV rotationally invariant irregularity of the optical surface or system,
`RMSx` = type of RMS deviation from the optical surface or system, and
`D` = RMS deviation from the optical surface or system.

The RMS deviation of the optical surface or system can be measured in multiple ways. There are three subscripts that can be used to describe the RMS wavefront (`t`, `i`, or `a`). If the RMS wavefront is being used as the quantifier, the subscript is necessary. When describing a surface or system figure, multiple types of RMS deviation are acceptable. These specifications must be broken up with a semicolon between them. The three subscripts when describing RMS deviation are

(1) `t` = total RMS deviation from the optical surface or optical system,

(2) `i` = RMS deviation from the optical surface or optical system with power removed, and

(3) `a` = asymmetric variant RMS surface irregularity.

- `RMSt`: The functional RMS surface deviation as measured by an interferometer. This is the deviation that is output from the interferometer with the tilt removed.
- `RMSi`: The RMS deviation with the tilt and power terms removed. This type of RMS deviation is the surface found from taking the functional RMS deviation (`RMSt`) and subtracting the power.
- `RMSa`: The asymmetric RMS deviation with the wavefront aspheric approximation, power, and tilt removed. When using the asymmetric RMS deviation the surface measured is typically aspheric.

If the optical surface being measured is cylindrical (rotationally variant), there are slight differences that would occur from the information provided above. For a basic form, the surface quantifiers become

`3/ AX:AY(B/CX:CY)`

As is with the rotationally invariant surfaces, the `AX` and `AY` terms are the surface or system power. In this case, the two power terms are for the orthogonal axes. The irregularity for a cylindrical surface is still a single term, except it is used for the irregularity for cylindrical surfaces or systems.

Like surface or system power, the rotationally invariant irregularity can have orthogonal terms for both the x and y orientations.

3.3.3 Peak-to-valley quantifiers

The peak-to-valley (PV) deviation can be specified instead of the RMS deviation for a surface or system by using a separate system than the power and irregularity. The quantifier for specifying the `PV` deviation is either

`3/ PV(Q)` or `3/ PVr(R)`

where:
`Q` = maximum PV surface deviation, and
`R` = maximum robust PV surface deviation.

The PV surface deviation is a simple explanation of a description. This type of surface deviation, similar to the total RMS deviation from a surface (`RMSt`), takes the direct output from an interferometer and compares that value. Another more recent method of defining the PV is the robust peak-to-valley (`PVr`).[7,8] The `PVr` is a peak-to-valley specification that was proposed by Dr. Chris Evans in 2008 and has been shown to minimize measurement error between interferometers, and to be more reproducible than other means of defining PV. Instead of taking the direct output from the interferometer and using that as the specification, the PVr takes the PV of the interferogram when fit to a 36-term Zernike polynomial (a 10th-order Zernike polynomial, plus the 12th-order spherical term) and adds 3× the RMS deviation from the residual. The PVr specification is a more-robust method of specifying PV than other means of doing so as it takes differences in the measurement equipment or setup out of the potential errors.

3.3.4 Slope quantifiers

If the surface deviation is being measured as a slope, rather than the general interferometric form, a different notation is used. The slope of a surface can be specified as either the maximum or RMS slope deviation. Both the maximum surface slope and the RMS surface slope have similar quantifiers. The overall form for the slope is shown below. The terms listed following the quantifiers are similar enough between the maximum slope and the RMS slope that they are listed together.

3/ ΔSv,w(F/G/H) or 3/ RMSΔSv,w(K/L/M)

where:
ΔS = maximum slope deviation,
RMSΔS = RMS slope deviation,
v = number of dimensions considered for slope (1-dim, 2-dim, or nothing),
w = orientation of slope in either cartesian coordinates (x / y) or polar coordinates (ρ / φ),
F or K = maximum slope deviation value or RMS slope deviation value,
G or L = sampling length for one dimension or the edge length of the sampling area when considering two dimensions, and
H or M = spatial sampling interval.

Measurement of the slope deviation can be performed in multiple directions. Consequently, the slope quantifier has the v term designating the number of dimensions that should be considered. Noting the number of dimensions for a slope tolerance should be performed as either 1-dim or 2-dim depending on the number of dimensions being considered. If the tolerance has no dimensions that are being considered, the v term need not be listed. Along with the number of dimensions, the coordinate system that is being used for the dimensions themselves can be listed. This listing is done in either Cartesian (x / y) or polar (ρ / φ). If no coordinate system must be used (i.e., the specification directions on the surface are arbitrary), the w term need not be considered.

The actual slope deviation itself is listed as either the F or K quantifier, depending upon the type of slope tolerance (maximum or RMS). When measuring the slope quantifier, the sampling length must be defined to calculate the slope deviation. Note that the quantifiers `G` and `L` can both indicate either sampling length or sampling area. The difference between the quantifiers comes with the usage for slope evaluated in one or two dimensions. Over the sampling distance, the interval or spacing for the slope deviation must be listed. The sampling length is the spacing between different measurement points. The quantifiers for sampling interval are either `H` or `M` depending on which type of slope is being toleranced, either maximum slope or RMS slope, respectively.

3.3.5 Zernike coefficient quantifiers

In certain cases, it is useful to tolerance a surface based on the amplitude of specific terms or groups of terms in a Zernike polynomial fit of the form error. In this situation there are two different methods for specifying the amplitudes of the Zernike terms. The terms may be taken individually, where the coefficient in each term is toleranced, or they may be evaluated as a group.

When defining the individual Zernike terms, their quantifier is:

```
3/ Z(n,m)(PV < O;RMS < P)
```

where:

`Z(n,m)` = index of the Zernike term to evaluate,

`O` = the peak-to-valley (PV) deviation described by the indexed Zernike term, and

`P` = RMS deviation described by the indexed Zernike term.

Note that the Zernike terms are not listed in a *one-to-n* index as is often seen in optical design or metrology software. Because there is no consensus on a single indexing system (Zernike/Born and Wolf, OSA/ANSI, University of Arizona/Fringe, Wyant, Noll, etc.) such a system lends itself to misunderstanding. The index for the Zernike terms must be indexed by both and integers.[9] The purpose of using two index terms is to minimize confusion with the orientation for each Zernike term. A good description of the index system for the ISO 10110-5 series compared to other Zernike notations was written by Schwiegerling.[10] The standard terms to be used are given in ISO 14999-4, Annex B.[11] The ordering in that table is most similar to a fringe Zernike description. The Zernike surface or system deviation error can be listed as either a PV or an RMS value.

It is often advantageous to group terms and specify the magnitude of the resultant fit to the surface form error. Here, they can be grouped as pairs of indices or by order. The quantifiers for groups of pairs are:

$$3/Z(n_1,m_1),Z(n_2,m_2),\ldots Z(n_i,m_i)(PV < O;RMS < P)$$

where:
$Z(n_1,m_1)\ldots Z(n_i,m_i)$ = Zernike terms to evaluate,
O = PV deviation described by the group of Zernike terms, and
P = RMS deviation described by the group of Zernike terms.

For example, it may be useful to specify the sum of the power, spherical, and fifth-order spherical terms together. Such a specification would be, for example

$$3/\ Z(2,0),Z(4,0),Z(6,0)(RMS < 100\ nm)$$

It is also common to group terms by order and to provide a tolerance for the PV or RMS of that order or group of orders. The quantifiers for orders of Zernike terms are:

$$3/\ Z(N=x)(PV < O;RMS < P)$$

where:
Z(N=x) = Zernike terms of order x,
O = PV deviation of the fit to the terms of order x, and
P = RMS deviation of the fit to the terms of order x.

As another example, consider specifying the RMS of the 10th order fit to the surface form error. Such a specification would be:

$$3/\ Z(N \le 10)(RMS < 100\ nm)$$

It is also possible to combine these systems. For example, to remove piston, tilt, and power from the fit, and add the 12thorder spherical term, you can specify:

$$3/\ Z(2 < N \le 10),Z(12,0)(RMS < 100\ nm)$$

If there are many Zernike terms or groups of terms that are being listed, it may be helpful to use a table format instead of the surface form indicator. When using a table for the Zernike terms, the quantifier should look like Table 3.1 for individual terms in peak-to-valley. It is also possible to list individual terms for the RMS deviation in a table, but that is more typically achieved by specifying a group of terms together. The units in the table format must be specified. It is suggested to use units of nanometers (nm).

Both indications on a drawing for Zernike quantifiers convey the same result. The positive aspect of using the table format to describe surface

Table 3.1 Surface deviation quantifier for Zernike terms.

Z	PV
Z(n,m)	O

notation with Zernike terms is the additional space it provides on a drawing.

3.3.6 Table notation

Rather than describe the surface form tolerance across the entire surface, it is possible to list the surface deviations in a table rather than in the surface form indicator, `3/`. When describing the surface form deviation with a table, the deviation quantifier value is listed in a separate column of the reference table. Either a Cartesian or polar coordinate system is necessary because it is possible to have deviation tolerances at varying locations across the surface. This indication lists where on the surface a slope or position deviation occurs as a point cloud of data for the surface figure.

3.4 Surface Figure

ISO 14999-4 is often used in conjunction with ISO 10110-5 to assist in describing the surface figure and provide a description of measurement techniques.

Historically, a test plate was used to measure the power and irregularity of an optical surface interferometrically. In this case an optical surface with the near perfect radius was manufactured of the negative curvature, known as a test plate. The optical surface to be measured would then be placed against this surface and viewed under a nearly monochromatic light source. Under these viewing conditions the difference in surface form could be seen with Newton's rings, or fringes. The power measurement is the number of fringes found and the irregularity is the difference in number of fringes between orientations (vertical versus horizontal). An example of how a test plate would be used is shown in Fig. 3.1. In both orientations shown in Fig. 3.1, the difference in curvature from the ideal radius, or test plate, is the surface power. The difference in curvature between the two orientations of this example (*xz* and *yz*) is the irregularity of the optical surface.

Describing surface figure in its most basic form uses the general notation described in Sec. 3.3.1, where the power and the irregularity of the surface are

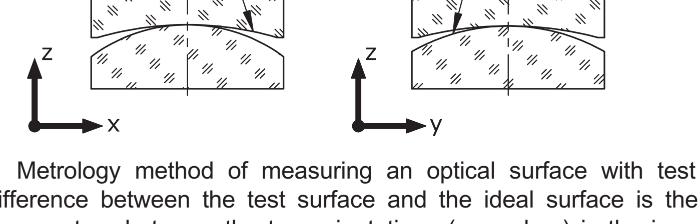
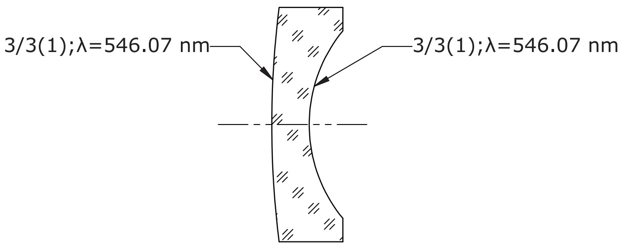


Figure 3.1 Metrology method of measuring an optical surface with test plates. The curvature difference between the test surface and the ideal surface is the power. The difference in curvature between the two orientations (*xz* and *yz*) is the irregularity of the optical surface.



(a)

(b)

Figure 3.2 Example optical element (a) showing the definition for surface figure on left and right optical surfaces with (b) the sample interferogram found with this type of surface figure measurement.[12]

the quantifiers. In the case of Fig. 3.2(a), the surface figure errors are described in terms of fringes. As such, the wavelength for measuring the figure is provided as well.

An example interferogram that a surface with this figure error might produce is shown in Fig. 3.2(b). In Fig. 3.2(b) you can count the number of fringes for the power and the difference between the orientations for the irregularity. A surface that produced this interferogram would barely pass the specification given in Fig. 3.2(a). Note this interferogram would indicate a perfect surface if it is null.

With the wide availability of computer-controlled interferometers, specifying surface errors with an RMS error method (that includes figure) is common. Measuring the optical surface for a random error does not indicate that the irregularity is uniform and easy to determine by simple visual inspection. Using a computer-controlled interferometer, the resulting measurements incorporate the surface figure errors in terms of RMS surface error and can be analyzed for different spatial frequencies.

Regarding RMS surface error, there are three different ways to specify RMS surface figure error. As shown in Fig. 3.3, the surfaces are specified with 3 fringes of power and 0.1 fringes of RMS surface figure error. In the case of

Fig. 3.3, the notation `RMSi` surface figure error is meant to exclude the power measurement.

Figure 3.4 shows the type of surface error that might be controlled with an `RMSi` surface error notation rather than a simple astigmatic irregularity. Figure 3.4(a) shows the surface figure error including the power of the optical surface. Note that with the power present it is difficult to determine the RMS surface error. Figure 3.4(b) shows this same surface figure error with the

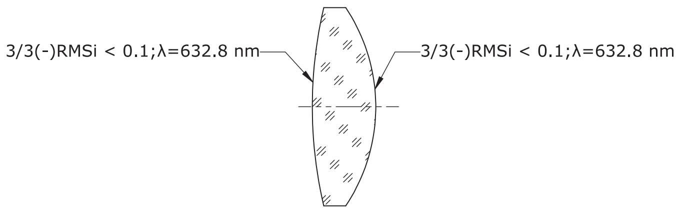


Figure 3.3 Example optical element showing the definition for surface figure of power and RMS irregularity error.

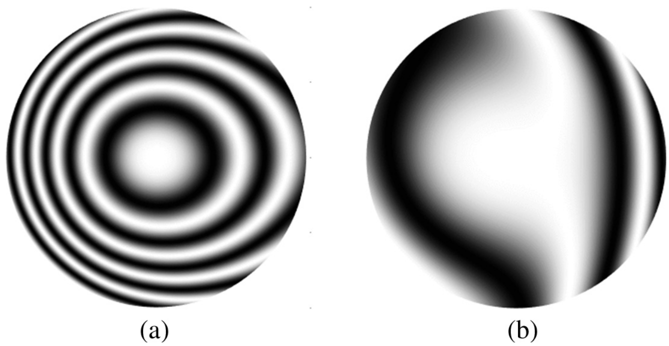


Figure 3.4 Example interferograms showing surface figure (a) with power and (b) without power.[12]

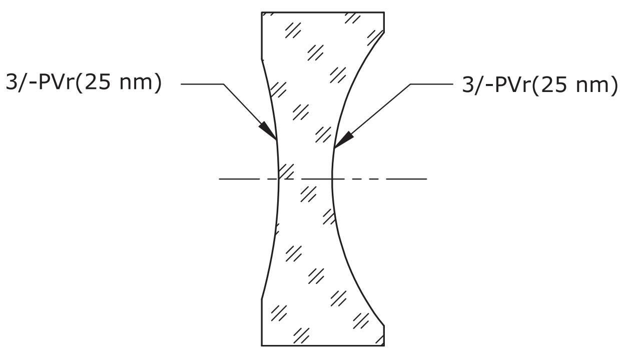


Figure 3.5 Example optical element showing the definition for surface figure of robust peak-to-valley error.

power removed. Finally note that these types of fringes are formed to null error versus methods to view error intentionally with tilt fringes.

There are two other ways to specify a surface for the total peak-to-valley error in ISO 10110-5. Section 3.3.3 describes these two types of peak-to-valley measurement methods. Figure 3.5 shows an optical element with both surfaces using the robust peak-to-valley method to define surface figure. The two surfaces in Fig. 3.5 have a specification of 25 nm. This notation specifies the total surface figure including power in the surface as designated with the preceding dash to the `PVr` quantifier.

3.5 Zernike

Unlike the previous types of surface figure specifications, a surface specified with Zernike terms can be much more complex. In Fig. 3.6, the left surface is specified as a standard peak-to-valley surface less than 30 nm. For the right surface, the surface figure is targeted at a peak-to-valley surface error for each of the Zernike terms `Z(2,0)`, `Z(3,3)`, and `Z(3,-3)`.

An interferogram of a surface that would barely pass the specification of the right surface described in Fig. 3.6 is shown below in Fig. 3.7. As noted in

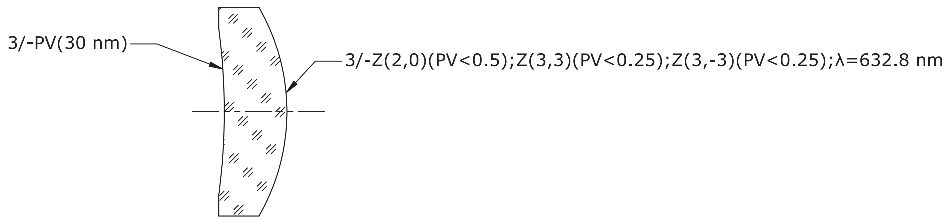


Figure 3.6 Example optical element showing the left optical surface with peak-to-valley error and the right optical surface with PV errors for individual Zernike terms.

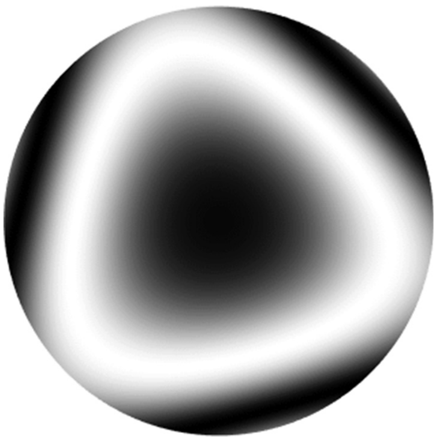


Figure 3.7 Example interferogram for Zernike terms `Z(2,0)`, `Z(3,3)`, and `Z(3,-3)`.[12]

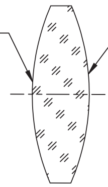


Figure 3.8 Example optical element showing both surfaces with peak-to-valley error based on a combination of Zernike terms.

Sec. 3.3.5, this surface figure is specified by three individual Zernike terms. Term `Z(2,0)` is the power of the optical surface, which appears to dominate the amount of surface error. Terms `Z(3,3)` and `Z(3,-3)` are trefoil in perpendicular axes. The effect from the trefoil is clearly visible with the nodes at the edges of the interferogram.

When specifying a surface with Zernikes, the individual Zernike terms need not be targeted one at a time. It is possible to specify a surface with a combination of Zernikes. Figure 3.8 shows a surface specified with the combination of three different Zernike terms that must sum to be less than a peak-to-valley of 40 nm.

3.6 Component or System Wavefront

The figure error for an entire optical element or system is described by the transmitted system wavefront deviation. When describing an entire component or system wavefront, the drawing symbol is a `1` preceding the standard `3/` for error; i.e., `13/`.[13] The same quantifiers that are used for surface figure error may also be used to describe the transmitted wavefront error—except that the default notation for power, irregularity, and rotationally varying irregularity is waves, rather than fringes. Note that it is not necessary to have a tolerance on the transmitted wavefront error. In the case that both surface figure errors and system wavefront errors are specified on a drawing, both specifications shall be satisfied.

The wavefront error must be shown in the drawing field where the notation and quantifier are directed at the optical axis. The pupil or stop location must be specified on the optical axis along with the wavefront error. The aperture location is necessary for fabrication, specifically to notify the manufacturer how to align and test the component or system. Figure 3.9 shows an example of an optical element where the wavefront error is specified with RMS wavefront error with the power removed, signified `RMSi`.

If the specified system wavefront error is not for infinite conjugate space, the object and/or image position(s) must be noted on the optical system

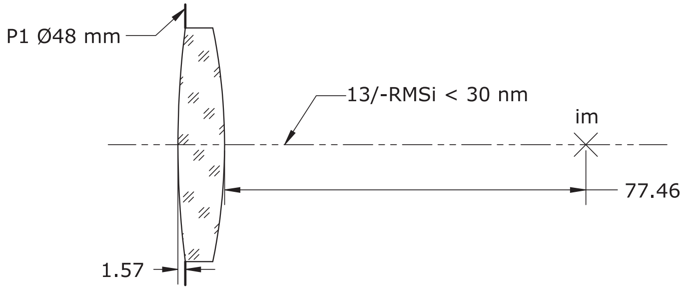


Figure 3.9 Example of notation used for system wavefront error for an optical element. The entrance pupil location and size are shown. The transmitted wavefront is shown on the optical axis.

drawing field. The only time that power is considered for a system wavefront error is when either the object or image position is listed on the drawing.

In Fig. 3.10, the drawing assumes collimated light is entering the optical system and focusing at the image position indicated. The entrance pupil location and diameter are shown relative to the vertex of the front optical surface. The image position is specified relative to the rear vertex of the second optical element.

Considerations must be made for how the optical element or system wavefront error will be measured. System wavefront error specified on a drawing is intended for a single-pass configuration of an interferometer. If the

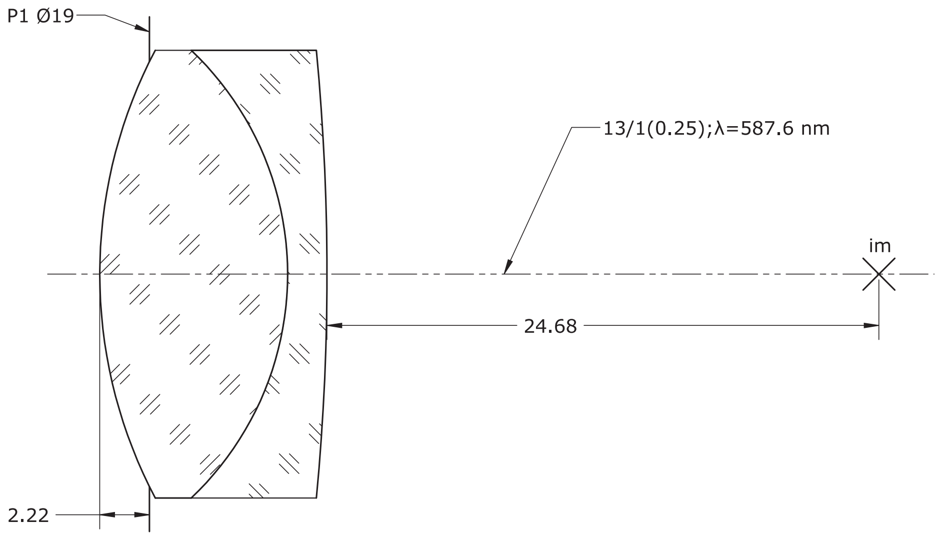


Figure 3.10 Example of notation used for system wavefront error for an optical element with both power and irregularity.

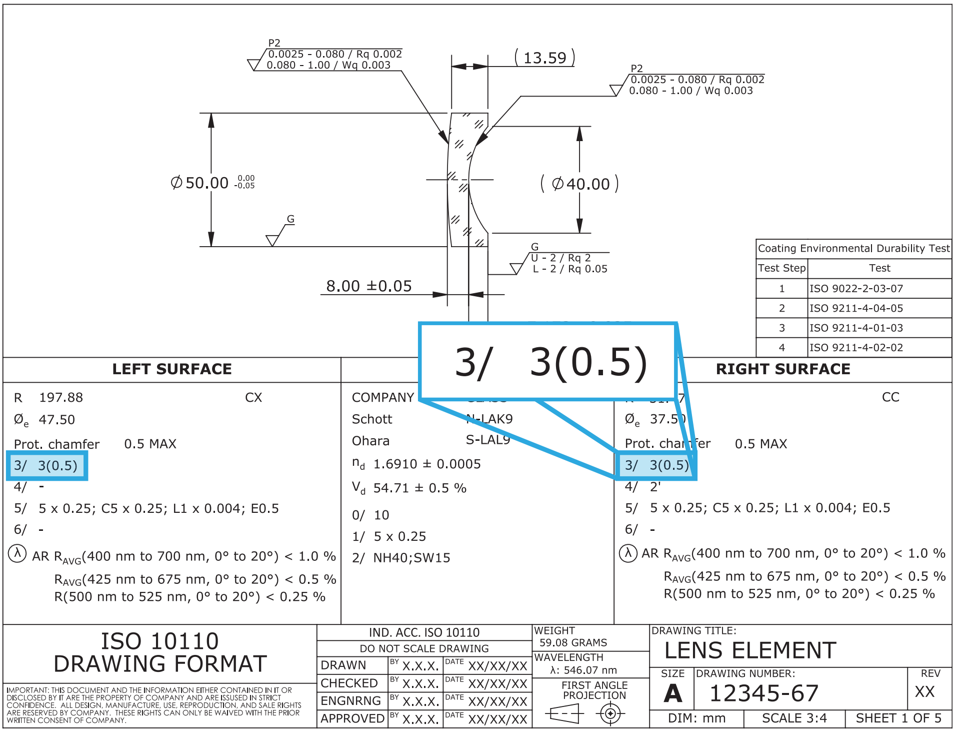


Figure 3.11 Example drawing with spherical surface figure of power and irregularity on both the left and right surfaces.

testing that is performed is with a double-pass configuration, considerations must be made to calculate the system wavefront error.

3.7 Drawing Example

Example drawings highlighting surface figure are shown in Figs. 3.11 and 3.12. The first drawing, Fig. 3.11, highlights the basic form notation for figure tolerance of power and irregularity on both surfaces. The surfaces both have a power of 3 fringes and an irregularity of 0.5 fringe. Even though it is not specified in the surface figure notation (3/), the wavelength is specified on the drawing sheet to be 546.07 nm.

The second example, Fig. 3.12, shows an aspheric optical element. For the spherical surface (the left surface), the figure is specified using a typical power and irregularity notation. On the aspheric surface (the right surface), the power is specified in terms of fringes. The irregularity is specified as an RMS irregularity of the surface rather than the typical irregularity.

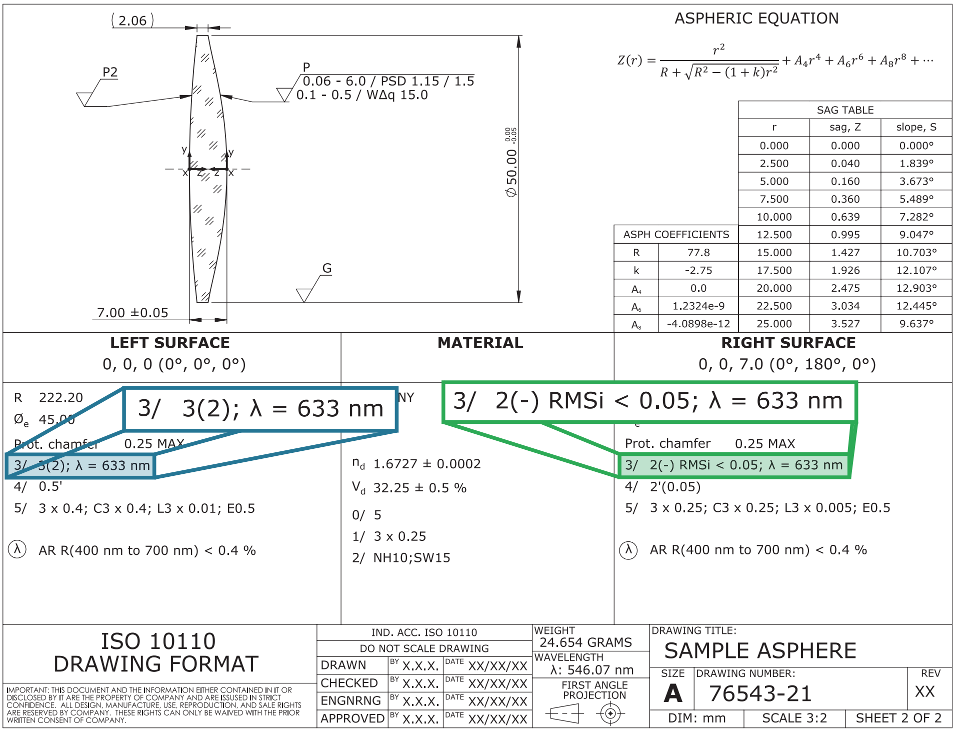


Figure 3.12 Example drawing with a spherical surface on the left surface having power and irregularity,and an aspheric surface on the right surface having power and `RMSi` irregularity tolerance; along with a sag table.

References

1. D. Malacara, *Optical Shop Testing*, 3rd ed., Wiley (Hoboken, New Jersey: 2007).
2. W. J. Smith, *Modern Optical Engineering*, 3rd ed., McGraw–Hill (New York: 2000).
3. E. P. Goodwin and J. C. Wyant, *Field Guide to Interferometric Optical Testing*, SPIE (Bellingham, Washington: 2006).
4. International Organization for Standardization (ISO), *Preparation of Drawings for Optical Elements and Systems—Part 5: Surface Form Tolerances*, ISO 10110-5 (Geneva: 2015).
5. ISO, *Preparation of Drawings for Optical Elements and Systems—Part 14: Wavefront Deformation Tolerance*, ISO 10110-14 (Geneva: 2018).
6. ISO, *Optics and Photonics—Interferometric Measurement of Optical Elements and Optical Systems*, ISO 14999 (Geneva: 2019).
7. C. J. Evans, “Robust estimation of PV for optical surface specification and testing,” *Proc. Opt. Soc. Am. Front. Opt.*, 2008, OWA4 (2008).

8. C. J. Evans, "PVr—A robust amplitude parameter for optical surface specification," *Opt. Eng.* **48**(4), 043605 (2009).
9. ISO, *Interferometric Measurement of Optical Elements and Optical Systems—Part 2: Measurement and Evaluation Techniques*, ISO 14999-2 (Geneva: 2019).
10. J. Schwiegerling, "Review of Zernike polynomials and their use in describing the impact of misalignment in optical systems," *Proc. SPIE* **10377**, 103770D (2017) [doi: 10.1117/12.2275378].
11. ISO, *Interferometric Measurement of Optical Elements and Optical Systems—Part 4: Interpretation and Evaluation of Tolerances Specified in ISO 10110*, ISO 14999–4 (Geneva: 2015).
12. J. C. Wyant, "Using webMathematica to solve optics problems," *accessed September* 16, 2021, https://wp.optics.arizona.edu/jcwyant/mathematica/webmathematica/.
13. ISO, *Preparation of Drawings for Optical Elements and Systems—Part 1: General*, ISO 10110-1 (Geneva: 2019).

Chapter 4
Surface Texture: Roughness and Waviness

Surface roughness tolerances are requirements for the level of smoothness of the polish of an optical surface, typically expressed as a maximum allowed RMS surface height error and evaluated with a profilometer. The advent of highly deterministic small-tool polishing has led to increased importance of spatial frequencies lower than traditional roughness scale-lengths but higher than those more easily controlled using surface form tolerances described in Chapter 3. To cover this need, the standard includes notation to control such mid-spatial frequency errors with more rigor.

ISO 10110-8 provides the notations for a variety of tolerancing methods of varying complexity and sophistication.[1] The simplest form of surface texture tolerance is a simple polishing grade based on a standard level of RMS roughness. If the standard polish grades are not suitable, any level of RMS roughness tolerance can be defined. Moreover, the scale-lengths over which the RMS roughness requirement is to be evaluated can be customized. This same tolerance methodology can be used to specify deliberately rough surfaces, such as the ground edge of a lens, using both upper and lower bounds on the roughness. In addition to these fairly conventional notations, the standard provides for more complex tolerances often used to control mid-spatial frequency surface form errors, or waviness, based on peak slope, RMS slope, and even a power spectral density (PSD). The most recent edition of the standard adds the areal (2D) versions of these notations.

4.1 Background

Tolerances on surface texture are not unique to the optics industry. Indeed, specifying the smoothness of a bearing surface is quite old. Today, the study of friction, wear, and lubrication is referred to as tribology; and is greatly concerned with characterizing the texture of a surface. Leonardo da Vinci (1452–1519) is considered by some to be the first tribologist.[2] His notebooks include many observations about friction and indicate that he was very concerned about how friction limited efficiency in his machines. He also observed that smoother materials resulted in less friction. His notes are often credited as the principal influence on the "laws of friction" published by French physicist Guillaume Amontons in 1699.

Quantification of the smoothness of a finished surface became of great interest in the 1930s, when a Bentley engine (made with famously smooth cylinders) seized up during the Le Mans 24-hour race. Not long after that, E. J. Abbott invented what is thought to be the first analog profilometer to be used to characterize smooth surfaces and assign a quantitative value for surface texture. Many of the notations in use to specify surface texture on optics today trace their origins to profilometry and Abbott's work to develop the bearing area curve, also known as the Abbott–Firestone Curve.[3]

4.1.1 The language of roughness

Surface texture, and particularly surface roughness, is usually discussed using the language of surface profilometry. In many cases this is appropriate because the surface texture tolerances are evaluated using one or more measurements with a tactile or optical profiler. However, in the optics industry most of us do not have a background in tribology and the terminology can be confusing. Several key profilometry terms and definitions are given in Table 4.1.

4.1.2 Form, waviness, and roughness

We often think of surface form and surface roughness as very different requirements. Surface form is specified in units of waves and fringes—with terms such as power, irregularity, and Zernike profiles—and evaluated interferometrically (be it a test plate or a Fizeau interferometer). In contrast,

Table 4.1 Key profilometry terms used with surface texture specifications.

Profilometry term	Definition
Total profile	Representation of a surface obtained by a profiler
Primary profile	Total profile after applying a long-pass filter to remove noise
Spatial wavelength	Peak-to-peak distance of a surface undulation, especially when decomposed using Fourier analysis
Spatial band	Range of spatial wavelengths to be considered in a specification
Surface form	Polynomial fit of a measured profile over long scale-lengths
Waviness profile	Primary profile over the waviness spatial band
Roughness profile	Primary profile over the roughness spatial band
Evaluation length	Length of the profile to be evaluated
Sampling length	Longest spatial scale length to be considered for an evaluation. The evaluation length of the primary profile is the sampling length
Lay	Direction of the profile to be evaluated with respect to the surface. For example, on a round part, *R* indicates a radial lay, while *C* indicates a circumferential lay

surface roughness is specified in Angstroms (Å) or nanometers (nm)—with terms such as band-limited RMS or PSD—and evaluated with a surface profiler. Surface form errors affect image quality,[4] while surface roughness affects transmission and in some cases scatter.[5,6] However, surface form and surface roughness are just errors in the shape of an optically smooth surface evaluated over different scale-lengths, as shown graphically in Fig. 4.1. Surface errors that are measured over the entire surface are considered form errors while surface errors measured over very short distances are considered roughness. The spatial scale lengths between these two extremes are referred to as mid-spatial frequencies or waviness. Seen this way, surface form, surface waviness, and surface roughness are just different parts of the continuum of surface errors, and differentiation between them is arbitrary. Diverse applications will have different cutoffs between the three types of surface errors, and specifications are driven more by the impact on the performance and the measurement method, rather than the physical scale of the error.

4.1.3 Origin of roughness symbols in ISO 10110

A historical standard symbol for surface texture is shown in Fig. 4.2. Figure 4.2 comes from MIL-STD-10 (published in 1949), but has been converted into SI units.[7] The symbol likely predates that standard; however, drawing symbols for mechanical features in the MIL-Specs system can be traced to Ordnance Corps ORD 30-1-7—which first appeared in 1946—and surface texture specifications go back to 1940. MIL-STD-10A[8] was ahead of

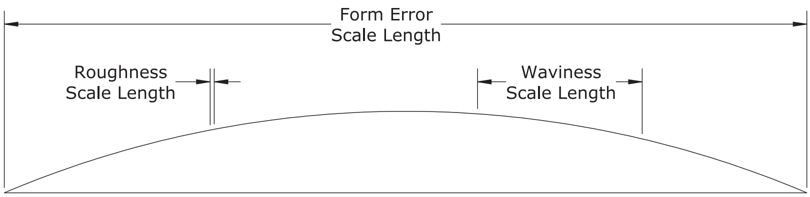


Figure 4.1 Surface scale lengths for form, waviness, and roughness.

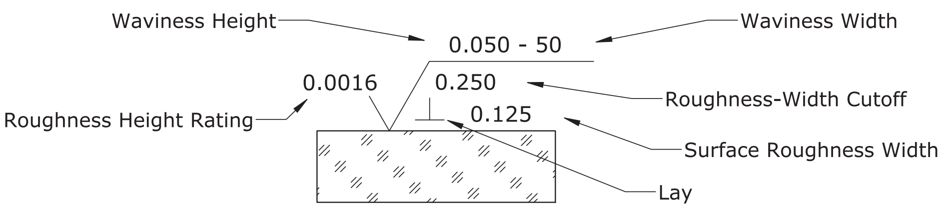


Figure 4.2 Surface texture symbol from MIL-STD-10A, a variant of which is still in use today. Values have been translated from inches to mm. The origin of the symbol likely precedes this standard.[8]

its time; it specified the use of average roughness (Ra) rather than the RMS roughness (Rq) used in ISO 10110, but had many excellent features. It provided a notation for waviness in addition to roughness, used spatial bands (e.g., from “roughness cutoff” to “roughness width”) to clarify the meaning of the texture specification, included a symbology for lay, and provided default scale-lengths depending on the allowed magnitude of the roughness.

In 1963, MIL-STD-10A was canceled and replaced by ASA B46.1-1962, the American Standards Association equivalent. That standard continued the same symbology and notation as that used in MIL-STD-10A. Since then it has been revised seven times, most recently in 2019 and is active today as ANSI/ASME B46.1-2019.[9] Unfortunately, during the many revisions the symbology and drawing notation was split off to ASME 14.36, and the default spatial bands were lost entirely. It is important to note that the committee that manages B46.1 also interacts with the US delegation in ISO/TC 213, which is responsible for the ISO surface texture standards in the geometrical product specifications (GPS) system—including ISO 1302.[10] The committees try to maintain harmony between the US standards and the ISO standards for defining and evaluating surface roughness on surfaces. For optical surfaces, ISO 10110 uses a notation similar to that of the GPS standard system as given in ISO 1302, which can be traced back to the original symbology from the MIL standards in the 1940s. It is worth noting that ISO 1302 is slated to be replaced by the ISO 21920 series in the first half of the 2020s.

Most of the initial drafts of parts of ISO 10110 were based on the German DIN standard in the 1990s. One of the exceptions is ISO 10110-8, which drew both on the roughness part of DIN 3140 and its qualitative surface polish grades, and the band-limited notations from ASME 14.36 and ISO 1302. As a result, there have always been two somewhat complimentary notations used in ISO 10110-8.

4.2 Indications on Drawings

The drawing symbols for surface texture in ISO 10110 are shown in Fig. 4.3. Figure 4.3(a) shows the symbol for a surface made using a material removal

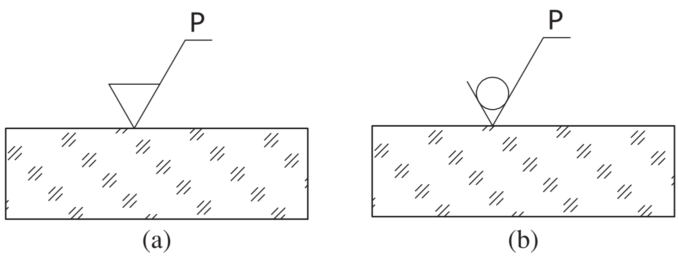


Figure 4.3 Surface texture symbols from ISO 10110-8 for (a) polished surfaces and (b) molded surfaces.

process such as grinding and polishing, while Fig. 4.3(b) shows the symbol for a molded surface.

A ground surface also uses the symbol shown in Fig. 4.3(a). Above the symbol is either `G` or `P`, or a polish grade. No additional quantifiers are necessary; however, if additional texture requirements are needed, they are placed below the horizontal line in the symbol—as shown in the example in Fig. 4.4.

4.2.1 Polish grades

The simplest form of a surface texture specification is a polish grade. An example of a surface specified to polish grade 3 (`P3`) is shown in Fig. 4.5. Historically, polish grades corresponded to qualitative polish samples and were loosely quantified based on a number of micro-defects per 10 mm trace, where each defect was counted if it "differed substantially" from the baseline. In 2010 the standards committee voted to add an informative Annex to the standard, correlating the polish grades to a surface roughness amplitude. In the 2018 edition, the references to micro-defects were removed and the surface roughness grades from the Annex were made normative. A correlation between the polish grades and the corresponding surface roughness is given in Table 4.2.

It is important to note that the spatial bands for the surface roughness of the standard polish grades (0.002 mm to 1.000 mm) are not the same as either of the default spatial bands for surface roughness (0.0025 mm to 0.080 mm) or

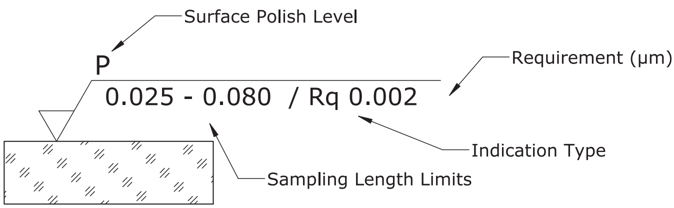


Figure 4.4 Surface texture with additional quantifiers.

Table 4.2 Surface roughness values corresponding to the four polish codes.

Polish Grade	Surface roughness (RMS) for scale lengths from 2 μm to 1000 μm
P1	≤8 nm
P2	≤4 nm
P3	≤2 nm
P4	≤1 nm

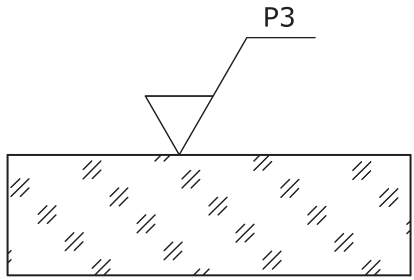


Figure 4.5 Surface texture with polish grade.

surface waviness (0.080 mm to 2.500 mm). This differentiation is discussed more in Sec. 4.3 below.

4.2.2 Band-limited RMS roughness and waviness notations

In some applications, the surface roughness given by the standard polish grades is not useful and a more complex roughness specification is required. The most fundamental of these is the band-limited RMS. If the surface roughness is "well-behaved" or the height errors are statistically well-distributed across the surface, the RMS value `Rq` can be related to the magnitude of the scattered light.[11]

When specifying the surface texture using a band-limited RMS roughness, the specification is placed under the horizontal line in the symbol as was shown in Fig. 4.4, and takes the form

`A – B / Rq C`

where:
`A` = short scale-length limit of the spatial band in mm,
`B` = long scale-length limit of the spatial band in mm, and
`C` = RMS roughness tolerance in μm.

This same notation can be used for ground or polished surfaces; the only difference is whether there is a `P` or a `G` above the horizontal line. Furthermore, the exact same format is used for a waviness specification, but the `Rq` in the roughness specification is replaced with `Wq`. Multiple roughness and waviness specifications over different spatial bands can be placed in a column beneath the horizontal line as needed.

For example, in Fig. 4.6, the surface texture specification is given for surface roughness and waviness together. The surface roughness shall be less than 0.002 μm RMS, when evaluated over spatial scale-lengths from 2.5 μm to 80 μm. The surface waviness shall be less than 0.003 μm when evaluated over spatial scale-lengths from 0.080 mm to 1.00 mm.

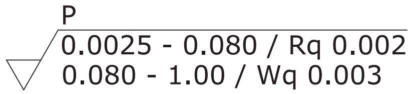


Figure 4.6 Surface texture specification with both Rq and Wq indicated.

With ground surfaces, it is sometimes necessary to specify both a minimum and a maximum amount of roughness allowed. This can be done with the U and L prefixes on the notation. Specifically,

U A – B / Rq C
L D – E / Rq F

where:
A = short scale-length limit of the spatial band in mm,
B = long scale-length limit of the spatial band in mm,
C = maximum RMS roughness tolerance in μm,
D = short scale-length limit of the spatial band in mm for the low limit,
E = long scale-length limit of the spatial band in mm for the low limit, and
F = minimum RMS roughness tolerance in μm.

If there is no lower limit, the U designation is not required. If there is only one Rq specification, it is presumed to be the upper bound of the RMS roughness for the surface. Further, the lower limit of the spatial bands is not required. For example, in Fig. 4.7, the minimum and maximum roughness specification is given for a ground surface. The surface is given as a ground surface, with at least 0.05 μm RMS roughness and no more than 2 μm RMS roughness for all spatial scales less than 2 mm.

As is discussed further in Sec. 4.3, the RMS specification is incomplete without spatial band limits. For ISO 10110-8, if no spatial band limit is provided, a default spatial band applies. For RMS texture errors specified with Rq, the default spatial band is 0.0025 mm to 0.080 mm. For RMS texture errors specified with Wq, the default spatial band is 0.080 mm to 2.50 mm.

4.2.3 Slope tolerance notations

As deterministic figuring has become more common in the past 20 years, the problem of surface waviness has become increasingly important. In some cases, a band-limited RMS waviness may be insufficient. It has been observed that surface slopes can be a significant predictor of system performance.[12]

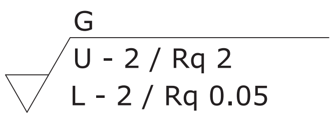


Figure 4.7 Roughness specification with both the upper and lower limit indicated.

Specifications of surface slope therefore may be a preferred method of specifying the waviness tolerance. This can be achieved by using an RMS slope specification.

Instantaneous slope is the derivative of the surface shape. In one dimension this slope can be defined as the derivative of the profile, but the local slope magnitude of a discretized profile is highly dependent on the ordinate spacing. The local slope is calculated using Chetwynd's formula.[13] Chetwynd's formula for local slope is given in Eq. (4.1), where Δx is the spacing between adjacent profile points, and z_i is the height of the ith point in the profile:

$$\frac{dz_i}{dx} = \frac{1}{60\Delta x}(z_{i+3} - 9z_{i+2} + 45z_{i+1} - 45z_{i-1} + 9z_{i+2} - z_{i-3}) \tag{4.1}$$

The local slope is unitless but we express the slope as the arctangent of the surface slope in microradians. When specifying the surface texture using a band-limited RMS slope, the specification is placed under the horizontal line in the symbol as shown in Fig. 4.8, and takes the form:

`A – B / RΔq C`

where:
`A` = short scale-length limit of the spatial band in mm,
`B` = long scale-length limit of the spatial band in mm, and
`C` = RMS slope tolerance in μrad.

For example, in Fig. 4.8, the waviness tolerance is specified in terms of RMS slope. Here, the RMS slope shall be less than 0.6 μrad, for all spatial wavelengths from 0.01 mm to 5.0 mm.

As is discussed further in Sec. 4.3, the RMS slope specification is incomplete without a spatial band limit. For ISO 10110-8, if no spatial band limit is provided, a default spatial band applies. For RMS texture errors specified with `RΔq`, the default spatial band is 0.080 mm to 2.50 mm.

4.2.4 Power spectral density notations

While the band limited RMS and RMS slope specifications are quite powerful, they do not control the magnitude of the specific spatial frequencies present on the finished surface. A single, high-fidelity periodic waviness can

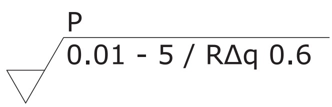


Figure 4.8 Waviness specification using the RMS slope notation.

have the same RMS or RMS slope as a surface with more randomness but lower fidelity waviness. In some applications, the fidelity of the periodic shapes is important. [14,15] In other cases, the sensitivity to periodic features may fall off versus spatial frequency in a well-understood way. For these types of situations, a power spectral density (PSD) function may be a better model for the impact of the surface errors.[16,17] ISO 10110-8 provides a notation for tolerancing the PSD of the finished piece.

When specifying the surface texture using a PSD function, the specification is placed under the horizontal line in the symbol as shown in Fig. 4.9, and takes the form:

`A – B / PSD C/D`

where:
`A` = short scale-length limit of the spatial band in mm,
`B` = long scale-length limit of the spatial band in mm,
`C` = value of the PSD limit function at the spatial frequency 1 mm^{-1}, and
`D` = power of the spatial frequency term in the PSD limit function.

Namely, the power spectral density of the surface shall be less than or equal to

$$PSD \le \frac{A}{f^B} \text{ for all } \frac{1}{D} \le f \le \frac{1}{C} \tag{4.2}$$

where:
f = spatial frequency in mm^{-1}.

Note that when plotted on a log–log graph, the PSD function becomes a straight line with a slope of `D` and a value `C` nm^2-mm at a 1 mm spatial period. For example, in Fig. 4.9, the texture tolerance is specified in terms of a PSD. Here, when plotted on a log–log graph, the PSD of the surface at all points shall be less than the line defined by Eq. (4.3) for all spatial wavelengths from 0.002 mm to 1.0 mm:

$$PSD = \frac{1.0}{f^2} \text{ in units of nm}^2\text{-mm} \tag{4.3}$$

For RMS texture errors specified with the PSD, the default spatial band is 0.080 mm to 2.50 mm.

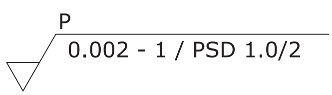


Figure 4.9 Texture specification using the power spectral density (PSD) notation.

4.2.5 Areal versions of roughness and waviness specifications

In 2018 the standard was revised to include the areal versions of the roughness and waviness specifications. While they are conceptually the same as the linear versions, the mathematics are different since the data set is now a 2D array of z-height information rather than a 1D line of z-height information. As a result there is a different notation for the areal equivalents of the various texture specifications. In the case of band-limited RMS, the notation `Sq` is used instead of `Rq` or `Wq`. In the case of slope, `SΔq` replaces `RΔq`. In the case of PSD, `APSD` replaces `PSD`. Note that for areal RMS specifications the default spatial band is the same as that for `Rq` specifications: 0.0025 mm to 0.080 mm, while for areal slope the default band is 0.080 mm to 2.50 mm.

Figure 4.10 shows an example of the areal equivalent of the band-limited RMS. In this case, the areal RMS surface texture shall be less than 3 nm for spatial scales from 0.002 mm to 1.00 mm. Figure 4.11 shows an example of the areal equivalent of the RMS slope. In this case the areal RMS surface slope shall be less than 0.6 μrad, for spatial scales from 0.01 mm to 5.00 mm. Figure 4.12 shows an example of the areal equivalent of the PSD. Here, when plotted on a 2D log–log graph, the PSD of the surface at all points shall be less than the cone defined by Eq. (4.4) for all spatial wavelengths from 0.001 mm to 1.0 mm:

$$APSD = \frac{1.0}{f^2} \text{ in units of nm}^2\cdot\text{mm}^2 \tag{4.4}$$

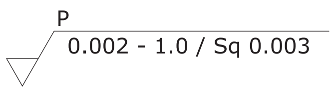


Figure 4.10 Surface texture specification using the band-limited areal RMS.

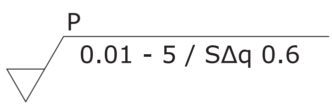


Figure 4.11 Waviness specification using the areal RMS slope notation.

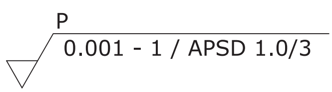


Figure 4.12 Texture specification using the areal power spectral density (APSD) notation.

4.2.6 Table notation

Typically, the surface texture notations are shown in the drawing field—with the tip of the texture symbol in contact with the line representing the surface, or a subsidiary line as shown in a subsequent section. In tabular drawings, however, it is also possible to put the texture symbol in the appropriate table sub-field for a given surface.

4.2.7 Differences in slope using Part 5 and Part 8

It is possible to specify surface slope tolerances in two ways; as discussed in this chapter using the ISO 10110-8 slope notation, or with the ISO 10110-5 notation discussed in the previous chapter. It is important to be aware that in the two cases the notations are quite different and the allowed tolerances are not the same. In the texture slope case, only RMS slope and areal RMS slope can be specified. This treatment is appropriate for surface undulations that are expected to be somewhat uniformly distributed across the surface and are therefore amenable to statistical function tolerances. In the case of the surface form slope, both the peak slope and RMS slope can be specified using either linear or areal calculations. Moreover, the slope is calculated differently. For texture, the slope values are calculated using Chetwynd's smoothing formula, while for the form slope tolerances no smoothing is applied to the instantaneous slope prior to the spatial filtering of the band limits. The differences are subtle and the slope tolerance will—for the most part—work equivalently for most surfaces. Care must be taken, however, since the two calculations can provide slightly different results for the same surface data.

The intention is for the texture slope tolerance described here to be used for short scale-length errors in the roughness and waviness domain, and for the surface form slope tolerance described in ISO 10110-5 (chapter 3) to be used for the longer scale-length errors. As a rule of thumb, for surface errors with less than five cycles across the surface, ISO 10110-5 slopes should be used; while for surface slope errors of a higher spatial frequency, the texture slope described in ISO 10110-8 should be used instead.

4.3 Meaningful Surface Texture Specifications

Most tolerances on optical elements can be derived or calculated from the application requirements using computer-aided optical design programs. There are few guidelines or tools, however, to facilitate selection of appropriate tolerances for surface roughness and waviness. Often a simple legacy specification (e.g., 3 Å RMS) is used, with little thought for either the cost of achieving the specification or the penalty for failing to achieve it. These legacy specifications are often ambiguous, inadequate, and in some cases completely meaningless. This section provides some basic rules and equations to lend insight into the meaning of surface texture errors, and their impact on performance and guidance on how to select appropriate tolerances for roughness and waviness.

4.3.1 Spatial bands and RMS calculations

When considering surface texture specifications, it is important to understand the relationship between a specified roughness and the spatial band over which that roughness is to be evaluated. This has been discussed in the literature[5,6] but the issue is still not widely understood.

Simply put, for any typical, well-behaved optical surface, the band-limited RMS will increase for larger spatial bands. As an example, a surface that achieves 1 nm RMS for the spatial band of 0.001 mm to 0.01 mm will have a larger RMS for a spatial band of 0.001 mm to 0.1 mm.

This is best understood through the power spectrum of the surface profile. The power spectrum of a profile can be thought of as the square magnitude of the Fourier Transform of the profile, utilizing an appropriate weighting function. For a digitized profile with an evaluation length L and with N equidistant points separated by a sampling interval Δx, the function is approximated by Eq. (4.5)[9]

$$PSD(f) = \left(\frac{\Delta x}{N}\right)\left|\sum_{j=1}^{N} Z_j e^{-i2\pi f(j-1)\Delta x}\right|^2 \tag{4.5}$$

where:
$f = \frac{K}{L}$ and
K = an integer that ranges from 1 to $\frac{N}{2}$.

Mathematically it can be shown[18,19] that the sum of the band-limited PSD is the square of the RMS evaluated over the same band. That is, the RMS increases for larger spatial bands and decreases for smaller spatial bands. Consider the case where the PSD of the profile of the surface is given by Eq. (4.3).

We can evaluate this function to determine the RMS of the surface over various spatial bands of interest—for example, the default bands of ISO 10110-8: 0.0025 mm to 0.080 mm and 0.080 mm to 2.50 mm; and the band used for the polish grades: 0.0025 mm to 1.00 mm. For this analysis we will take Δx to be 0.001 mm for simplicity. The results are shown in Table 4.3.

This case is being provided as a theoretical example—we do not expect a surface to have a perfect f^2 falloff PSD—but the results are consistent with the PSD of real surfaces. What we can see is that the magnitude of the surface RMS is highly dependent on the spatial band used for evaluation for the *very same surface*. Thus, an RMS roughness or waviness specification without a spatial band is meaningless.

Table 4.3 Square root of the area under the discrete PSD defined by Eq. (4.3) for different spatial bands.

Band used	Short cutoff	Long cutoff	RMS
Default Roughness Band	0.0025 mm	0.080 mm	0.5 nm
Default Waviness Band	0.080 mm	2.500 mm	1.7 nm
Polish Grade Band	0.002 mm	1.000 mm	1.4 nm

4.3.2 Meaningful roughness and waviness specifications

Assuming the surface texture obeys some statistical distribution properties, the RMS error value can be related to scatter.[20] For a well-behaved polished surface of a lens, the fraction of light scattered (S) is given by Eq. (4.6), where n is the index of the material, σ is the RMS, and λ is the wavelength of light. For mirror surfaces, set n equal to -1:

$$S = \left(\frac{2(\mathrm{n}-1)\pi\sigma}{\lambda} \right)^2 \tag{4.6}$$

From this we can see that the total scatter goes as the square of the ratio of the roughness and the wavelength. For example, a Germanium surface ($n = 4.00$), which has an RMS roughness of 20 nm over the scale-lengths of interest and is used in the long-wavelength infrared band ($\lambda = 8$ μm), will scatter 0.06%; not a significant factor. The same surface used as a mirror in the visible ($n = -1$, $\lambda = 0.56$ μm) would scatter 5% of the incident light. The inverse square wavelength factor is why residual surface texture errors are rarely a problem in the infrared but are often a concern in the visible and become extremely important in the ultraviolet.

4.3.3 Considerations when choosing a spatial band

The impact of scattering from surface texture can be quite different depending on where the scattered light goes. For high-spatial-frequency surface texture (i.e., roughness) the scatter tends to be treated as a loss since the angles associated with the high spatial frequencies tend to be large. For lower-spatial-frequency texture (i.e., waviness) the scatter angles can be closer to zero, resulting in a veiling glare or stray light impact on the system performance. Determining the appropriate spatial bands for a surface texture specification requires an understanding of this in the context of the application.

It has been shown elsewhere[11] that the scattered light follows the grating equation (again with many assumptions)

$$\theta_2 = \sin^{-1}\left(\frac{\lambda}{L}\right) \tag{4.7}$$

where:
L = spatial period of interest and
λ = wavelength of light.

A shorter spatial period corresponds to a higher scattering angle. Thus the texture can be assumed to scatter light into a host of angles related to the spatial periods covered by the spatial band. For visible light (e.g., λ = 0.56 μm), the default roughness spatial band from ISO 10110-8 equates to roughness that scatters light at angles from 0.4 degrees to 12.8 degrees. For many applications this is perfectly appropriate for a roughness specification that is to be treated as a loss. For other applications the near-angle end of the scatter distribution may cause a veiling glare that is far more egregious than simple surface texture, and multiple specifications for various spatial bands of surface texture (e.g., roughness and waviness) may be required.

Another consideration when selecting a spatial band range is the expected profile based on the manufacturing method. For deterministic small-tool figuring methods, a signature residual surface form error is expected. For diamond-turned surfaces, this tool signature is typically in the 5 μm to 40 μm spatial wavelength range. For small-tool polishing systems such as those used in magnetorheological finishing (MRF) or computer-controlled optical surfacing (CCOS) the spatial wavelength of the tool signature is much larger; typically 0.5 mm to 5 mm or more. If no smoothing is to be applied the small-tool signature can be so pronounced that it dominates the surface texture RMS. In this case a judicial choice of a cutoff between the waviness and roughness tolerance can be helpful in controlling the surface texture. In Fig. 4.13 we see an example of a surface texture with a dominant tool signature. In the first case a longer cutoff wavelength is used and the tool

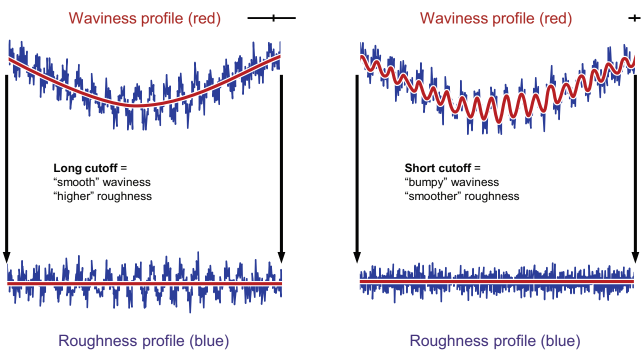


Figure 4.13 Choosing an appropriate evaluation length for a given expected surface profile.[21]

signature shows up in the roughness band. In the second case a shorter cutoff wavelength is used, resulting in a bumpier waviness profile and a smoother roughness. It is typically best to choose the cutoff that will best isolate the tool signature such that the tool signature can be controlled with a specific texture tolerance, without driving the cost through a tight specification on waviness scale-lengths with little or no impact on the performance.

One final concern when choosing a spatial band is to be aware of the metrology method to be used to verify compliance. Tactile profilers can have fairly long traces and cover significant spatial bands with high fidelity, but with very little surface area coverage. Optical profilers are increasingly common for surface texture measurements but have specific limitations on the spatial band that can be evaluated using a given configuration. Many optical profilers are based on microscope configurations—and will have valid data for spatial bands that are of order-three undulations across the field of view at the low frequency end and perhaps five pixels per undulation at the high-frequency end.[22] For example, if the roughness specification is to be validated using a scanning white-light interferometer with a 10× objective that has a 0.5 mm square field of view and a 1 MP camera, the spatial scale length data

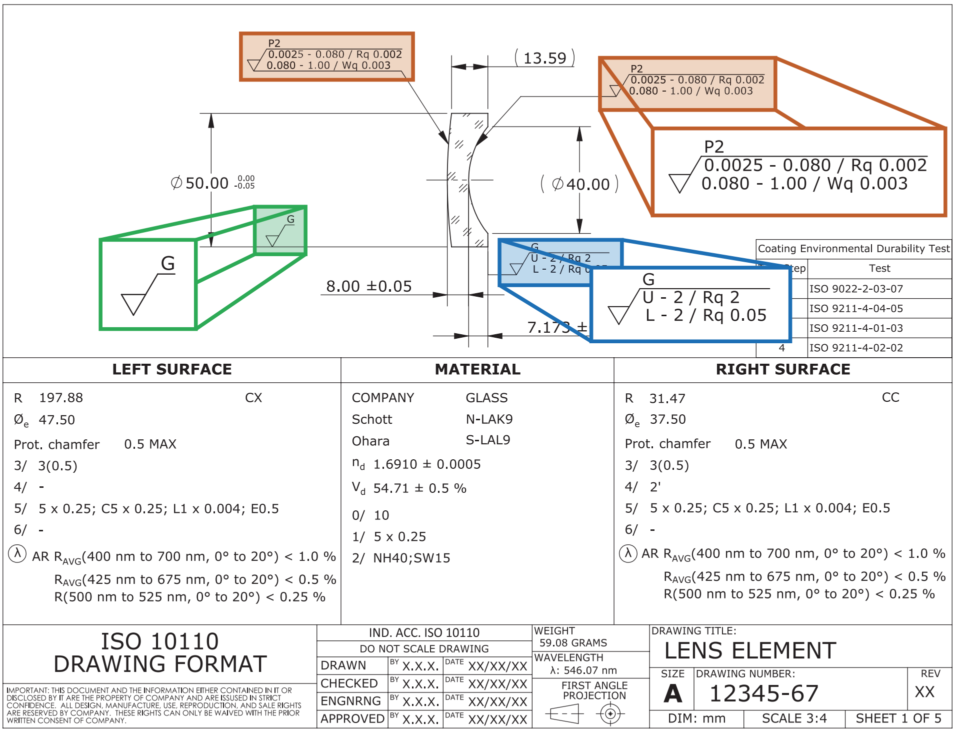


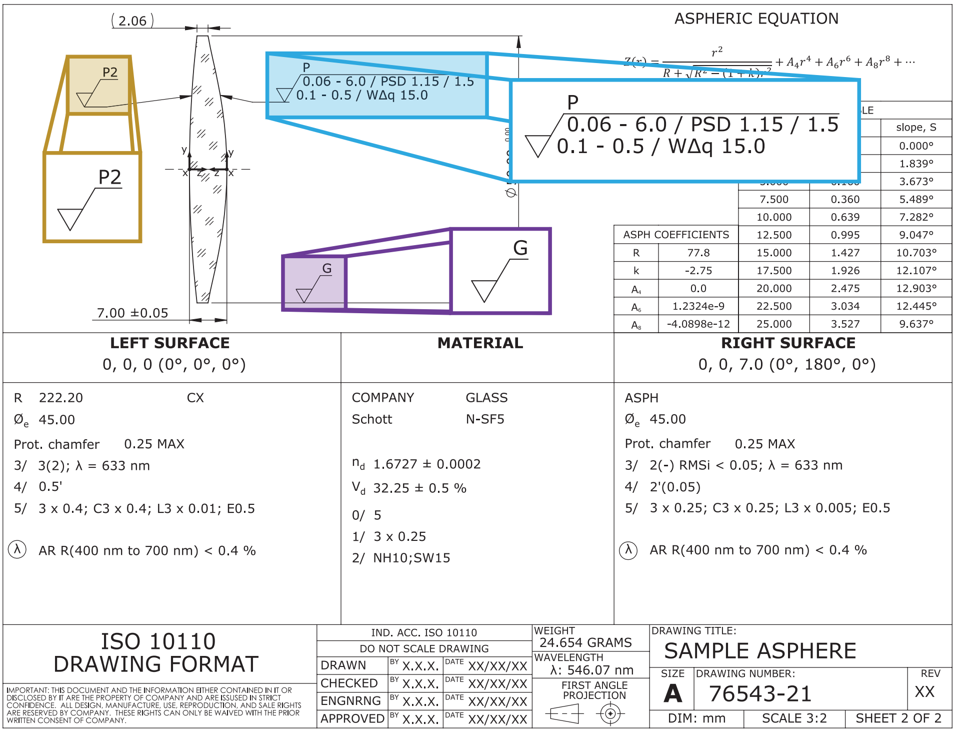
Figure 4.14 Example drawing with ground surface specifications on the edges and face flat; and roughness, waviness, and polish-grade specifications on the polished surfaces.

range for that measurement would be only 0.0025 mm to 0.16 mm or even less.

4.4 Drawing Example

Example drawings with some specific tolerance notations and their meaning are given below. In addition, ISO 10110-8 has an extensive Annex B containing many more examples.

Figure 4.14 shows an ISO 10110 drawing with four surface texture notation symbols. The face-flat on the right surface is given as a ground surface, with at least 0.05 μm RMS roughness and no more than 2 μm RMS roughness for all spatial scales less than 2 mm. The cylindrical edge is defined simply as a "ground" surface, with no particular requirement on roughness. The two optical surfaces are each specified with three different surface texture requirements. The first is that the surfaces be polish grade 2 or better; that is, no more than 4.0 nm RMS over the spatial band from 0.002 mm to 1.0 mm. The second requirement is that the surface roughness over the spatial band from 0.0025 mm to 0.080 mm shall be less than 2.0 nm RMS. The third



7.500	0.360	5.489°
10.000	0.639	7.282°
12.500	0.995	9.047°
15.000	1.427	10.703°
17.500	1.926	12.107°
20.000	2.475	12.903°
22.500	3.034	12.445°
25.000	3.527	9.637°

Figure 4.15 Example drawing with a spherical left surface having polish grade P2 and an aspheric right surface, with the texture specified using an RMS slope waviness and an additional surface PSD requirement.

requirement is that the surface waviness for spatial wavelengths from 0.080 mm to 1.000 mm shall be less than 3.0 nm RMS. The fact that these surface textures overlap is not a concern; the finished surface must comply with all three criteria simultaneously.

The second example, Fig. 4.15, shows an aspheric optical element. The cylindrical edge is identified as "ground" with no specific texture requirement. For the spherical surface (the left surface), the texture is specified using the standard polish code `P2`. On the aspheric surface (the right surface), the texture is specified with an RMS slope specification and a power spectral density tolerance. The waviness slope specification requires the surface slope be less than 150 μrad over spatial scale lengths from 0.1 mm to 0.5 mm. The PSD tolerance requires the power spectrum of the optical component, when plotted on a log–log scale from 0.06 mm to 6.0 mm, to be below the line defined by PSD $\leq 1.15\ \text{nm}^2\text{-mm} / f^{1.5}$, where f is the surface spatial frequency.

References

1 International Standards Organization (ISO), *Preparation of Drawings for Optical Elements and Systems—Part 8: Surface Texture*, ISO 10110-8:2019 (Geneva: 2019).
2. I. M. Hutchings, "Leonardo da Vinci's studies of friction," Wear 360–**361**, 51–66 (2016).
3. E. J. Abbott and F. A. Firestone, "Specifying surface quality: a method based on accurate measurement and comparison," *Mech. Eng.* **55**, 569–572 (1933).
4. A. E. Conrady, *Applied Optics and Optical Design: Part 2*, 2nd ed., Dover Publications (London: 1988).
5. J. E. Harvey and A. K. Thompson, "Scattering effects from residual optical fabrication errors," *Proc. SPIE* **2576**, 155–174 (1995) [doi: 10.1117/12.215588].
6. R. N. Youngworth and B. D. Stone, "Simple estimates for the effects of mid-spatial-frequency surface errors on image quality," *Appl. Opt.* **39**(13), 2198–2209 (2000).
7. US Department of Defense (USDOD), *Surface Roughness, Waviness and Lay*, MIL-STD-10 (Arlington, Virginia: 1949).
8. USDOD, *Military Standard: Surface Roughness, Waviness and Lay,* MIL-STD-10A (Arlington, Virginia: 1963).
9. American Society of Mechanical Engineers (ASME), *Surface Texture (Surface Roughness, Waviness, and Lay)*, ASME B46.1 (New York: 2019).
10. ISO, *Indication of Surface Texture in Technical Product Documentation*, ISO 1302 (Geneva: 2002).

11. J. M. Bennett and L. Mattsson, *Introduction to Surface Roughness and Scattering*, 2nd ed., Optical Society of America (Wasnington, DC: 1999).
12. J. R. Rogers, “Slope error tolerances for optical surfaces,” *Proc. SPIE* **10316**, 1031605 (2007) [doi: 10.1117/12.725057].
13. D. G. Chetwynd, “Slope measurement in surface texture analysis,” *J. Mech. Eng. Sci.* **20**(3), 115–119 (1978).
14. D. M. Aikens, A. Roussel, and M. Bray, “Derivation of preliminary specifications for transmitted wavefront and surface roughness for large optics used in inertial confinement fusion,” *Proc. SPIE* **2633**, 350–360 (1995) [doi: 10.1117/12.228287].
15. M. Shibuya et al., “Classification of undulated wavefront aberration in projection optics by considering its physical effects,” *Opt. Eng.* **46**(5), 053001 (2007) [doi: 10.1117/1.2734995].
16. J. K. Lawson et al., “Specification of optical components using the power spectral density function,” *Proc. SPIE* **2536**, 38–50 (1995) [doi: 10.1117/12.218430].
17. R. N. Youngworth, B. B. Gallagher, and B. L. Stamper, “An overview of power spectral density (PSD) calculations,” *Proc. SPIE* **5869**, 58690U (2005) [doi: 10.1117/12.618478].
18. M.-A. Parseval de Chenes, “Dissertation on series and on the complete integration of a second order linear partial difference equation with constant coefficients,” *Memories Present. by Inst. Sci. Lett. Arts*, by Var. Sch. read its Assem. Sci. Math. Phys., 638–648 (1806).
19. M. Plancherel and M. Leffler, “Contribution to the study of the representation of an arbitrary function by definite integrals,” *Rend. del Circ. Mat. di Palermo* **30**(1), 289–335 (1910).
20. D. M. Aikens, “Meaningful surface roughness and quality tolerances,” *Proc. Int. Optical Design Conf.* **2010**, ITuD3 (2010).
21. M. C. Malburg, “Three steps to understanding surface texture,” *Digital Metrology*, May 6, 2014, https://digitalmetrology.com/tutorials/3-steps-to-understanding-surface-texture.
22. E. L. Church, T. V. Vorburger, and J. C. Wyant, “Direct comparison of mechanical and optical measurements of the finish of precision machined optical surfaces,” *Opt. Eng.* **24**(3), 388–395 (1985) [doi: 10.1117/12.7973494].

Chapter 5
Surface Imperfection Tolerances

Surface imperfections are localized flaws, such as scratches and digs, on the surface of an optical element. Surface imperfections that are not localized—i.e., cover the entire surface—are considered texture and are toleranced using the notation given in Chapter 4. Typically, when considering surface imperfections we include other features such as edge chips and coating imperfections. The ISO 10110 standard includes two methods of indicating surface imperfections. The first method is analogous to the scratch and dig visibility standard found in US MIL-PRF-13830B and the related visibility method described in ANSI/OEOSC OP1.002. The second method is analogous to the withdrawn DIN 3140's dimensional specification.

5.1 Background

Surface imperfection specifications are among the most misunderstood and misapplied standards in the world of optical engineering.[1–3] Most indications fail to even reference a standard, trusting instead that the manufacturer will interpret the notation in the intended manner with no further guidance than `60-40` scratch and dig or `5/ 4 × 0.1`. The long and complex history of these specifications combined with their frequent citation as cause for rejection requires more attention from the optical engineer in preparing a surface quality tolerance. In the event that one should attempt to develop a meaningful surface quality specification from existing standards, one is often confronted with bad options: a standard that is expensive to implement, is often misinterpreted, or might not fit the required application.

The surface imperfection notation for optical surfaces, often referred to as "scratch and dig," is found today in Section 3.5 of MIL-PRF-13830B,[4] an active standard supported by the US military. It is based on MIL-O-13830A, first released as a military standard in the US in 1954 and revised once in

1963. It has since been used as the basis of ANSI/OEOSC OP1.002,[5] most recently (at the time of this writing) revised in 2017.

Because most surface imperfections are not a function of the fabrication process but are instead caused by mishandling and processing errors, it is typical to inspect 100% of the parts rather than batch sampling. In addition, these imperfections are small and localized, thus most parts must have 100% of their area inspected. Such inspection is tedious and expensive, and requires skilled personnel. One practical idea to speed up inspection using a visual classification and grading system was proposed by MacLeod and Sherwood of Eastman Kodak in 1945.[6] This method—based on a subjective comparison of a surface imperfection versus a standard artifact, set to determine the visibility or "grade" of the imperfection—was adopted by the US military and incorporated into MIL-O-13830. This approach is a cost-effective way to control surface imperfections for cosmetic purposes[7] but has limited application in specifying the surface quality for functional applications.[8] A dimensional grading system with wide acceptance is the German DIN 3140 standard, Part 7, where imperfections are compared to chrome on glass rectangles and line width standards of specific sizes.[9]

These two systems, described in more detail in the technical report ISO/TR 21477, remain in wide use throughout the world.[10] In the 2017 edition of ISO 10110, both are included with an available notation, supported by visual inspection techniques described in ISO 14997.[11]

Continuing practical problems associated with both methods persist. Some of the continuing problems are

(1) Availability of comparison standards and quantitatively categorizing their visibility,
(2) Subjectivity of visual inspection,
(3) Difficulty of measuring the actual widths of small scratches due to diffraction effects, and
(4) Increasing difficulty in finding highly skilled inspectors to do the comparisons.

These problems have led to the development of several instruments to aid in objectively quantifying the inspection of optical elements using machine vision. As of this writing, a new technical report (ISO TR 14997-2) is under development to aid in the standardization of these machine vision systems. While machine vision systems are increasingly available, most optical components specified in either the visibility or dimensional notation of ISO 10110-7 are still inspected visually by trained technicians. This costly and challenging inspection renders the surface imperfection specification the greatest cause for rejection of optical components, and the most vexing and misunderstood of all the optical tolerances.

5.2 Indications on Drawings

The drawing symbol for surface imperfections is `5/`. Appropriate quantifiers are added after the `5/` depending on which of the two specification systems are being used. The two methods of quantifying surface imperfections are mutually exclusive. The first method, referred to as the dimensional method, is a quantitative grading system related to the imperfection size (i.e., dig area or scratch width). If the quantifier is of the form $N \times A$, where N and A are numbers separated by an $\times$, the dimensional method is being used. The second method, referred to as the visibility method, is an inspection approach based on the visibility or appearance of surface imperfections when observed under specific lighting conditions. If the quantifier is the form `S-D`, where `S` and `D` are numbers separated by a hyphen, the visibility method is being used.

5.2.1 Dimensional method drawing notation and interpretation

In the dimensional method, surface imperfections are controlled according to their size. When the dimensional method is used, there are five categories of quantifiers. The notation is of the form:

$$5/\ N_g \times A_g;W\ A_w;C\ N_c \times A_c;L\ N_l \times A_l;E\ A_e$$

where:
N_g = number of general imperfections of the maximum size grade allowed,
A_g = maximum allowed size grade of a general imperfection,
A_w = maximum extent of any general imperfection,
N_c = number of coating imperfections of the maximum size grade allowed,
A_c = maximum allowed size grade of a coating imperfection,
N_l = number of long imperfections (i.e., scratches $\geq$2 mm long) of the maximum width grade allowed,
A_l = maximum allowed width of a long imperfection, and
A_e = maximum allowed extent of any edge chip.

Note that only the first pair ($N_g \times A_g$) is required; all other tolerances are optional. If no additional tolerances are indicated, then all surface imperfections (general, long, and coating) are subject to the rules for a general imperfection. This dimensional method is based on the DIN 3140 standard commonly used in Germany, which has been withdrawn.

General Imperfections

In the dimensional method, the quantifier N_g is the maximum number of allowed imperfections of the maximum specified size, A_g. N can be any positive integer but would usually be in the range of 1 to 5 for small components less than 25 mm diameter, rising roughly as the area of the

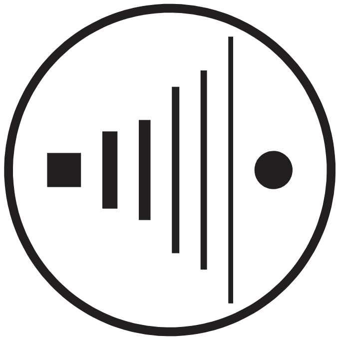


Figure 5.1 Imperfections of the same grade, varying in aspect ratio from 1:1 to 100:1.

component. For example, a 50 mm part could allow 4 to 20 imperfections or more, depending on the application.

The quantifier A_g is the square root of the area of an allowed imperfection; for example, the length in mm of the side of a square imperfection of the maximum allowed area. It must be a member of the Renard R5 series of numbers,[11] given in Table B.1 of ISO 14997; that is, 0.4, 0.25, 0.16, 0.1, 0.063, 0.04, 0.025, 0.016, 0.01, 0.0063, and 0.004. Thus, if A_g = 0.1, the area of the maximum allowed imperfection is 0.01 mm^2. The imperfection may not be square, and in general will not be. Figure 5.1 illustrates several imperfections of the same area but having varying aspect ratios.

Coating Imperfections

The indication for coated surfaces need not be used on a drawing for a coated element. If there is no specification for coating imperfections, the general imperfection specification is taken to apply to the finished part. It is sometimes useful, however, to tolerance coating imperfections separately from uncoated surface imperfections. The format for coating imperfections is essentially the same as for general imperfections, with N_c replacing N_g and A_c replacing A_g, except the specification is preceded by the letter `C`. When the coating imperfection specification is used, the general imperfection specification is taken to apply to the substrate prior to coating.

Long Imperfections

When using the `L` designation for long imperfections, the number A_l refers to the maximum allowed width of all long imperfections in mm. A long imperfection is defined as one in which the length is greater than or equal to 2 mm. The number N_l is simply the number of maximum grade long scratches allowed independent of their length. Long scratches less than the maximum width are allowed if the combined width of all scratches does not exceed N_l times A_l, where scratches of width less than `0.25A` are not counted.

Maximum Extent of Imperfections

It is possible to constrain the aspect ratio of all imperfections using the W indication. When the W indication is used, the number A_w refers to the maximum dimension of any imperfection in mm. That is, all imperfections regardless of grade must be shorter than A_w in the long dimension. This notation may be helpful when it is desirable to only allow a certain length of scratches on the component. Since these notations are fundamentally incompatible, the W requirement and the L requirement cannot be used together.

Accumulation of Imperfections

In the dimensional method, imperfections smaller than the maximum allowed grade are allowed, but subject to the maximum accumulation rule. The total area of all imperfections may not exceed N times A^2, the total allowed area of all the maximum allowed imperfections. In this accumulation, imperfections with a size grade of 0.16A or less shall not be counted in figuring the total area.

The total area of all imperfections must not exceed the area of the original indication. This requirement can be extremely stringent. If there were only two maximum-sized imperfections allowed, the total allowed area obscured would be 2 times A^2. Table C.1 of ISO 14997 can be helpful in showing how many smaller imperfections have the same total area as a larger imperfection. One can think of the unused maximum grade imperfections are being "exchanged" for smaller imperfections. For example, if the specification were 2×0.4, and no 0.4 grade imperfections were present, they could be exchanged for 5 (2.5 times 2) imperfections of grade 0.25, the next grade down, because they have the same area. Alternatively, if no 0.25 grade imperfections were present either, then the two 0.4 grade allowed imperfections could be exchanged for ~13 (6.3 times 2) imperfections of grade 0.16, the next lower grade in the R5 system, and so on.

There is a limit to accumulation. If A is the largest allowed imperfection, imperfections of 0.16A and smaller are not counted. This limit is reasonable because there is the possibility of subdividing a single imperfection A into 16 imperfections of a size 0.25A. In some cases, it may be necessary to add a note to limit the accumulation of imperfections to a larger grade to prevent unnecessary yield loss.

Concentration of Imperfections

In addition, while imperfections may be subdivided, concentrations of imperfections are not permitted. A concentration occurs when more than 20% of the allowed imperfections (by area) fall inside an area that is 5% of the area of the surface under inspection. If there are ten or fewer imperfections, a concentration is present when two or more maximum allowable imperfections (or smaller imperfections of the same area) lie within a 5% sub-area.

This requirement is also stringent (very tight). If the surface has a unit diameter, then no two imperfections may lie within a distance of 0.224 of the diameter of each other. This constraint may not be applicable to all applications. Since it is a tight requirement, it is often wise to override this requirement by adding a note to the drawing that indicates: `concentration rules do not apply`.

Edge Chips

Edge chips have a single quantifier A_e that indicates the maximum distance that a chip may extend from the physical edge of the element toward the center. The measurement of the chip length does not consider the size of the chamfer and does not depend on the location of the clear aperture. Note that even if there is no edge chip specification, edge chips are not permitted to enter the clear aperture. The edge chip quantifier is preceded by an E.

Any number of edge chips are allowed—provided they do not exceed the specified length from the edge, measured along the surface.

5.2.2 Visibility method drawing notation and interpretation

In the visibility method, surface imperfections are controlled according to a "scratch and dig" visibility system, where the imperfections are graded using a visual comparison artefact. When the visibility method is used, there are three categories of quantifiers. The notation is of the form:

$$5/\ S\text{—}D;C\ S_c\text{—}D_c;E\ A_e$$

where:
S = maximum visibility grade of an allowed scratch,
D = maximum size grade of an allowed dig,
S_c = maximum visibility grade of an allowed coating scratch,
D_c = maximum size grade of an allowed coating "dig", and
A_e = maximum allowed extent of any edge chip.

Additionally, a note on the drawing must be included to state what the comparison standard that will be used for the final inspection, including the make and model.

Note that because the appearance of the imperfection standards is subjective and could vary from set to set, the drawing note of the make and model of the comparison standard to be used is **required**.

As with the dimensional method, the coating and edge chip specifications are optional. If there is no coating specification, all coating imperfections are judged against the scratch and dig specification S—D.

The visibility method is based on the MIL-PRF-13830B standard commonly used in the US.

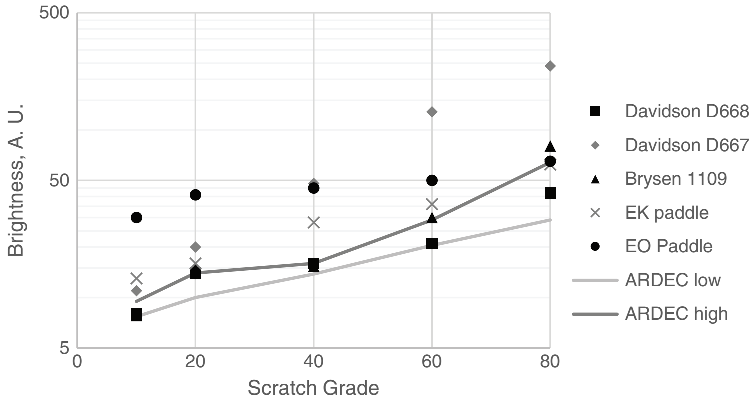


Figure 5.2 Relative brightness or visibility of some of the more common comparison standards.[12]

General Imperfections

In the visibility method, the quantifier S for uncoated or coated scratches and D for uncoated or coated digs is the maximum visibility grade permitted, based on the comparison set referenced in the note. While any comparison set can be used, it is best to use the standards with some degree of traceability back to the US Army limit standards.

Dig visibility is determined by apparent size; that is, how "big" does the dig appear, when compared to the comparison standard. The visibility of scratches, however, is based on the apparent brightness under specific illumination conditions.

The quantifier S is an arbitrary grade number associated with a specific comparison set. In the case of US Army-issued comparison standards, and presumably any standard that has been certified to be Army-traceable, the grades refer to a set of master scratches retained at the Picatinny Arsenal in New Jersey. However, it is not safe to assume that a scratch—which has determined to be a 60 according to the Davidson D668 comparison standard, for example—will also be a 60 when evaluated according to the Edmund Optics comparison standard.*

The quantifier D also refers to a specific comparison set but refers to a particular size dig. The dig grade of the comparison set is intended to be the diameter of the dig in units of tens of microns. The #50 dig should be 500 μm in diameter. As this requirement is fairly easy to understand, the digs in most comparison sets for the visibility method are fairly close to these values. Even

*Davidson D668 comparison standard and Edmund Optics comparison standard are examples of commercially available visibility standard sets.

so, this should only be presumed the case if the set is US Army-issued or is certified to be Army-traceable.

Relative brightness or visibility of the scratches in some of the more popular comparison standards are shown in Fig. 5.2. The brightness values, measured over several 1 mm lengths with a linear photodetector, are shown in arbitrary units on a log scale for clarity. The two gray lines are the limits based on the US Army limit standards. In theory, all Army-issued and Army-traceable comparison sets will have brightness between these two lines. However, as can be seen in Fig. 5.2, many of the popular comparison standards (see Table 5.1) deviate quite a bit from this requirement.

If you are trying to specify the cosmetic quality of your optics, then the scratch–dig tolerance is a good choice. The key is to specify a scratch and dig that fits your application. Most people are not familiar with the visibility of the scratches in the scratch–dig standard. Table 5.2 is a qualitative assessment of scratch visibility by number for a US Army-traceable set, and the most common scratch–dig pair for a given level of desired visibility.

Table 5.1 Some common visibility comparison sets in circulation and their allowed grades.

Make	Model	Army traceable?	Scratch grades	Dig grades
US Army ARDEC	C7641866	Yes	80, 60, 40, 20, 10	50, 40, 20, 10, 5
Brysen Optical,FLIR Systems	C7641866	If certified	80, 60, 40, 20, 10	50, 40, 20, 10, 5
Davidson Optronics	D668	Yes	80, 60, 40, 20, 10	50, 40, 20, 10, 5
Davidson Optronics	D667	No	80, 60, 40, 20, 10	50, 40, 20, 10, 5
Eastman Kodak	EK Paddle	No	160, 120 80, 60, 40, 20, 10	100, 70, 50, 40, 20, 10, 5
Jenoptik plastics	EO Paddle (#53-197) Thor Labs SAD paddle	No	160, 120 80, 60, 40, 20, 10	100, 70, 50, 40, 20, 10, 5

Table 5.2 Scratch visibility by number, and some common applications.

Scratch number	Scratch visibility under normal viewing conditions	Typical callout	Application
10	Very, very hard to see	`10-5`	Not recommended; dimensional method is more appropriate
20	Barely visible	`20-10`	Not recommended; dimensional method is more appropriate
40	Similar to the #20; slightly more visible	`40-20`	Close to image plane or reticle, or dark-field illumination
60	Somewhat visible	`60-40`	Precision optics
80	Easy to see	`80-50`	Commercial optics

Coating Imperfections

The indication for coated surfaces need not be used on a drawing for a coated element. If there is no specification for coating imperfections, the general imperfection specification is taken to apply to the finished part. Using the indication for a coated surface affords the designer additional flexibility in specifying surface imperfections, in that the imperfections will appear more or less visible than on bare glass, depending on whether the coating applied was reflective or antireflective. The format for coating imperfections is the same as for general imperfection but the specification is preceded by the letter `C`. When the coating imperfection specification is used, the general imperfection specification is taken to apply to the substrate prior to coating.

Number of Allowed Maximum Grade Imperfections

The allowed quantity of maximum grade imperfections is calculated based on the diameter of the inspection zone, usually the clear aperture. Specifically, the number of maximum grade digs allowed is the diameter of the zone in millimeters (mm) over 20, rounding up. The combined length of all maximum grade scratches allowed is one-quarter of the diameter of the zone.

Accumulation of Imperfections

In addition to the maximum grade imperfections allowed, the scratch and dig specification also requires all visible scratches and digs (when viewed according to the methods described in ISO 14997) to be evaluated and accumulated. The sum of the lengths of all the scratches, multiplied by the ratio of its grade number and the maximum allowable grade, must be less than or equal to half of the diameter of the clear aperture.

$$\sum\left(l_i \times \frac{s_i}{S}\right) \leq \frac{\text{Ø}}{2} \tag{5.1}$$

where:
l_i = length of a scratch i,
s_i = grade of a scratch i,
S = maximum scratch grade allowed, and
Ø = diameter of the clear aperture.

If there are no maximum grade scratches present, then this sum can be up to the full diameter of the clear aperture.

As with scratches, the digs (visible when viewed according to the methods described in ISO 14997) must all be evaluated and accumulated. The sum of the values of all the digs visible on the part must be less than twice the product of the maximum dig grade and the number of maximum digs allowed.

$$\sum d_i \leq 2 \times N \times D \tag{5.2}$$

where:
d_i = grade of dig i,
N = number of the maximum allowed dig, and
D = grade of the maximum allowed dig.

Concentration of Imperfections

Concentrations of imperfections are not allowed. Specifically, no more than four scratches can be found within a 6.35 mm circle, and the sum of all the grades in any 20 mm diameter cannot exceed 2 times D.

Note that these concentration rules differ slightly from those of the US MIL-PRF-13830B and ANSI/OEOSC OP1.002. In both of those standards, the scratch concentration rule applies only on parts in which the scratch and dig specification is `20-10` or tighter. In addition, with those standards, all digs must be separated by at least 1 mm if the specification is `20-10` or tighter. In the ISO system, the scratch concentration rule always applies, and the additional dig separation rule never applies.

Edge Chips

As with the dimensional notation, edge chips have a single quantifier `A` that indicates the maximum depth that a chip may extend from the physical edge of the element. See the edge chip paragraph in section 5.2.1 for the rules regarding edge chips.

5.3 Meaningful Surface Imperfection Specifications

5.3.1 Cosmetic blemishes

In nearly all cases, surface imperfections on optical components do not impact the function of the optical instrument and can be considered merely cosmetic blemishes. However, imperfections can still be interpreted as indicators of poor workmanship and be a cause for rejection, even though the imperfection itself has no influence on the image quality or resolution. Typically, the outermost surface in an optical instrument is specified more tightly than the internal components since the end-user will only view the outer surfaces. Since this casual inspection by the user will be most likely visual and without magnification, the visibility specification may be the best choice for specifying cosmetic imperfections. In infrared instruments, this visual evaluation will likely be done in reflection. In some cases, the instrument—especially for example virtual reality/augmented reality sets, rifle sights, binoculars, and telescopes—may be held 200 to 400 mm from the eye by a consumer with a store light as a background. In cases where the instrument or component is very expensive, such as an expensive one-off system for a spacecraft, it may be worth the expense of additional time for inspection to obtain the objectivity of the dimensional method, even though the concern is still one of cosmetics.

An area where cosmetic appearance more closely relates to instrument functionality is where an optical element is in very close proximity to an image plane conjugate, as in the case of a field lens or reticle. Surface imperfections on field lenses and reticles may be visible when viewed through the eyepiece of an instrument, depending on the magnification at which the element is viewed. With reticles that are designed to be viewed under high magnification, it is preferrable that they have only the faintest surface imperfections. In the case of field lenses and reticles, they are in practice viewed visually in transmission. This would render the visibility method, tested in transmission as in the arrangement shown in ISO 14997 Figures A.1 or A.2, the best candidate. Even so, visual inspection may not be sensitive enough to pick up imperfections small enough to still be objectionable, and a dimensional method using magnification would be more objective and quantitative.

5.3.2 Scatter and diffraction effects

In an imaging system, the functional impact of surface imperfections is to scatter nonimage-forming light into the image plane. In such a case, a surface imperfection is detrimental in two ways. First, light is scattered out of the image in proportion to the area of the imperfection relative to the area of the beam at the surface. This can cause a loss of signal at the image point of interest. The Strehl ratio, for example, is the peak of the point spread function for a point object relative to a geometrically perfect point image (i.e., a perfect lens). If we think of the Strehl ratio as representing the amount of light in the central lobe of the point image relative to what would be there if a perfect lens were used, we can imagine that if the imperfection is large enough with respect to the area of the beam at that surface, it is feasible to calculate the effect of the imperfection on the Strehl ratio.

Second, the scattered light may still be incident on the image plane near the intended central peak. This scattered light contributes to the background of the image, reducing contrast and potentially contributing veiling glare about the image of bright objects. If the application involves very high-contrast imaging or detecting a faint signal in close proximity to a bright signal, scattering from surface imperfections must be considered.

Finally, in cases involving coherent systems such as lasers, a sufficiently large imperfection can cause diffraction effects that can disrupt the beam uniformity over long distances. Such diffraction can even be the cause of laser-induced damage. In laser systems, care must be taken to calculate the ramifications of large-area imperfections on components in the optical train.

5.4 Inspection for Surface Imperfections

5.4.1 Test method

The test method for inspection of surfaces for imperfections is described in ISO 14997. This standard provides a comprehensive guide to inspection of optics to both the dimensional specification given in this book's Sec. 5.2.1 and the visibility specification given in this book's Sec. 5.2.2. The specific test configurations and methods are provided in Annex A of the standard, as is the illuminance required in the inspection station to conduct the comparison. These test configurations correspond exactly to the inspection setups defined in the MIL and ANSI standards. However, the specification of the illuminance required—rather than a specific type or wattage bulb—is a significant improvement.

The metrology standard also defines a hierarchy of inspection levels, as shown in Table 5.3. For the dimensional specification, the inspection levels are $\mathtt{IV}_{\mathrm{D}}$, for visual inspection without a comparison standard; $\mathtt{IS}_{\mathrm{D}}$, for subjective comparison against a dimensional reference artifact; and $\mathtt{IM}_{\mathrm{D}}$, for inspection with magnification and measurement. For visibility, the inspection levels are $\mathtt{IV}_{\mathrm{V}}$, for visual inspection without a comparison standard; and $\mathtt{IS}_{\mathrm{V}}$, for subjective comparison against a visibility artifact. Since the visibility system does not reference any physical attribute of the imperfections, no IM level using magnification and measurement is allowed.

5.4.2 Inspecting to the dimensional method

In the dimensional method it is expected that in most cases a part can be found to be free of surface imperfections with a quick visual inspection (inspection level $\mathtt{IV}_{\mathrm{D}}$). When one or more suspect imperfections are found, a comparison plate of bright chrome dots, rectangles, and lines on glass will be used in a visual comparison in reflection in arrangement A.3 (inspection level

Table 5.3 Inspection levels defined in ISO 14997.

Inspection level	Arrangement	Comparison Standard	Default conditions
DIMENSIONAL: Visual Inspection ($\mathtt{IV}_{\mathrm{D}}$)	Any	None	N/A
DIMENSIONAL: Comparison Inspection ($\mathtt{IS}_{\mathrm{D}}$)	View in reflection	Dimensional standard	5/ 1 × 0.63; L 1 × 0.1 or looser
DIMENSIONAL: Inspection with Magnification ($\mathtt{IM}_{\mathrm{D}}$)	Loupe or microscope	Dimensional standard	Tighter than 5/ 1 × 0.63; L 1 × 0.1
VISIBILITY: Visual Inspection ($\mathtt{IV}_{\mathrm{v}}$)	Any	None	N/A
VISIBILITY: Comparison Inspection ($\mathtt{IS}_{\mathrm{v}}$)	Given in Annex A	As indicated on drawing	When no inspection level is specified

$\mathtt{IS_D}$). The imperfections are compared based on the apparent area for general and coating imperfections. The maximum width of long imperfections must be quantified by observing how wide or narrow the scratch appears compared to the width of a chrome line. In case of conflict or uncertainty, or for specifications where the width of the scratches cannot be determined visually, the imperfection can be viewed under a loupe or even a microscope, since it is the dimensions that are of concern (inspection level $\mathtt{IM_D}$).

5.4.3 Inspecting to the visibility method

As with the dimensional method, in the visibility method most parts can be found to be free of troubling surface imperfections with a quick visual inspection (inspection level $\mathtt{IV_V}$). When one or more suspect imperfections are found, a subjective inspection with a comparison standard is required (inspection level $\mathtt{IS_V}$). The comparison standard used in the inspection is that identified on the drawing. The illumination arrangement is described in Annex A of ISO 14997, and any of the three arrangements can be used. These arrangements are the same as those described in the ANSI specification for surface imperfections and are based on the methods defined in the MIL-PRF-13830B Section 4.2.2 (inspection methods 1 and 2), and Annex C, Section C.4.5.3.1. The scratch imperfections are judged based on the visibility or brightness, and not the width; while the digs are evaluated based on their apparent size compared to the reference standard.

5.5 Drawing Example

As there is a large amount of data to consume with both the dimension and visibility methods for surface imperfections, seeing examples of each method on an ISO 10110 drawing should solidify this information.

Figure 5.3 provides an example of a surface imperfection notation with the dimensional method. The indication `5/` is present prior to any qualifiers. This example shows notations of `5 × 0.25` prior to optical coating, `5 × 0.25` imperfections after coating, 1 long scratch of 4 μm in width (each with their associated accumulation and concentration requirements), and 0.5 mm maximum edge chip depth.

Figure 5.4 provides an example of a surface imperfection notation with the visibility method. The indication `5/` is present prior to any qualifiers. This example shows a scratch of `60` and a dig of `40`, again with associated accumulation and concentration requirements. The comparison standard that is required to verify the surface imperfections is a Davidson Optronics scratch-and-dig set. This note also allows the fabricator to use an equivalent comparison standard if necessary.

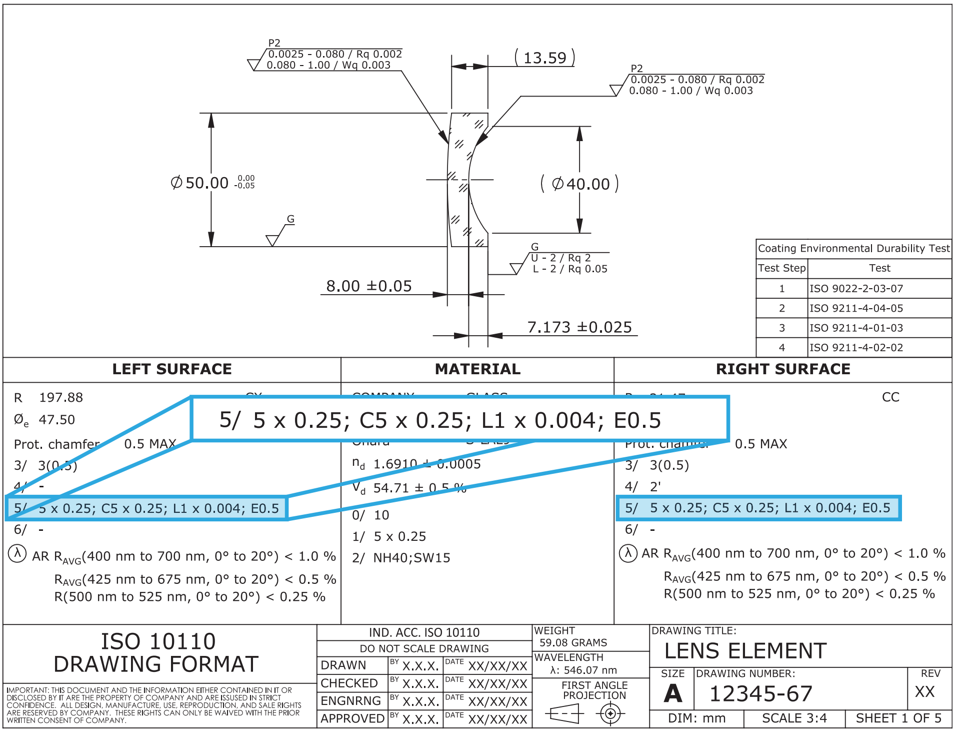


Figure 5.3 Example for dimensional method for surface imperfection specification.

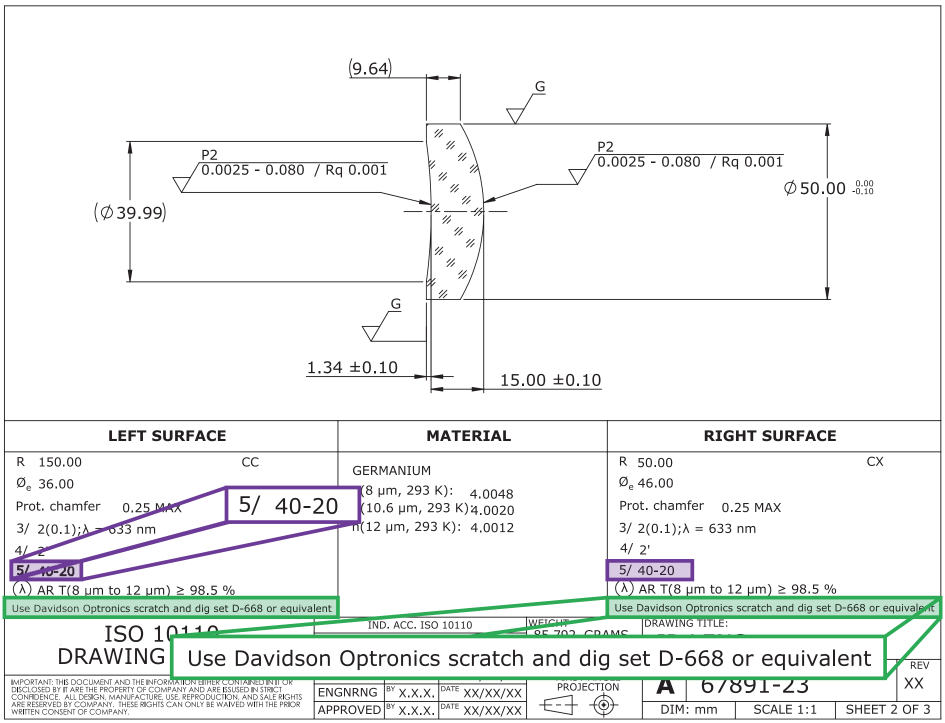


Figure 5.4 Example of visibility method for surface imperfection specification.

References

1. R. E. Fischer, "An interview with Matt Young on the scratch and dig standards for optics," *OE Rep.* **44** pp 1, 6A (1987).
2. D. M. Aikens, "Meaningful surface roughness and quality tolerances," *Proc. Int. Optical Design Conf*. **2010**, ITuD3 (2010).
3. D. M. Aikens, "The truth about scratch and dig," *Proc. Optical Fabrication and Testing*, 2010, OTuA2 (2010).
4. US Department of Defense, *Performance Specification: Optical Components for Fire Control Instruments; General Specification Governing the Manufacture, Assembly, and Inspection of*, MIL-PRF-13830B (Arlington, Virginia: 1997).
5. American National Standards Institute, *Optical Elements and Assemblies —Surface Imperfections*, OEOSC OP1.002 (New York: 2017).
6. J. H. McLeod and W. T. Sherwood, "A proposed method of specifying appearance defects of optical parts," *J. Opt. Soc. Am.* **35**(2), 136–138 (1945).
7. M. Young, "The scratch standard is only a cosmetic standard," *Proc. SPIE* **1164**, 185–190 (1989) [doi: 10.1117/12.962822].
8. I. T. Lewis, A. G. Ledebuhr, and M. L. Bernt, "Stray-light implications of scratch/dig specifications," *Proc. SPIE* **1530**, 22–34 (1991) [doi: 10. 1117/12.50493].
9. Deutsches Institute fur Normung (DIN), *Inscription of Dimensions and Tolerances for Optical Components—Surface Defects*, DIN 3140-7 (Berlin: 1978).
10. International Organization for Standardization (ISO), *Preparation of Drawings for Optical Elements and Systems—Surface Imperfection and Measurement Systems*, ISO/TR 21477 (Geneva: 2017).
11. ISO, *Test Methods for Surface Imperfections of Optical Elements*, ISO 14997 (Geneva: 2017).
12. D. M. Aikens, "New options for optical quality tolerances," *Proc. Int. Optical Design Conf*. **2017**, IM4A.3 (2017).

Chapter 6
Laser Damage

Laser damage is a major concern if an optical element or optical system involves high-power or high-energy lasers. The concern with optics in these systems is that a flaw in the optic may cause light to scatter or be absorbed and subsequently cause damage. Designation of a laser-induced damage threshold (LIDT) for these optical elements is intended to convey the message to a fabricator that additional care must be taken when manufacturing these types of optical elements or systems.

6.1 Background

For an ISO 10110 drawing, the notation for laser damage threshold is specified in ISO 10110-17.[1] Additional resources for understanding laser-damage thresholds and methods for evaluating laser-damage results are listed in the ISO 21254 series.[2–5] It is necessary to use the four parts of the ISO 21254 series with ISO 10110-17 to understanding laser-damage threshold specifications, testing conditions, and evaluation techniques.

High-power or high-energy laser light passing through an optical system may induce damage to the optical system or optical components. This is due to the laser power density or energy density—depending on the pulse duration, duty cycle, and type of laser. Laser damage will create irreversible damage due to ageing, microdamage, and induced defects.[3] The laser induced damage threshold is defined as the highest quantity of laser radiation incident on the component for which the extrapolated probability of damage is zero.[2] Due to the statistical nature of this extrapolation, the LIDT should not be considered the level below which damage will not occur; more that it is the level of laser radiation below which damage probability is less than the critical risk level.[6]

Laser damage for optics is typically a higher concern for optical surfaces rather than the bulk optical material. This is due to imperfections in the polishing of the optical surface or imperfections in coatings causing localized absorption. Typically, an optical element or system that will be used for high laser power density or energy density applications will list tighter specifications than would be used for low-power applications for each of these aspects. That is not to say that imperfections in the bulk optical material are not important. Bubbles and inclusions in the bulk optical material are still concerns, in addition to the bulk absorption coefficient. In extreme cases, laser damage sites could result in unintended light paths occurring and harming the user.

6.1.1 A word of caution

As of this writing, the laser induced damage threshold notation for ISO 10110 is in serious need of revision. ISO 10110-17 was first published in 2004 and has not been successfully revised since, in spite of (or perhaps because of) the controversy surrounding it. Prior to 2004, laser induced damage threshold was specified according to ISO 10110-13, published in 1995.[7]

ISO 10110-13 was in many ways simpler to understand; lasers were specified as either "high-energy" (short-pulse lasers) or "high-power" [long-pulse or continuous wave (CW) lasers], using the designation $\mathtt{H}_{\mathtt{th}}$ or $\mathtt{E}_{\mathtt{th}}$, respectively. For pulsed lasers the units were given as energy per unit area, and additional guidance for testing was required—such as the pulse duration, repetition rate, and number of sites to be tested as well as the number of pulses to be applied to each site. For CW or "high-power" lasers, the units were given as power per unit area, and the number of test sites to be evaluated.

In 2004, ISO 10110-13 was withdrawn and replaced by the current version of ISO 10110-17. In it, the notations of $\mathtt{H}_{\mathtt{th}}$ and $\mathtt{E}_{\mathtt{th}}$ were removed; the type of LIDT is now identified with its units, i.e., $\mathrm{J{\cdot}cm^{-2}}$ or $\mathrm{W{\cdot}cm^{-2}}$. A wavelength and a single pulse duration are also given. No additional information is to be included, such as number of test sites or pulses per site. Moreover, the units of $\mathrm{W{\cdot}cm^{-2}}$ specifically refers to pulsed power and cannot be used to indicate CW or long-pulse lasers. A new "notation" is provided, wherein the CW or long-pulse damage threshold must be given in $\mathrm{W{\cdot}cm^{-1}}$, or linear power. The thinking here was that for these types of laser systems, damage thresholds were more likely to be dependent on the spot size and total power than on the fluence.

There are many problems with the current revision. Damage testing has continued to evolve and is no longer consistent with the current notation. The lack of additional testing parameters is a significant step backward. Further, the inability to specify the CW laser-damage threshold as power per unit area is a significant problem for many laser systems. Moreover, the standard is nearly 20 years old and represents a fairly naïve view of laser-induced damage. Increasingly, damage thresholds for high-energy or high-power laser components are expressed as a probability function, rather than as a single number. Thus, ISO 10110-17 is not current with the technology or industry, and is in serious need for an update pending updates to the ISO 21254 series.

As a result, the user of this part of ISO 10110 is cautioned to be careful regarding the meaning of the LIDT indicated and the method of testing to be employed to verify compliance. It may be that this part is perfectly acceptable for their application, or it may be dangerously inappropriate.

6.2 Units

There are various mechanisms that can lead to laser-induced damage. With pulsed lasers, the laser induced damage threshold can be expressed as a peak power density ($W \cdot cm^{-2}$) or as a peak energy density ($J \cdot cm^{-2}$). If a laser emits for a period longer than 0.25 seconds then it is considered a CW laser, and in ISO 10110-17 the damage threshold is described differently.

In addition to laser type and wavelength, the physical aspects of the output from the laser have an impact on laser damage; e.g., beam diameter and beam quality. Within ISO 11145 there are two different methods of classifying a beam.[8] These can be either as the encircled power/energy or the second moment of power/energy density distribution function. The encircled power/energy is a relative measurement of the beam that contains a certain percentage of the total beam power/energy. The second moment of power/energy is the full density distribution function without a relative scaling. When specifying laser induced damage threshold, the expected energy or power distribution on the component must be considered.

6.3 Indications on Drawings

The drawing symbol for laser damage threshold is `6/`. The quantifiers following the drawing symbol vary based on the type of laser that is being specified. If the optical system is specifying for a pulsed laser the quantifiers are either:

$$6/\ H_{th};\lambda;\tau_{eff} \text{ or } 6/\ E_{th};\lambda;\tau_{eff}$$

where:

H_{th} = threshold energy density ($J \cdot cm^{-2}$),
E_{th} = threshold power density ($W \cdot cm^{-2}$),
λ = wavelength specified (nm), and
τ_{eff} = effective pulse duration (s).

Specification of the units for either energy density or power density differentiates whether H_{th} or E_{th} is specified. This notation is apparently to ensure the proper measurement is defined, and to not confuse energy or power.

In the case that the laser damage threshold specification is for a long-pulse or CW laser, the indication is similar. In that instance, the specification is:

$$6/\ \mathtt{F}_{\mathtt{th}};\lambda;\tau_{\mathtt{eff}}$$

where:
$\mathtt{F}_{\mathtt{th}}$ = threshold linear power density ($W{\cdot}cm^{-1}$).

The distinction for a pulsed laser to be a long pulse refers to when the thermal transit distance is on the same order as the test spot. This distinction is shown in Eq. (6.1). The approximate equality between the parameters is intended to show that if these terms are on the same order of magnitude, a long pulse should be considered.

$$(2D\tau_{\mathrm{eff}})^{(1/2)} \approx d_{(T,eff)} \tag{6.1}$$

where:
D = thermal diffusivity,
τ_{eff} = effective pulse duration (seconds), and
$d_{T,\ eff}$ = test spot diameter.

In the case that the time of the pulse is greater than 0.25 s, the laser damage is considered CW. For these types of lasers, the effective pulse duration is not specified. Instead, τ_{eff} becomes the exposure duration of the CW laser test.

6.4 Test Methods

ISO 10110-17 references ISO 11254-1 and -2, both of which have been withdrawn.[9,10] A schematic of the testing configuration described in ISO 11254 is shown in Fig. 6.1.

Laser damage may be performed in one of two different methods, depending on the type of laser in question and information needed. Damage testing may be done in a 1-on-1 test where a single shot occurs on a single unexposed part of the optical component in question. The other test

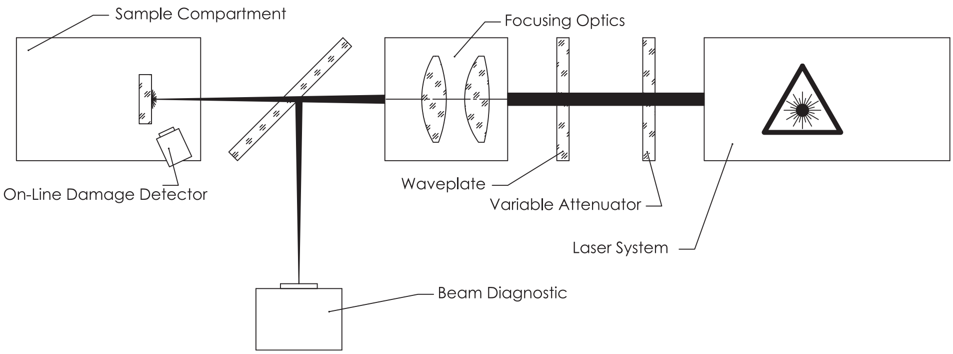


Figure 6.1 Basic approach to damage testing.

performed is an S-on-1 test where a series of pulses with a constant energy density is applied to an unexposed part of the optical component. In both of these testing situations, the number of shots and sites must be specified. While the notation for this was given in ISO 10110-13, it is not provided in ISO 10110-17. If the user desires a minimum number of sites and/or a specific number of shots per site to be tested, these must be indicated in a note.

The ISO 11254 series was withdrawn and replaced with the ISO 21254 series in 2011. This newer standard series provides much more information regarding the methods of performing damage testing and for calculating the LIDT results. In the simplest approach a series of sites are irradiated on the sample at various fluences and power densities, and each site is evaluated before and after irradiation to determine laser-induced damage. The fluences are chosen such that some sites have zero probability of damage and others have a 100% probability of damage. These results are graphed, as shown in Fig. 6.2, and the damage results are extrapolated to the point where zero damage is likely to occur.

In Fig. 6.2, the results of a 1-on-1 test are shown. Fluences range from 120 mJ, where no damage was observed; to 400 mJ, where 100% of the exposed sites showed the initiation of laser-induced damage. Between these two extremes, other fluences resulted in increasing probability of damage. A linear fit to these intermediate data points extrapolates to zero at a fluence of 185 mJ. The damage threshold can be calculated from this energy and the

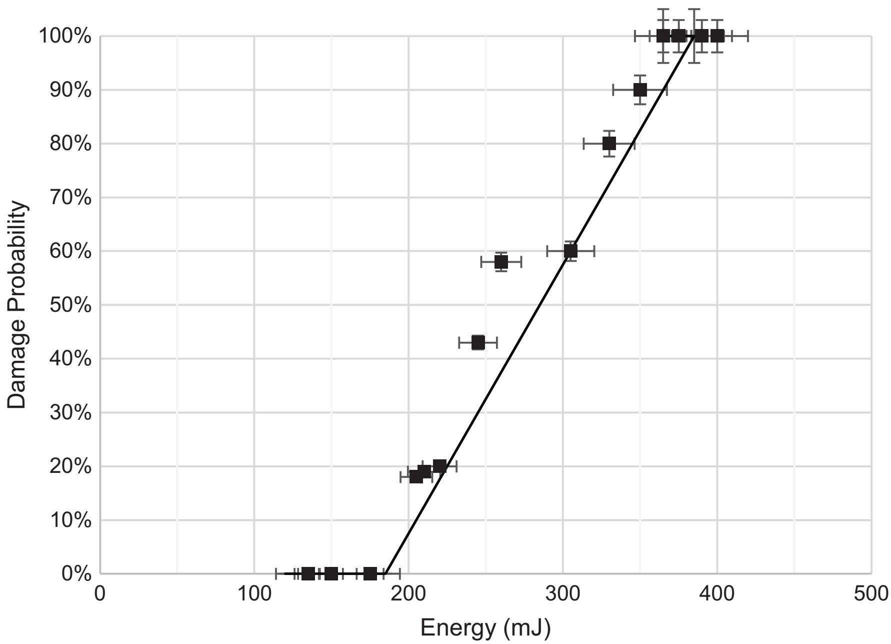


Figure 6.2 Example of a graph of LIDT testing results, showing the extrapolation of LIDT to the zero crossing at 175 mJ.

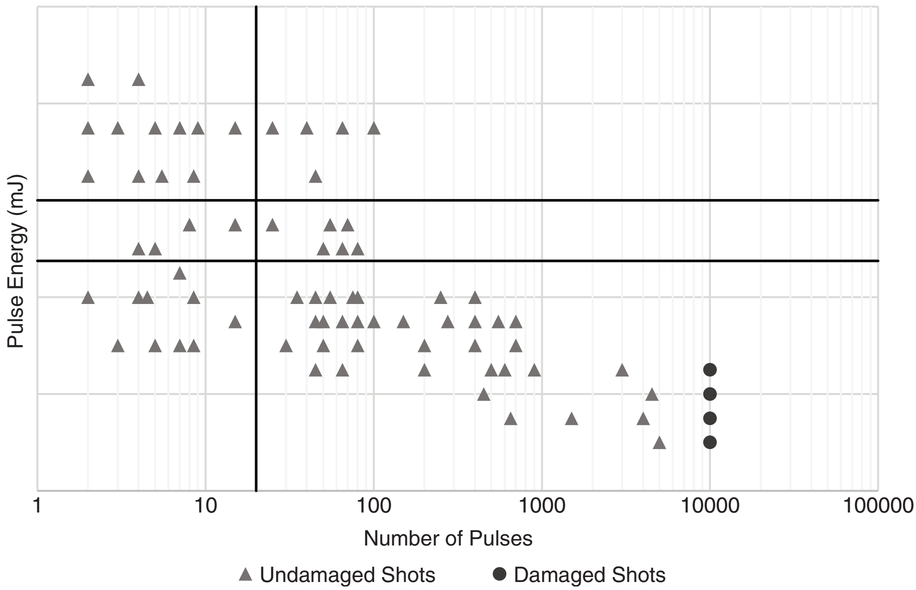


Figure 6.3 Example of damage vs. number of shots and pulse energy.

effective pulse area into the energy density in J·cm^{-2} above which laser-induced damage can be expected to occur.

For an S-on-1 test, an additional variable can be included. By varying the number of shots per site and the number of fluences tested, an array of data can be created where the probability of damage is a function of both number of shots per site and the fluence. An example of the raw data for such a test is shown in Fig. 6.3. From this example, a plot of the damage fluence for various probabilities of damage can be generated. Figure 6.4 shows an example of such a plot.

Note that at the time of this writing, the entire 21254 series is also under extensive revision. The user is cautioned to monitor publication of the new versions of these important standards.

6.5 Drawing Example

An example drawing is shown in Fig. 6.5, highlighting laser-damage-initiation threshold. As previously described, the specification for LIDT has the drawing symbol of `6/`. Of note, along with the specification of laser-damage threshold, the optical drawing lists an optical coating specifically for laser damage.

This drawing shows that the optical element has a laser-damage threshold requirement on each optical surface of 20 J·cm^{-2} for a 10 ns period using a 532 nm laser. The table on the drawing field indicates that this should be an

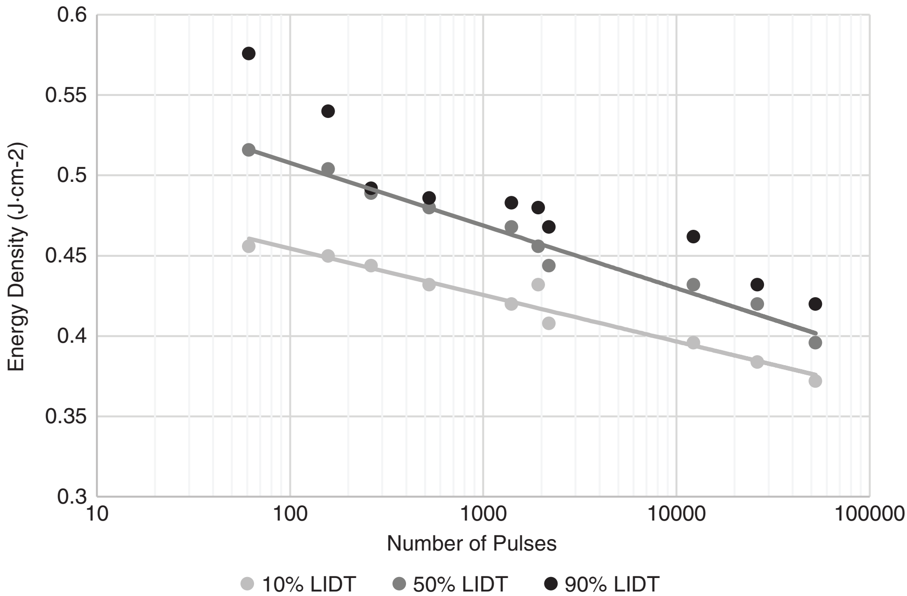


Figure 6.4 Example of fluence plotted for various probabilities of damage vs. number of shots.

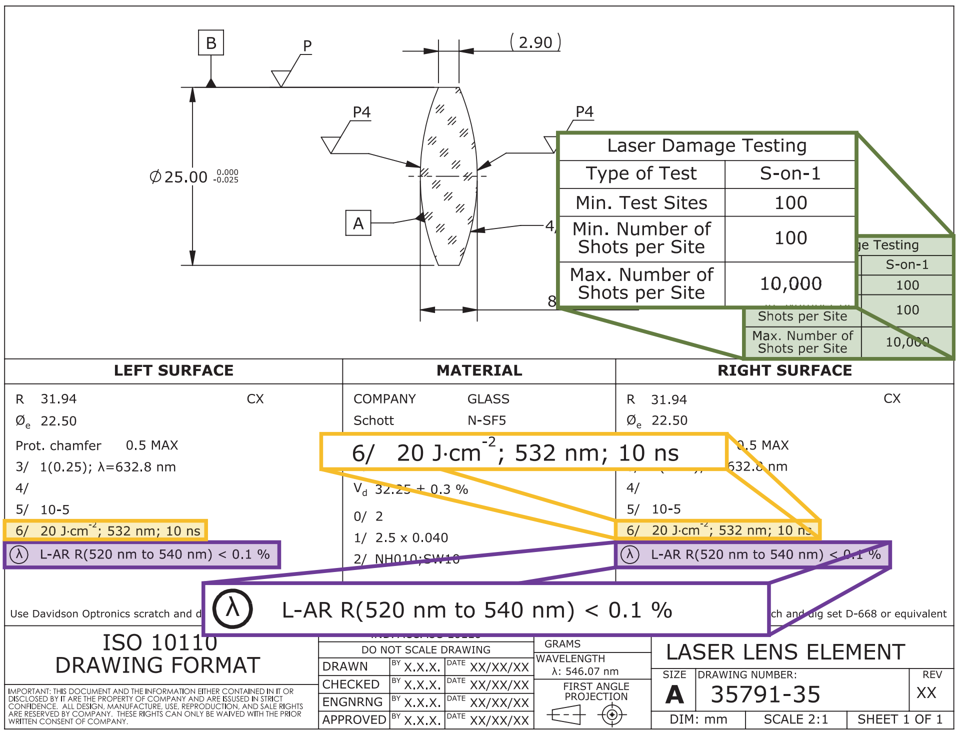


Figure 6.5 Example of an optical element ISO 10110 drawing, highlighting laser-damage threshold and a laser antireflective coating.

S-on-1 test occurring at 100 different sites. As this is an S-on-1 test, the minimum number and maximum number of shots are included.

Along with the laser-damage threshold, the antireflective coating is specified for a laser with the preceding `L-` for the notation. A deeper discussion on optical coatings and how they are specified is shown in chapter 7.

There are other aspects of this optical drawing that are indicative of a laser optic. The tight specification for a surface imperfection, the minimal number of bubbles, and the surface roughness all have a place in ensuring that the optical element or system is either not damaged or does not cause damage.

References

1. International Organization for Standardization (ISO),*Preparation of Drawings for Optical Elements and Systems—Part 17: Laser Irradiation Damage Threshold*, ISO 10110-17:2004 (Geneva: 2004).
2. ISO,*Test Methods for Laser-Induced Damage Threshold—Part 1: Definitions and General Principles*, ISO 21254-1:2011 (Geneva: 2011).
3. ISO,*Test Methods for Laser-Induced Damage Threshold—Part 2: Threshold Determiniation*, ISO 21254-2:2011 (Geneva: 2011).
4. ISO,*Test Methods for Laser-Induced Damage Threshold—Part 3: Assurance of Laser Power (Energy) Handling Capabilities*, ISO 21254-3:2011 (Geneva: 2011).
5. ISO,*Test Methods for Laser Radiation-Induced Damage Threshold: Part 4: Inspection, Detection and Measurement*, ISO 21254-4:2011 (Geneva: 2011).
6. Edmund Optics, "Understanding and specifying LIDT of laser components," accessed September 17, 2021, https://www.edmundoptics.com/knowledge-center/application-notes/lasers/understanding-and-specifying-lidt-of-laser-components.
7. ISO, *Optics and Optical Instruments—Preparation of Drawings for Optical Elements and Systems—Part 13: Laser Irradiation Damage Threshold*, ISO 10110-13:1997 (Geneva: 1997).
8. ISO, *Lasers and Laser-Related Equipment—Vocabulary and Symbols*, ISO 11145:2018 (Geneva: 2018).
9. ISO, *Determination of Laser-Induced Damage Threshold of Optical Surfaces—Part 1: 1-on-1 Test*, ISO 11254-1:2000 (Geneva: 2000).
10. ISO, *Determination of Laser-Induced Damage Threshold of Optical Surfaces—Part 2: S-on-1 Test*, ISO 11254-2:2001 (Geneva: 2001).

Chapter 7
Surface Treatment and Coating

Surface treatments and coatings are applied to an optical surface at the end of the component fabrication process. Optical coatings involve the addition of material on the polished optical surface while surface treatments are generally applied to the edges of the components. The type of optical coating depends on the functional purpose of the optical element in the application. In ISO standards, the ISO 9211 series and ISO 10110-9 work in conjunction with each other to provide detailed descriptions of optical coatings, as well as their notation for drawings.[1–9] ISO 10110-9 contains notation methods for specifying the optical coating on a drawing. While it is certainly possible to use this notation for the specification of a coating, a very detailed description of optical coatings and test methods is provided in the ISO 9211 series. Optical coatings are used to alter the transmission, reflection, or absorption properties of an optical element's surfaces. Further, these key properties of the surface may be described by wavelength regions, angles of incidence, phase, or polarization. Optical coatings are ubiquitous and a key factor for many optical drawings.

7.1 Background

Much has been written about the design, fabrication, and functionality of optical coatings without consideration for the rest of an optical system.[10–13] An optical coating is typically the only aspect of the optical fabrication process where additional material is deposited or added to an optical surface.[14] These optical coatings may be as simple as a single layer of an additional material or as complex as a multi-layer stack-up of various materials.[11] Many optical coatings are interference coatings wherein reflected or transmitted light from various layers produce different wavelength, angle, and polarization dependent constructive and destructive interference. Thus, a common function for optical coatings is to alter the transmission, reflection, or absorption for an optical element or system.

The discussion of optical coatings in ISO 10110-9 is solely the notation that should be used to specify an optical coating on a drawing. The details of optical coating specifications are covered in the ISO 9211 series. Each of the individual subsections from the ISO 9211 series are listed in Table 7.1.

The first two sections of the ISO 9211 series (ISO 9211 Parts 1 and 2) discuss the information needed to define an optical coating. The following two sections (ISO 9211 Parts 3 and 4) discuss how to specify the environmental durability and associated tests. Considerations for the durability of an optical coating are necessary to ensure that the build-up of coating layers adhere to the optical element. If the optical coating does not properly adhere to the optical element, peeling will occur on the finished optical element. Furthermore, environmental considerations can be critical to avoid coating delamination and crazing when exposed to temperature cycling and other environmental factors.

As with uncoated optical surfaces, surface imperfections are important for the final optical element. Additional considerations for uneven buildup of material are critical in ensuring that unwanted imperfections are not introduced.

The final four sections of the ISO 9211 series describe specific types of coatings. Because the coating information conveyed is narrow in scope, minimum requirements for the environmental durability are listed in each section. Additionally, ISO 9211 Parts 5, 6, and 7 contain a simplified notation for certain coatings. For example, ISO 9211 Part 5 includes a shorthand notation for antireflective (AR) coatings. Each of the simplified notations is for standard coatings, discussed further in Sec. 7.4.

In addition to functional optical coatings, there are supplementary or surface treatments that may be applied to a finished optical element. By surface treatments we mean, for example, protective paint (1) applied to a part of the optical element without optical properties, (2) used to protect an optical coating that would be susceptible to environmental effects, or (3) used to suppress stray light reflections. The notation for such surface treatments is specified in ISO 10110-9 as well.

Table 7.1 List of optical standards in the ISO 9211 optical coating series.

Standard Number	Latest Revision	Standard Name
ISO 9211-1	2018	Part 1: Definitions
ISO 9211-2	2010	Part 2: Optical properties
ISO 9211-3	2008	Part 3: Environmental durability
ISO 9211-4	2012	Part 4: Specific test methods
ISO 9211-5	2018	Part 5: Antireflecting coatings
ISO 9211-6	2018	Part 6: Reflecting coatings
ISO 9211-7	2018	Part 7: Neutral beam splitting coatings
ISO 9211-8	2018	Part 8: Coatings used for laser optics

7.2 Types of Optical Coatings

There are many types of optical coatings for optical elements. Because of the complexity in conveying the intent of these optical coatings, there is a key used to describe each coating on an ISO 10110 drawing. This description of a coating is critical for designer and fabricator communication, particularly for the fabricator to implement a manufacturable functional design. All optical coatings can be broken down to an amount of transmission, reflection, and absorption at different wavelengths and/or polarizations, and their tolerances. This description of an optical coating is shown in both Eqs. (7.1) and (7.2). The difference between these equations is the use of Greek characters or standard alphabetic terms. Both notations are acceptable on an ISO 10110 drawing. The character notation used must be consistent across the entire drawing.

$$\tau(\lambda) + \rho(\lambda) + \alpha(\lambda) = 1 \tag{7.1}$$

or

$$T(\lambda) + R(\lambda) + A(\lambda) = 1 \tag{7.2}$$

where:

τ or T = transmission at a specific wavelength,
ρ or R = reflection at a specific wavelength, and
α or A = absorption at a specific wavelength.

Understanding the purpose of each coating type is critical. Along with the purpose of the coating, it is necessary to list the transmission, reflection, or absorption properties for the coating. This description is what is listed on an ISO 10110 drawing.

In the following subsections, representative coating curves are shown. Data shown is meant to be representative of some typical categories of optical coatings. Specific details differ based on type of coating, specifications, application, fabrication process, and manufacturer. Many of the curves here have used data from Edmund Optics, Inc. without endorsement or any commercial implication; they are simply being used as examples.

7.2.1 Reflective

Reflective coatings are intended to increase the reflectance of an optical element at a specified wavelength or range of wavelengths. Reflective coatings can be broken down into three different categories: metallic, dielectric, and enhanced metallic. Metallic reflective coatings are usually made by the deposition (which can include electrochemical and dip processes) of a single metallic material to an optical surface. These types of optical coatings typically consist of aluminum, gold, or silver (Fig. 7.1).

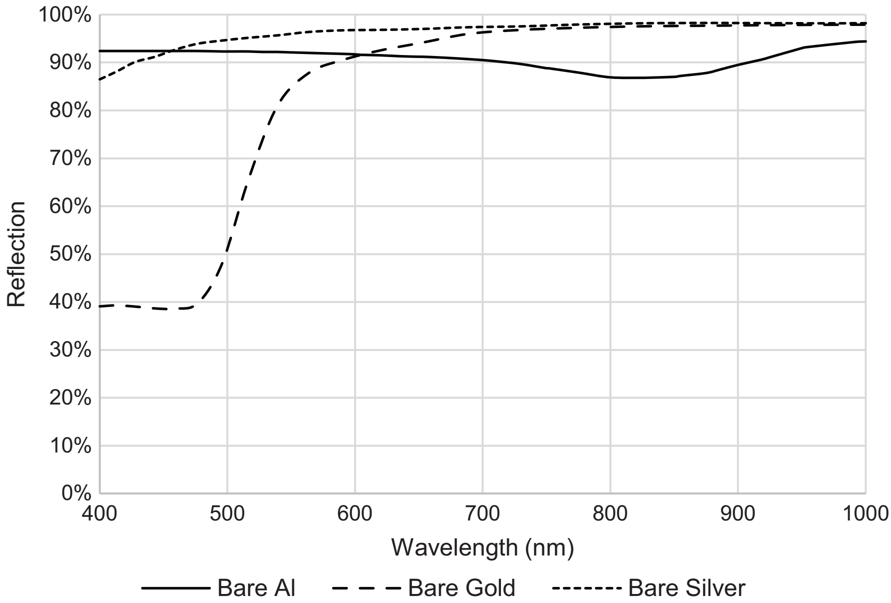


Figure 7.1 Comparison of bare aluminum, bare gold, and bare silver.[15]

Enhanced metallic coatings build on standard metallic coatings by adding additional materials on top of the metallic layer (Fig. 7.2). These coatings may be as simple as adding a protective layer on top of the metallic surface. Additional layers are meant to reduce environmental effects, such as oxidization on the bare metallic layer, while still maintaining high reflectivity. More-complex enhanced metallic coatings may use the single metallic layer for the base reflectance, but have additional layers on top to both provide an increased reflection across the wavelength range and can also provide environmental protection.

Coatings that are not built from a base metallic layer can be created and are known as dielectric reflective coatings (Fig. 7.3). These types of coatings are created with a mix of materials to create a dielectric stack of layers. Dielectric coatings can also include metallic materials in the stack. These types of dielectric coatings are used for other applications as well as a basic reflective surface. They may be used to select the wavelength range or angle of incidence being reflected.

7.2.2 Antireflective

Without an optical coating, when using a transmissive optic, Fresnel reflections will occur when light enters and exits an optical element.[16] These reflections are cumulative and reduce the total system transmission. Thus, an antireflective coating may be applied to optical surfaces to reduce Fresnel reflections.

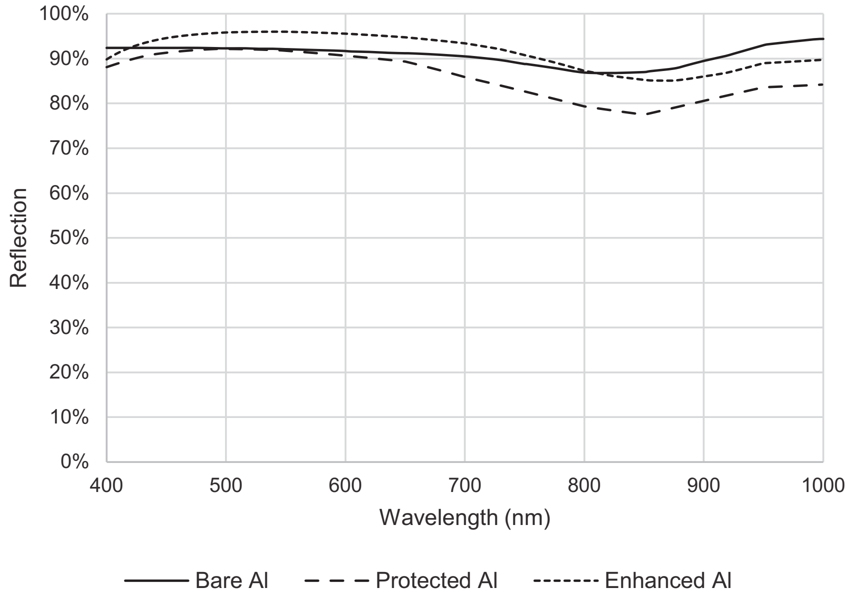


Figure 7.2 Comparison of bare aluminum, protected aluminum, and enhanced aluminum (protected and enhanced aluminum data for figure courtesy of Edmund Optics, Inc.).[15]

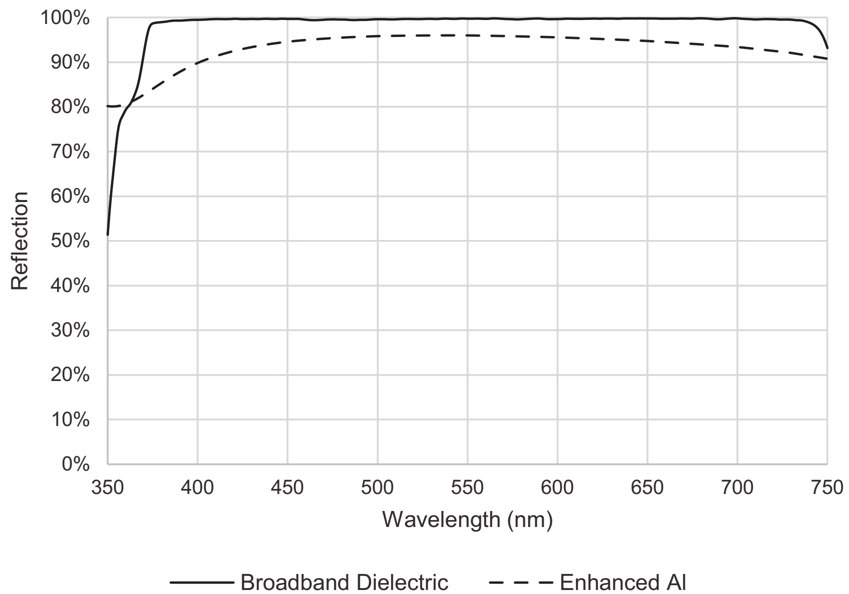


Figure 7.3 Comparison of enhanced aluminum and broadband dielectric coatings. Broadband dielectric coating is the average of the s and p polarization (coating data for figure courtesy of Edmund Optics, Inc.).

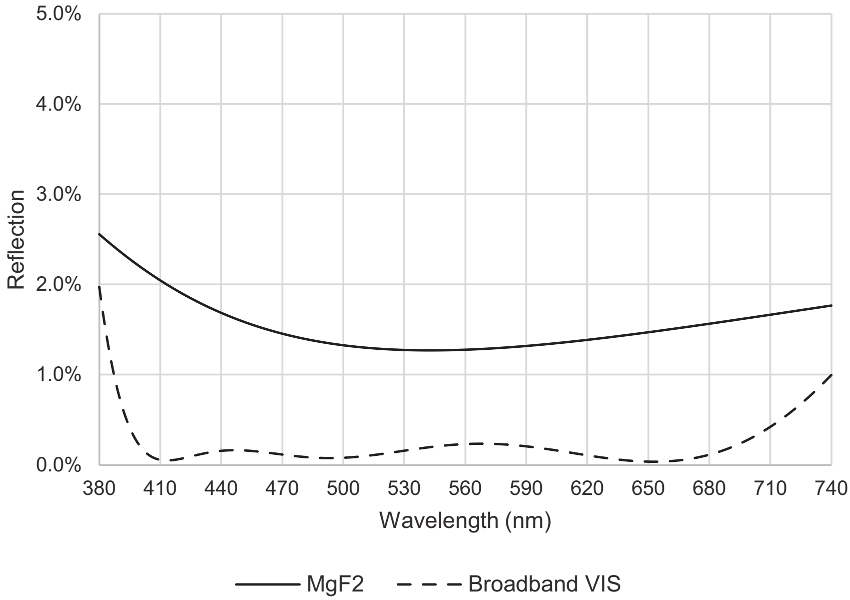


Figure 7.4 Reflection of MgF_2 and broadband visible coating (coating data for figure courtesy of Edmund Optics, Inc.).

Antireflective coatings may be designed as simple as a single layer of quarter-wave thick material onto an optical surface. The quarter-wave material and thickness are based on the primary wavelength for the intended application. For example, in the visible wavelength spectrum, typically a layer of magnesium fluoride (MgF_2) that is one-quarter of a wavelength thick at the center wavelength is deposited onto the optical surface to reduce the reflectivity of the surface (Fig. 7.4).

More complex antireflective coatings may be used to reduce Fresnel reflections at more than a primary wavelength. This is done by using multiple layers deposited to an optical surface instead of a single quarter-wave layer. The most common two types of designs are V-coats (popular in laser applications) and broadband antireflective (BBAR) designs that are often about eight layers thick (Fig. 7.4). Generally, the layers and materials in these coatings will have varying thickness instead of only quarter-wave thickness. Typically for an antireflective coating, as the number of layers increase, the amount of reflections across the wavelength range is reduced, while the cost and complexity increase.

7.2.3 Absorbing filters

Rather than transmitting or reflecting light, some optical systems require all light of a range of wavelengths to be absorbed. These types of filters will be used in situations where eliminating the light is required to reduce any

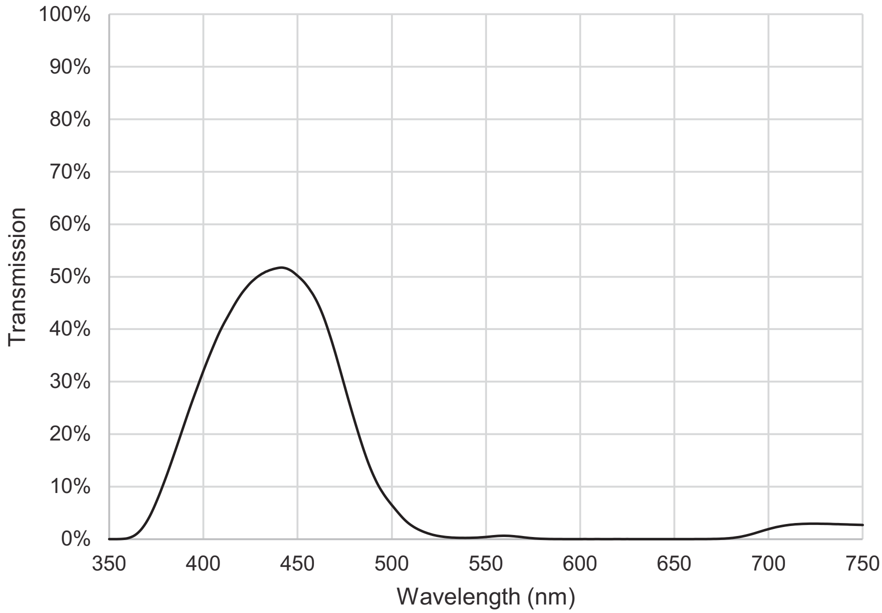


Figure 7.5 Transmission of absorbing glass filter (coating data for figure courtesy of Edmund Optics, Inc.).

potential unwanted light scatter or spectral interference. There are various types of absorbing filters. In some cases, a coating may be designed to perform this function. In other cases, a glass material may be used to absorb certain wavelengths (Fig. 7.5).

7.2.4 Attenuator

With visible light, it may be necessary to reduce the amount of light transmitted over multiple magnitudes. An attenuator coating works to reduce light across a wide wavelength range. An example of an attenuating coating is a neutral density filter (Fig. 7.6). These types of filters can reduce the transmission of visible light by orders of magnitude over a broad spectrum.

7.2.5 Beamsplitter

Combining reflective and antireflective coatings, some applications require different types or amounts of light to go down certain paths. These types of coatings may split the amount of transmitted and reflected light by wavelength or polarization (Fig. 7.7). Beamsplitter coatings typically occur on plano optical surfaces to reduce the number of varying angles of incidences. Beamsplitter coatings are common on plates and prisms.

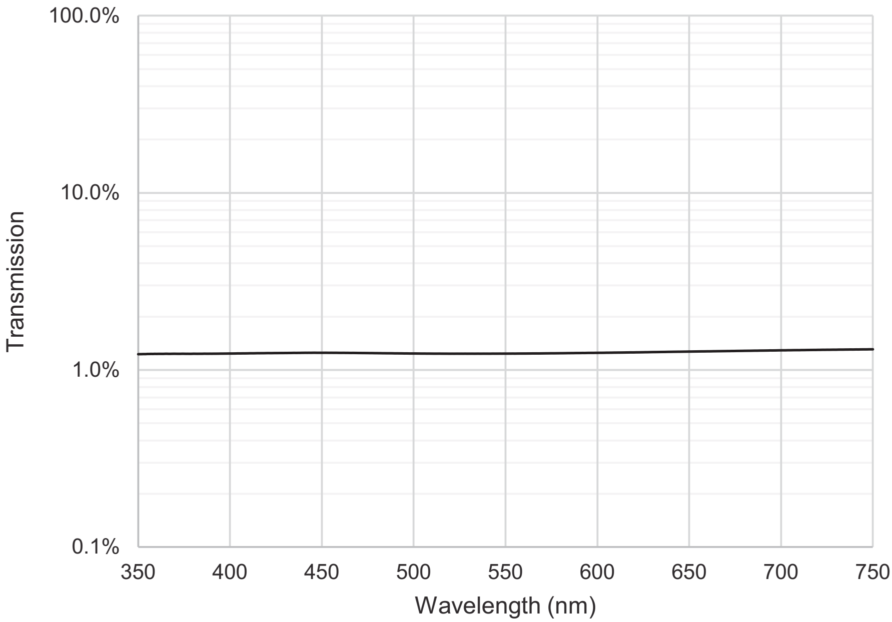


Figure 7.6 Transmission of attenuating or neutral density filter (coating data for figure courtesy of Edmund Optics, Inc.).

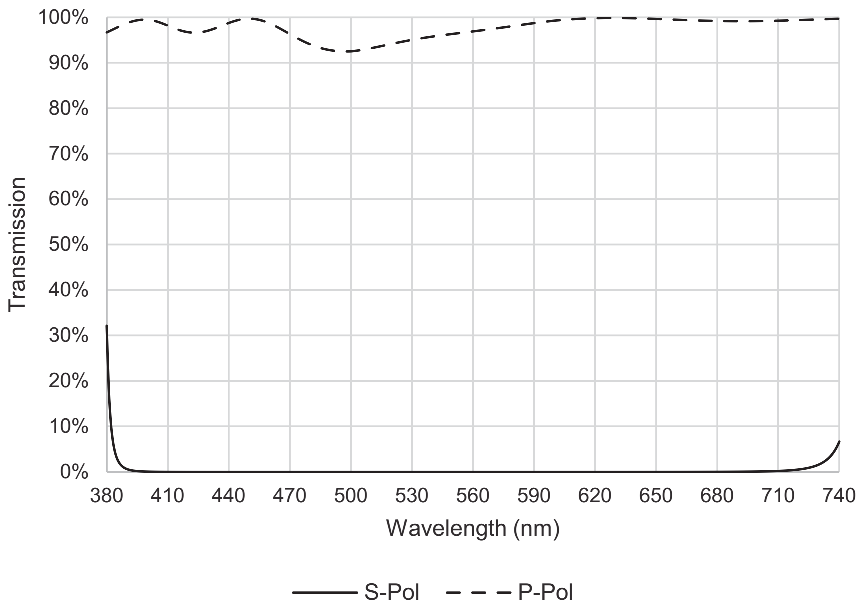


Figure 7.7 Transmission and reflection of a polarized beamsplitter (coating data for figure courtesy of Edmund Optics, Inc.).

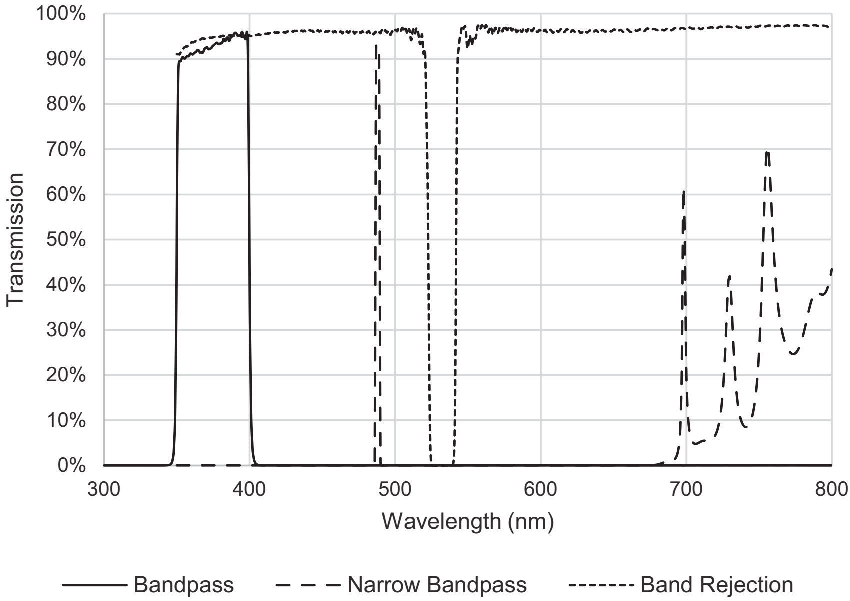


Figure 7.8 Transmission of broad bandpass filter, narrow bandpass filter, and band rejection filter (coating data for figure courtesy of Edmund Optics, Inc.).

7.2.6 Bandpass and band rejection filters

A bandpass filter will transmit light over a highly controlled wavelength region (Fig. 7.8). There will always be a minimum and a maximum wavelength for the transmission range. The light outside of the wavelength band in question will be either reflected or absorbed. These types of filters require multiple layers to cleanly transmit and block light as required. Typically, the more layers in the coating stack-up, the cleaner the transition (higher filter quality factor) will be between wavelengths blocked to transmitted to blocked again.

In contrast to bandpass filters, in certain cases, blocking a single wavelength range may be necessary. A band rejection filter will block a highly controlled wavelength region from transmitting through the optical element (Fig. 7.8). Band rejections will typically work over a wide wavelength region to keep the entire waveband transmitting, except for the rejected region.

7.2.7 Long and short pass filters

Optical coatings can also be designed to transmit long wavelengths (a long pass filter) or short wavelengths (a short pass filter) [Fig. 7.9]. Note the definition of long and short pass is thus with respect to the wavelength, versus many applications of filtering seen in analog circuits—which describe high and low pass filters in frequency space.

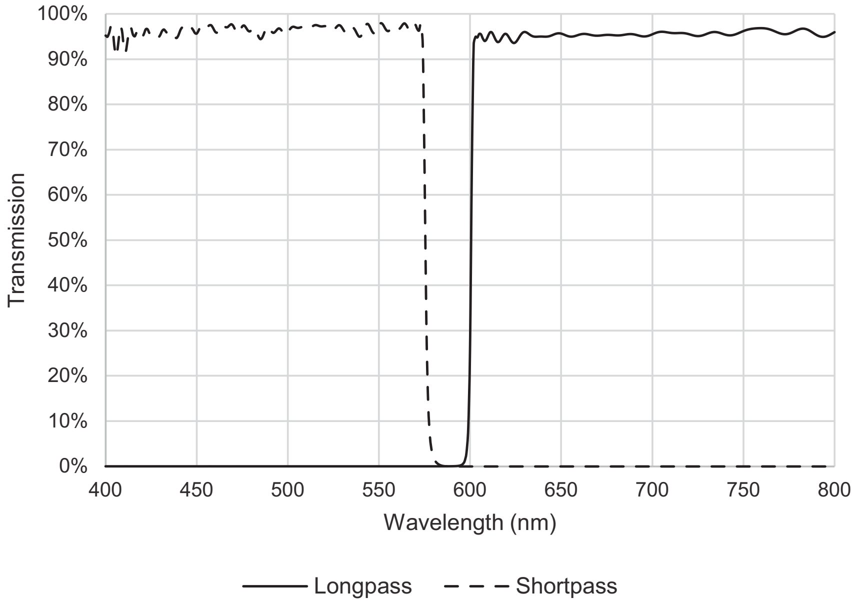


Figure 7.9 Transmission of long and short pass filters (coating data for figure courtesy of Edmund Optics, Inc.).

A long pass filter may also be termed a cold mirror in thermal applications, notably if used in the visible wavelength range. A cold mirror will reflect short or blue wavelengths and transmit long or red wavelengths. Likewise, a short pass filter may also be termed a hot mirror in thermal applications, notably if used in the visible wavelength range. A hot mirror will transmit short or blue wavelengths and reflect long or red wavelengths.

7.2.8 Polarizer

Of the previous filter coatings, the filtering is dependent on the wavelength region over which the tolerances are defined. Another type of filter will transmit or block light based on the polarization orientation (Fig. 7.10). These types of filters will still operate over a wavelength range, but their primary function is to filter based on the electric field polarization state.

7.2.9 Phase changing

Phase-changing filters are based on polarization orientations. These types of filters will transmit, block, or change the phase of a polarization orientation. Phase-changing filters are typically used in conjunction with a polarizer to cleanly transmit a specific orientation and block all others.

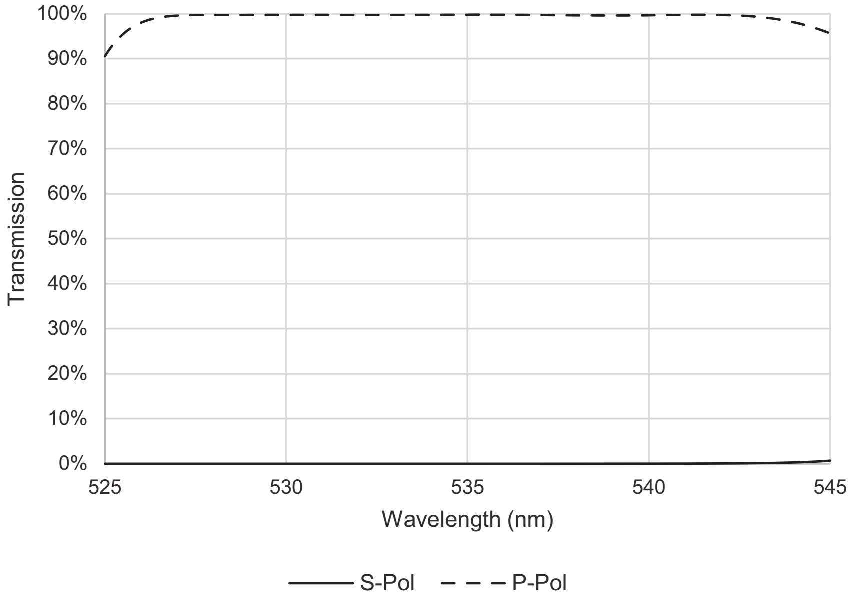


Figure 7.10 Transmission of dielectric polarization coating. This coating is designed to polarize primarily for 532 nm (coating data for figure courtesy of Edmund Optics, Inc.).

7.2.10 Laser optic

For optical systems that rely on high-power lasers to perform their intended function, the necessary large amounts of power may damage the optical elements. The coatings on these optical elements require special protective coatings to mitigate unwanted light and not destroy the optical elements in question. Laser-specific optical coatings will primarily be used for antireflective or reflective coatings (Fig. 7.11). Often, having a laser being absorbed onto an optical element will continually heat up the absorbing optical element, potentially causing damage to the coating.

7.3 Indications on Drawings

Specifying an optical coating on an ISO 10110 drawing is different than other aspects of the optical tolerances previously discussed. Unlike other optical properties specified with numeric drawing symbols, optical coatings are indicated with an encircled lambda symbol; i.e., Ⓛ as the drawing symbol. The quantifier for an optical coating may be done in many ways. These methods are all based on the type of optical coating and the tolerance. In comparison to the numeric drawing symbol quantifier, the dash (–) that is used for other drawing indicators is not necessary for an optical coating. There is no default set of tolerances for optical coatings (ISO 10110-11 does not cover this topic).

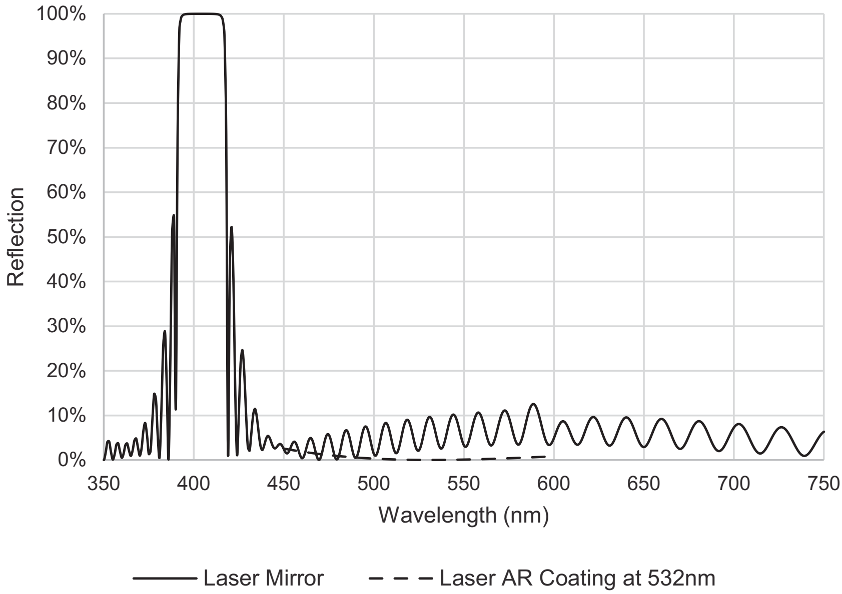


Figure 7.11 Laser optic designed coating for both a mirror and an antireflective coating. (Coating data for figure courtesy of Edmund Optics, Inc.)

Thus, no optical coating indicator on a drawing indicates that the surfaces are uncoated.

The indication for an optical coating may be done in either the tabular field for the associated surface or directly in the drawing field with the encircled lambda tangent to the associated surface. This is similar to the numeric drawing symbols on a drawing using the tabular or nontabular view.

The full indicator for an optical coating specification contains multiple quantifiers that may all be required to specify the requirements. The basic structure for the quantifier is consistent across all types of optical coatings. This structure is as follows:

`code (lower limit term) [< or ≤] (spectral optical property term) [< or ≤] (upper limit term)`

Each of the terms have overlapping notation consistent for all types of optical coatings. It may or may not be necessary to include both the lower and upper limits on a drawing, depending on the optical coating requirements.

The notation for drawing indications as described in Sec. 7.2 is shown in Table 7.2. Within each type of coating, the many spectral optical properties are uniform. Certain spectral optical properties are intended for specific types of optical coatings but are not required for said coatings.

For each type of coating, the upper limit, lower limit, and spectral optical property may be defined by the transmission, reflection, or absorption. The notation for spectral optical properties primarily uses the notation shown in Eqs. (7.1) or (7.2). Additional notation is used for change in phase ($\Delta\phi$ or $\delta\phi$)

Table 7.2 Description of functional coatings and their drawing indications.

Type of Coating	Drawing Indication
Reflective	`RE`
Antireflective	`AR`
Absorbing	`AB`
Attenuating	`AT`
Beamsplitter	`BS`
Filter: Bandpass	`FI-BP`
Filter: Band Rejection	`FI-PR`
Filter: Long Pass	`SC-LP`
Filter: Short Pass	`SC-SP`
Polarizing	`PO`
Phase Changing	`PC`
Laser Optic Antireflective	`L-AR`
Laser Optic High Reflective	`L-HR`

or optical density (D). The optical density is used in cases where the difference in transmission, reflection, or absorption occurs across multiple magnitudes. The optical density notation relates to the transmission, reflection, or absorption; as shown in Eq. (7.3):

$$D(\lambda) = -\log[X(\lambda)] \tag{7.3}$$

where:
D = density of the optical parameter,
λ = wavelength of light being specified in the spectral optical property, and
X = a placeholder term that can be represented by a transmission, a reflection, an absorption, a phase, or a density symbol.

The spectral optical property to be specified can either be defined for the average over the spectral range, or be absolute for every wavelength over the spectral range. Furthermore, the average or absolute property can be specified for the `s` and/or `p` polarizations, or for unpolarized light. If neither `s` or `p` is indicated, the specification is assumed to be for unpolarized light. The notation associated with an optical property is shown in Table 7.3. The `X` as shown in Table 7.3 is a placeholder for the transmission, reflection, absorption, phase, or density.

Most optical coatings will be defined for a constant transmission, reflection, or absorption. In some cases, it is necessary to specify an optical coating linearly changing over a wavelength range. In the case that a linearly changing value is used, an arrow (→) between the terms is used to designate the starting and ending limits. This type of tolerance notation is shown in Fig. 7.12.

Table 7.3 Notation used to describe an optical property.

Type of Optical Property	Optical Property Notation
Unpolarized Absolute	X
Unpolarized Average*	X_{ave}
Polarization (S), Absolute	X_s
Polarization (P), Absolute	X_p
Polarization (S), Average	$X_{s,ave}$
Polarization (P), Average	$X_{p,ave}$

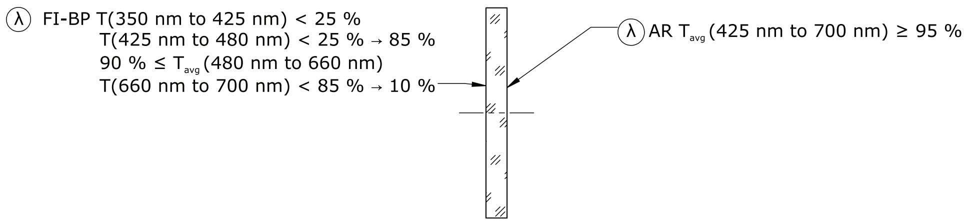


Figure 7.12 Notation for a bandpass filter coating on the left surface and an antireflective coating on the right surface. The bandpass filter contains two linear changes across the two different wavelength bands.

In some circumstances it may be best to describe the optical coating graphically. In this instance the coating curve may be provided, showing a graphical representation of what would be found using the notation description. When describing an optical coating graphically, there must be explicit callouts where the wavelengths are and the optical property being described. Describing what the upper and/or lower limits are for the optical coating should be done with either a notation of a `U` and an `L`, or arrows pointing in the direction of the tolerance limit. When using a graphical representation of an optical coating, the notation shown on the curve relates only to absolute values. If the coating has an average optical property, that must be listed separately with a written notation. Figure 7.13 graphically describes an example of an optical coating specification.

The written notation for the optical coating in Fig. 7.13 would be*

```
FI-BP  T(302.5 nm to 340 nm) <5%
       87.5% <T(352.5 nm to 397.5 nm)
       T(407.5 nm to 497.5 nm) <5%
```

*The subscript specifying the average of an optical property may be defined differently by country. ISO 9211-2 uses “`ave` for average. Other parts of the world use “`avg` to specify average.

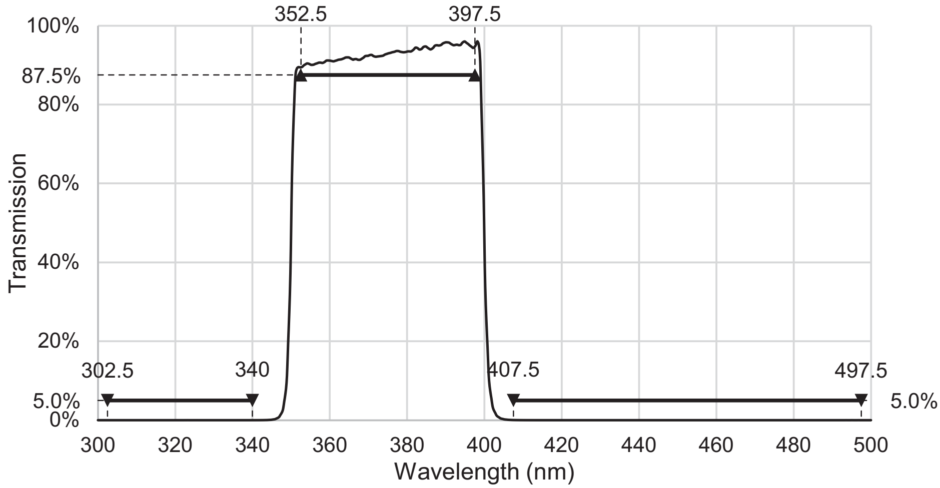


Figure 7.13 Notation for a graphic representation of a bandpass filter where the transmission band and rejection band are noted by the wavelength and transmission percentage (coating data for figure courtesy of Edmund Optics, Inc.).

7.4 Standard Coatings

As the ISO standards for optical coatings continue to evolve, three of the ISO 9211 standards (Parts 5, 6, and 7 for antireflective, reflective, and beamsplitter coatings respectively) specify standard optical coatings that may be used. Antireflective coatings are simply identified as a single layer (`U`), double layer (`V`), multilayer broadband (`W`), and other (`X`). The first three coatings [`U`, `V`, and `W`] are basic types of antireflective coatings that are named based on their reflectance curve shape. The `U` and `V` coatings are specified for a single wavelength. The `W` multilayer broadband coating is meant for a range of wavelengths. The reflectance designations for these three types of coatings are listed in Table 7.4. Lastly, an `X` coating type is a multilayer coating that does not have reflectance properties as described by `U`, `V`, or `W` coatings.

The coatings described for reflective surfaces are based on the metallic material that should be used. These specifications are shown in Table 7.5.

Table 7.4 Coded notation for antireflective coatings associated with ISO 9211-5.[5]

Coating Type	Refractive Index	Wavelength	Reflection
U	1.45 to 1.5	λ	$\leq$2%
	1.5 to 1.7	λ	$\leq$1.5%
	$\geq$1.7	λ	$\leq$1%
V	–	λ	$\leq$0.2%
W	–	λ_1 to λ_2 where $\frac{\lambda_2}{\lambda_1} \geq 1.57$	$\leq$0.5%

Table 7.5 Coded notation for reflective optical coatings associated with specific metallic materials.[6]

Coating Type	Spectral Range	Reflection
Al	400 nm to 600 nm	≥89%
	600 nm to 780 nm	≥86%
Al-GS	400 nm to 440 nm	≥82%
	440 nm to 620 nm	≥86%
	620 nm to 700 nm	≥82%
Al-RE	420 nm to 500 nm	≥88%
	500 nm to 600 nm	≥93%
	600 nm to 700 nm	≥88%
Ag-R	400 nm to 440 nm	≥90%
	440 nm to 500 nm	≥94%
	500 nm to 780 nm	≥96%

Table 7.6 Coded notation for beamsplitter optical coatings.[7]

Coating Type	Wavelength	Reflection	Transmission
D1	450 nm to 650 nm	50% ±5%	50% ±5%
D2	450 nm to 650 nm	70% ±5%	30% ±5%
D3	450 nm to 650 nm	20% ±5%	80% ±5%
D4	400 nm to 700 nm	50% ±3%	50% ±3%
M1	380 nm to 780 nm	30% ±5%	30% ±5%
M2	450 nm to 700 nm	45% ±5%	45% ±5%

Note that only three types of aluminum coating and one type of silver coating are defined. There is no standard optical coating for bare silver or for any type of gold coating.

Beamsplitter coatings can also be defined by a single coded notation. The definition for beamsplitter coatings is generally a balance between the transmission and reflection over a given wavelength range. Two of the coded optical beamsplitter coatings have some absorption based on their specifications. All of specifications for the coded beamsplitter coatings are shown in Table 7.6.

7.5 Surface Treatments

A surface treatment, meaning one that does not have a critical-to-function optical purpose, may be applied to a surface on an optical element. In general, this may be a paint or electronic coating. A paint may be applied to an optical element for either decorative or protective purposes. An oxidized coating on the edges of the lens may be used to electrically isolate an optical element in a complete system.

When specifying these types of surface treatments, the optical coating symbol ⓐ is not used and the treatments are instead noted on the drawing

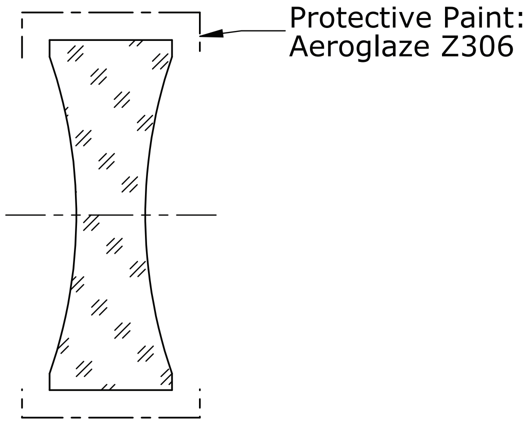


Figure 7.14 Example of a double concave lens with Aeroglaze Z306 protective paint applied to the cylindrical edge and annulus surfaces.

field. The specification is attached to a chain line (dash–dot–dash) outlining the surface or region of the optical element that will be coated. Figure 7.14 shows an example of an optical element with a protective paint on the two annulus faces and the outer diameter or cylindrical edge surface.

7.6 Testing and Durability

The completion of an optical coating will require testing for both optical and mechanical properties. Verifying the transmission, reflection, or absorption is typically performed with a spectrophotometer. A spectrophotometer can test in either transmission or reflection configurations, depending on the setup. By using Eqs. (7.1) or (7.2), the measurement of the amount of light found in transmission and reflection across wavelengths can be used to determine the amount of light being absorbed.

There are many different environmental and durability tests that can be performed on optical elements. These tests ensure the mechanical function use of the optical element. With environmental testing, a few test conditions are important for optical coatings. Particularly, various temperature and humidity cycles can cause stress in an optical coating. These stresses may cause problems for the adhesion of the coating to the substrate.

To ensure the environmental tests do not compromise the optical elements in question, use of witness samples is generally used. Since it is expected that the coatings of all of the components in a single coating run will be affected in the same way, witness samples can be used as a stand-in for the optical elements for testing optical and mechanical functions of the coating. Witness samples are ideally the same material as the optical element itself with an equivalent polish; the witness sample might be a sacrificial component in the run but is more often a part with no optical function, but the purpose of which is solely to be tested in the conditions the optical coating will face. In the case that one of the tests done to a witness sample becomes destructive, it is possible to then remove the coating from the optical elements in the run and subsequently recoat the parts without needing to start from the beginning of fabrication.

Even though optical elements may not experience rapid environmental changes, rapid cycle environmental changes can simulate the lifetime of an optical coating in a short time-period.

7.6.1 Durability testing indications on drawings

For specifying coating environmental durability on an ISO 10110 drawing, an environmental test code is necessary. Each coating test code has multiple parts, stating the type of test and the conditions under which the coating is to be tested.

The codes for coating durability are different than the tolerance notation for a normal ISO 10110 drawing. The test code for the coating durability is based on a single condition under which a test will occur. Multiple coating durability tests may be specified using the same notation. The coating durability notation below is similar to those for all environmental conditions, as will be discussed further in Chapter 11, but without the state of operation code, which would not apply to a component.[17] As there is overlap between typical environmental test conditions and coating environmental durability testing, some of the notations may appear the same. In the event that it is important to conduct the various tests in a specific order, the durability tests shall be numbered. An example of durability tests in a specific order can be found in Figure 7.16

`Coating environmental durability test ISO 9XXX -Y -AA -BB`

where:

`ISO 9XXX` = specific ISO standard that should be referenced; e.g., 9211 or 9022;
`Y` = part of the respective ISO standard previously listed;
`AA` = conditioning code; that is, the test to be performed; and
`BB` = severity level for the given durability test.

The specific ISO standards that will be referenced for coating environmental durability tests come from either the ISO 9022 series for environmental conditions or from ISO 9211-4 for specific test methods only used on optical coating durability.[4] As it may be confusing which environmental tests should be listed for a coating environmental durability, Table 1 in ISO 9211-3 provides guidance on the parts from the ISO 9022 series that should be referenced.[3] Within ISO 9211-3 there are seven different parts from the ISO 9022 series that could be specified. The conditioning methods specifically for coating durability are all listed in ISO 9211-4.

The conditioning code listed states the specific test to be performed. Each of the conditioning codes will also have their own degree of severity level. These are the levels at which the test should be performed. An example of a coating environmental durability test code is listed as follows:

`Coating environmental durability test ISO 9211 -4 -01 -03`

This coating specification is referencing `ISO 9211-4` for the test information. The explicit test that will be done is conditioning code `01`, abrasion. As there are different levels of abrasion that can be performed, described more in Sec. 7.6.3, the severity level is listed as `03`, eraser test (severe abrasion).

Since there are environmental tests that can be performed with a wider range of degrees of severity than necessary for coatings, ISO 9211-3 lists each of the tests used specifically for coating durability. In Table 1 of ISO 9211-3, the degrees of severity for each conditioning method applicable for a coating is stated along with a description of those test conditions.

7.6.2 Environmental changes

The main environmental effects that can compromise an optical coating are the influences of temperature, humidity, solubility or immersion, and exposure to solvents. The different temperature or humidity changes or introduction of water or chemical solvents may cause either surface damage or weakening of the coating adhesion.

Temperature and Humidity

Thermal change for an optical element, going from either hot to cold or vice versa over multiple cycles, simulates an environmental condition for the optical element being outdoors and being exposed to different seasons. The inclusion of either a dry or humid environment exposes the optical element to different conditions based on location.

When an optical element is exposed to hot or cold temperatures with humidity, moisture seeping into the optical coating may cause the optical coating to break up and start to separate from the optical element. In the case that there is no humidity in the atmosphere, the optical coating may have other kinds of adhesion damage. These conditions can happen with exposure to temperatures as cold as $-35°C$ and as hot as $+70°C$. From those two extremes in temperature, the optical element may be exposed to dry environments with humidity less than 40% or as humid as up to 95%. Covering extremes for variations of temperature and humidity simulates possible environmental conditions that an optical element would be exposed to in severe uncontrolled outdoor applications.

Exposing an optical coating to these types of environments may create a crazing effect. This will create a pattern of cracks in the optical coating. These cracks may then lower the adhesion of the optical coating to the optical element.

Solubility

In addition to temperature and humidity, optical elements may also be subjected to submersion in a liquid. Submersion in this case refers to either a water solution (salt water or distilled water, either boiling or at 23°C), a chemical corrosive, or a solvent.

The environmental testing of an optical element in a corrosive or a solvent is specified to ensure that exposing an optical coating to chemicals (e.g., acetone or ethanol) will not have an effect on the adhesion of the coating.

7.6.3 Adhesion and abrasion test methods

Abrasion and adhesion tests are used specifically to ensure the coating mechanical integrity and coating adhesion is acceptable. These tests can be applied either before or after environmental testing. The two tests, either abrasion or adhesion, are performed on the optical coating; and then the coating is visually inspected to ensure no changes have occurred.

Both tests (abrasion and adhesion) use relatively simple materials for their test; specifically, abrasion testing uses either cheesecloth or an eraser, while adhesion testing uses a piece of adhesive tape.

Abrasion

Abrasion testing of an optical coating involves rubbing a tester against the surface and then inspecting the sample to ensure that no damage has been imparted to the optical coating.

An abrasion test is performed with one of two testers, depending on the required test level. The optical coating is rubbed with either a piece of cheesecloth or with a rubber eraser. Both materials will be abrasive to the optical coating and will therefore test the mechanical integrity of the coating at a single zone of the optical surface. Selection of either the cheesecloth or the eraser will depend on the severity required for a durability test. Cheesecloth is the less-severe testing method for abrasion.

Adhesion

To test the adhesion of a coating to an optical surface, a tape test is performed. In the tape test, a piece of standard adhesive transparent tape is adhered to an optical coating and then pulled off from the surface. This type of test will typically be performed at various removal speeds to test the severity.

Testing an optical coating for adhesion may be performed at either the center of the optical element or at the edge. If the optical coating adhesion is unsatisfactory, then either a flaking (center of the optical element) or peeling (edge of the optical element) will occur.

7.7 Drawing Example

An example drawing highlighting the optical coatings is shown in Fig. 7.15, which shows an antireflective optical coating specification where the absolute transmission of the coatings across the wavelength region from 8 μm to 12 μm must be greater than or equal to 98.5%. Further examples of different specifications can be found in the ISO 10110 series, including Part 1.

An additional example drawing highlighting the optical coatings is shown in Fig. 7.16—which shows an antireflective optical coating specification with average reflection for a wide wavelength region less than 1%, a slightly less wide wavelength region less than 0.5%, and an absolute reflection for a narrow wavelength region less than 0.25%. This optical coating specification is listed for an angle of incidence range on the coating. Along with listing the multiple optical coating specifications, this example also shows the coating

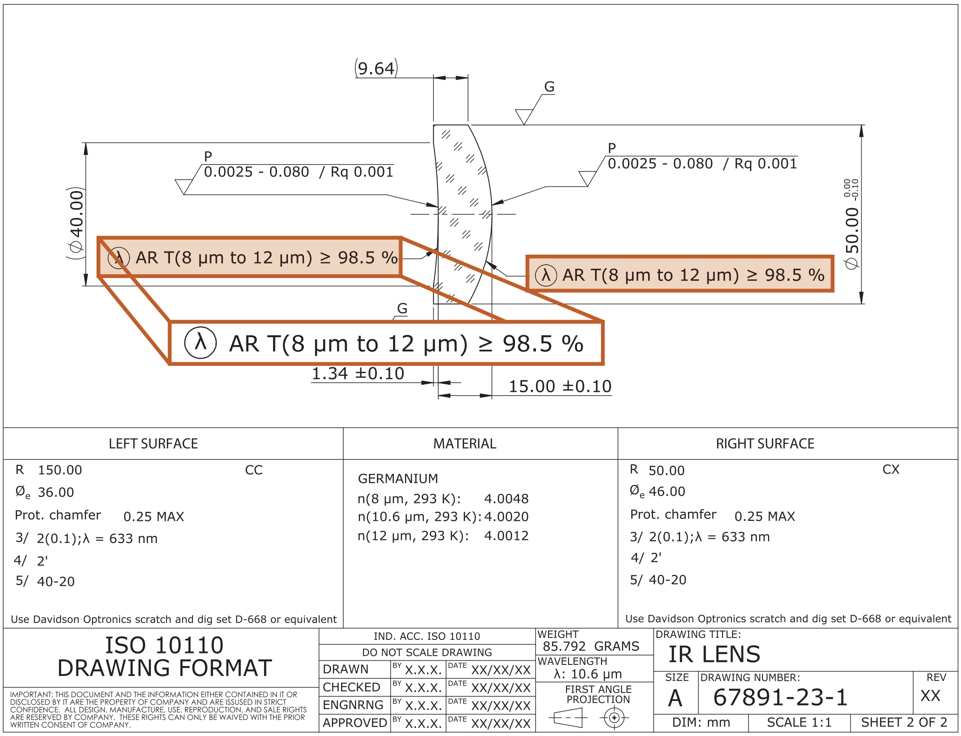


Figure 7.15 Example drawing with an antireflective coating on both the left and right surfaces.

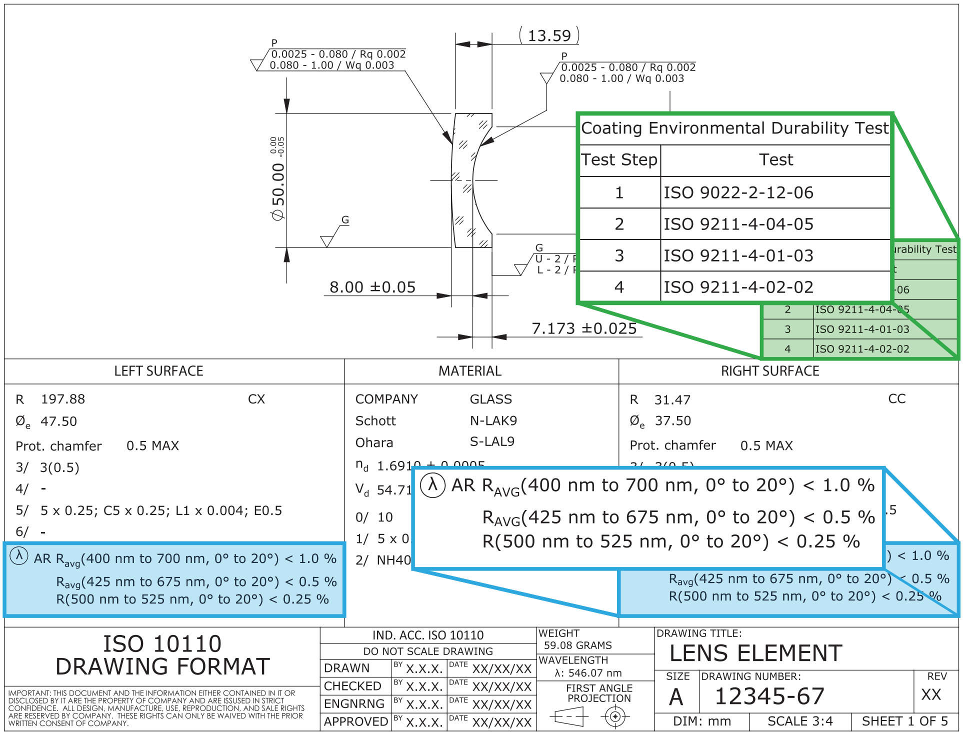


Figure 7.16 Example drawing with an antireflective coating on both the left and right surfaces, for specific wavelength ranges and angles of incidence. This drawing also shows coating environmental durability tests that should be performed.

durability tests that should be performed. In this instance, the coating environmental durability tests will be conducted in the order as specified in the test step.

References

1. International Organization for Standardization (ISO), *Optical Coatings—Part 1: Definitions*, ISO 9211-1:2018 (Geneva: 2018).
2. ISO, *Optical Coatings—Part 2: Optical Properties*, ISO 9211-2:2010 (Geneva: 2010).
3. ISO, *Optical Coatings—Part 3: Environmental Durability*, ISO 9211-3:2008 (Geneva: 2008).
4. ISO, *Optical Coatings—Part 4: Specific Test Methods*, ISO 9211-4:2012 (Geneva: 2012).
5. ISO, *Optical Coatings—Part 5: Minimum Requirements for Antireflecting Coatings*, ISO 9211-5:2018 (Geneva: 2018).
6. ISO, *Optical Coatings—Part 6: Minimum Requirements for Reflecting Coatings*, ISO 9211-6:2018 (Geneva: 2018).

7. ISO, *Optical Coatings—Part 7: Minimum Requirements for Neutral Beam Splitter Coatings*, ISO 9211-7:2018 (Geneva: 2018).
8. ISO, *Optical Coatings—Part 8: Minimum Requirements for Coatings Used for Laser Optics*, ISO 9211-8:2018 (Geneva: 2018).
9. ISO, *Preparation of Drawings for Optical Elements and Systems—Part 9: Surface Treatment and Coating*, ISO 10110-9:2016 (Geneva: 2016).
10. J. D. Rancourt, *Optical Thin Films: User Handbook*, SPIE (Bellingham, Washington: 1996).
11. A. Macleod, *Thin-Film Optical Filters*, 5th ed., CRC (Boca Raton, Florida: 2017).
12. R. R. Willey, *Field Guide to Optical Thin Films*, 1st ed., SPIE (Bellingham, Washington: 2006).
13. J. A. Dobrowolski, "Optical Properties of Films and Coatings," in *Handbook of Optics 1*, Chapter 42, 2nd ed., McGraw–Hill (New York: 1995).
14. R. Williamson, *Field Guide to Optical Fabrication*, SPIE (Bellingham, Washington: 2011).
15. E. D. Palik, *Handbook of Optical Constants of Solids*, 1st ed., Academic (San Diego: 1998).
16. J. E. Greivenkamp, *Field Guide to Geometrical Optics*, SPIE (Bellingham, Washington: 2003).
17. ISO, *Optics and Photonics—Environmental Test Methods*, ISO 9022:2020 (Geneva: 2020).

Chapter 8
Centering and Tilt Tolerances

Optical centering is the alignment of optical surfaces or elements with respect to a datum or datum system. The datum may either be a point, an axis, a surface, or a cylinder. The method of describing optical centering is unlike other sections of ISO 10110 because optical centering relies heavily on mechanical engineering methods and geometrical dimensioning and tolerancing (GD&T). ISO 10110-6 specifies surface tilts, surface decenters, runouts, and beam deviations that can be used to tolerance the centering of an optical surface, an optical element, or a subsystem.[1]

8.1 Background

Centering tolerances for an optical element are applied near the final stages of fabrication.[2] The process of centering relates the surface normal of the optical surfaces to a datum, which is frequently an axis defined by a cylindrical surface (in many cases, the outer diameter). Throughout the remainder of this chapter, all outer diameter surfaces will be referenced as a cylindrical edge surface. The centering tolerances of optical surfaces reference datums or datum systems by a tilt, or a tilt and a decenter. There are multiple ways to describe the surface tilt and/or decenter because of the different methods by which optical surfaces are manufactured and tested. Although referred to as centering, it is most often specified in the form of a tilt angle of the surface normal with respect to some reference axis, or a datum axis. The definition of the datum or datum system is critical to the centering tolerance and how the tolerance should be interpreted. An understanding of each of the different datum reference methods and the intended use of the optic are key to how the centering tolerance should be assigned.[3,4] The centering of an optical element is one of the most important aspects of the fabrication process to ensure that system alignment tolerances can be met.[5]

The method of centering an optical surface—and subsequently defining a tolerance with an ISO 10110-6 callout—depends on the type of surface or surfaces, and how the element will be mounted in the final assembly.

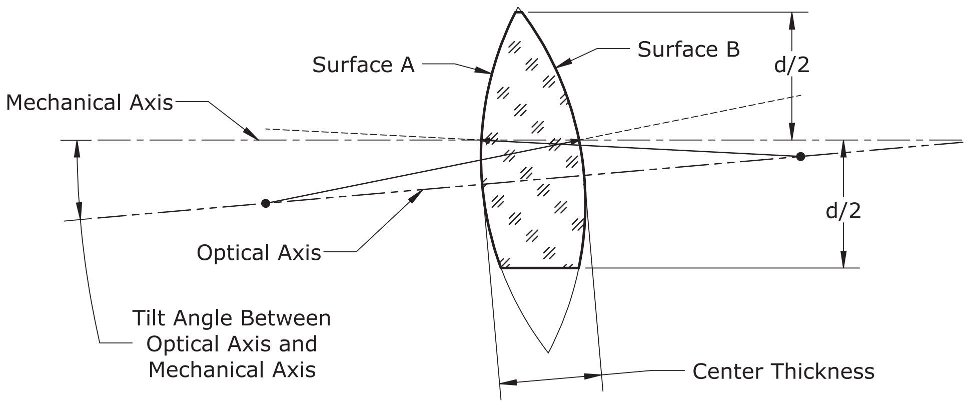


Figure 8.1 Drawing of tilt between two optical surfaces for fabricating an optical element.

Rotationally invariant optical element centering tolerances typically relate to the alignment of the optical axis and the mechanical axis. In cases where optical elements are not cylindrical on their outer periphery, the datum of an optical axis relative to a mechanical axis may relate to other mechanical references—such as points, axes, or planes of the optical element.

In general, rotationally invariant cylindrical periphery optical elements comprise two optical surfaces and the cylindrical edge surface. For this type of element, an optical axis is simply defined by the theoretical axis connecting the two optical surface centers of curvature. A mechanical axis is formed by a theoretical axis defined solely by the cylindrical edge surface. It is the difference between these two axes that we control with centering tolerances.

As an example, the optical element shown in Fig. 8.1 has a mechanical axis defined by the axis of its cylindrical edge surface. We can use that axis as the reference, or datum, to which other features are related. Additionally, the two optical surfaces create an optical axis established by connecting their centers of curvature. The angle between the optical axis and mechanical axis is shown on the left. The element wedge can be taken to be the angle between the lines shown from the two surfaces to their center of curvatures. There are multiple ways to specify tilt, wedge, and decenter within ISO 10110-6. Figure 8.1 is a straightforward example of the objective of a centering tolerance.

8.2 Datum Definitions

As previously discussed, a datum is an ideal point, axis, or plane that can be a reference. A datum for an optical element in ISO 10110-6 can be defined by a few different methods; e.g., the optical axis, the mechanical axis defined by a cylindrical edge surface, a single point, a mechanical plane on the optical element, or an optical surface alone. In most cases, the datum for a surface is an axis, and the centering tolerance is given as the angle of the surface normal at the point where the surface intersects that datum axis.

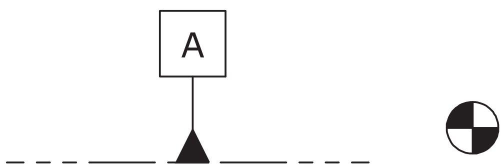
A datum feature is a physical feature found on the optical element. The feature serves as a basis to locate the datum. While the datum feature is not a geometrically perfect surface, each subsequent datum is constructed from the initial feature. For example, a cylindrical edge surface feature may be toroidal, conical, or have a lobed profile; but we can define an imaginary perfect cylinder as the smallest cylinder that can be circumscribed about the feature. In ISO 10110-6, such an ideal cylinder is defined as a cylindrical surface datum. The axis of that cylinder is defined as the cylindrical datum or datum axis. Similarly, an approximately spherical surface feature of an actual lens has a spherical surface datum defined as the best-fit sphere to that surface. The spherical datum is that sphere's center of curvature. This datum point can be combined with another datum point, such as the geometric center of the same surface, to define a datum axis.

8.2.1 Datum notation

The mechanical axis is the most common datum used in drawings of optical components. In the case that a cylindrical mechanical surface is used, the theoretical axis that this cylinder creates is the mechanical axis. An optical axis is the axis connecting the two optical centers of curvature. As discussed in chapter 1, the optical and mechanical axes are notated differently. The mechanical axis is shown as a dash–dot–dash notation and the optical axis is shown as a dash–dot–dot–dash type line. When describing an axis and/or a plane as a datum, they are explicitly noted with a datum indicator [Fig. 8.2(a)]. In the case that a single point is used for the datum notation, the point will be noted on a drawing with a level datum point [Fig. 8.2(b)]. To understand where a datum point is on an optical element, additional axes or planes may be necessary.

8.2.2 Datum construction

ISO 10110-6 allows two methods of definition for each tolerance: by either explicitly noting the datum and tolerance for the element, or solely noting the tolerances of the surface and implicitly defining the datum. As certain aspects



(a) (b)

Figure 8.2 Datum notation for (a) an axis, plane, or cylinder surface; and (b) a point.

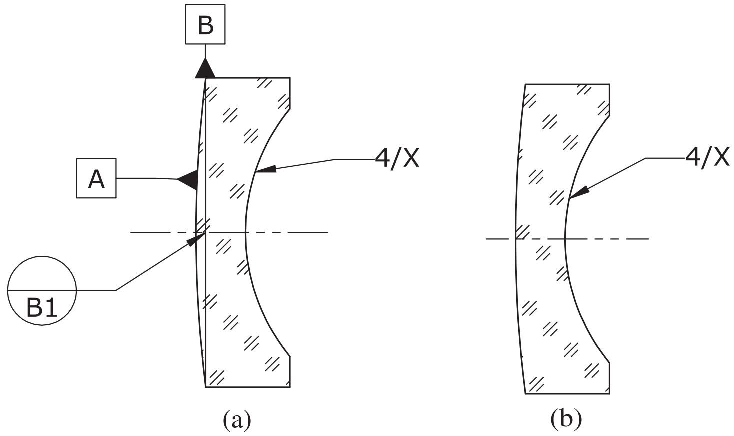


Figure 8.3 Sample optical element where the centering tolerance (surface tilt) for the right surface is (a) explicitly defined and (b) implicitly defined without the datum where the tolerance is the 4/x indicator and quantifier.

of an optical element are defined without confusion (e.g., mechanical axis for a cylindrical surface), such aspects need not be explicitly defined on the drawing. However, it is necessary to ensure the datum is clearly understood if being defined implicitly (Fig. 8.3).

In the case that the centering datums are explicitly defined in the drawing field of an ISO 10110 print, the drawing symbol and tolerance values may be noted in either the drawing field or the table field. If the centering tolerance is noted in the table field, the tolerance is associated with a single surface. Defining a centering tolerance of an optical element in an assembly, or a beam

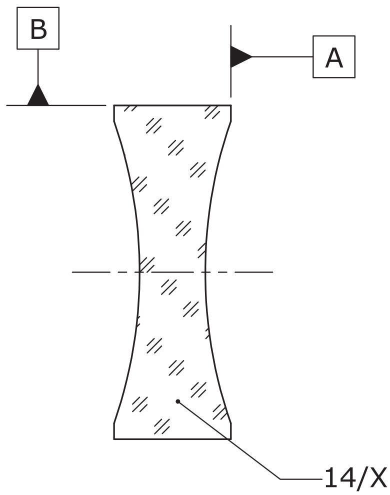


Figure 8.4 Optical element with datum hierarchy starting with the annular surface and followed by the outer diameter.

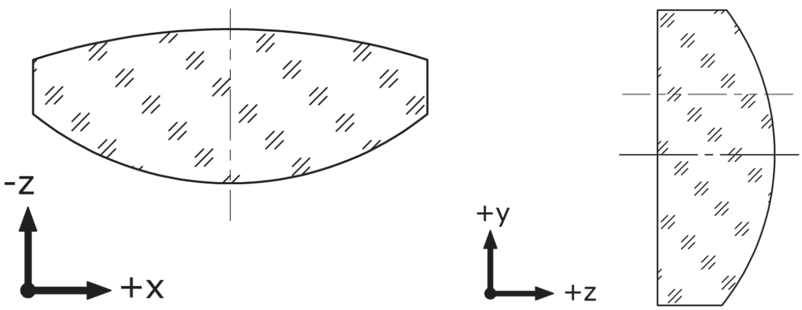
deviation tolerance of an element or assembly, must be shown in the drawing field itself.

When working with a datum system for any type of drawing, there is a hierarchy in place based on the alphabetic letter on the print. The first letter takes precedence and will drive tolerances on the remainder of the datum system. In the case of Fig. 8.4, the flat annular surface on the right-hand side of the optical element is where references should begin. Following that flat annular surface, the cylindrical edge surface is a secondary datum feature. The tolerance that is shown should be referenced against the axis of the smallest superscribing cylinder that is perpendicular to the flat annular surface.

8.3 Coordinate Systems

When specifying an aspheric or nonrotationally symmetric optical element, the centering in an optical element cannot solely be controlled with a tilt tolerance. In this situation, a tilt and decenter are both relevant to the centering tolerance. With rotationally symmetric aspheric surfaces, the standardized coordinate system is given in ISO 10110-12 and matches the typical coordinate system for a rotationally invariant spherical optical element. The optical axis is the z-axis and the coordinates are defined using the yz-plane. ISO 10110-19 provides a standardized coordinate system for general surfaces including freeforms. For right surfaces, the default coordinate system used in ISO 10110-12 may cause confusion because users in some countries (such as the US) sometimes define the $+z$-axis as going into the part on both sides of the element. Care must be taken in the drawing to be sure the z-axis used is understood by all parties. Nonetheless, these surfaces are discussed further in Chapter 9.

When engineering an optical element without rotational symmetry, defining the Cartesian coordinate system is necessary to ensure the centering tolerance is satisfied (Fig. 8.5). In that case, defining the datum explicitly is critical in ensuring the coordinate system is correctly interpreted.



(a) (b)

Figure 8.5 Coordinate definition in a nonrotationally symmetric optic in two different planes.

8.4 Aspheric and Freeform Element Centration Factors

Unlike spherical optical elements, the centering of an aspheric or freeform optical surface cannot create an optical axis by two points. In the simplest case, a rotationally symmetric aspheric surface feature alone references two points (the center of curvature for the best fit asphere and the aspheric surface vertex).[6] Because of this added complexity, the centering of the aspheric or freeform surface requires at an additional degree of freedom for a decenter of a reference point (Fig. 8.6).

The addition of the three points (R_1 center of curvature, R_2 best fit asphere center of curvature, and R_2 vertex) renders the creation of a single optical axis difficult. Minimizing the decenter tolerance as well as the surface tilt is critical. In some cases, minimal aspheric departure from the best fit sphere may be small enough that the decenter tolerance may not be necessary.

Moving from a rotationally symmetric aspheric surface to a nonrotationally symmetric freeform surface brings additional degrees of freedom.[7] These degrees of freedom are required for optical elements with different radii in each axis. To simplify this understanding, a cylinder lens is analyzed. With this type of optical element, the tilt of the surface is indistinguishable from a decenter because the cylindrical surface is circular. As with spherical optical elements the tilt is present, but for this case surface tilt is independent in each orientation. Because of the differences in radii for different orientations, the notation for the cylinder optical element is shown where the tilt is for the respective Cartesian coordinates. When we move to a freeform optical system, the addition of the decenter becomes present. As seen in Fig. 8.7 the *yz*-plane shows a tilt of the cylindrical surface. Moving to the *xz*-plane, a tilt is present in the cylindrical surface, but it does not have the same tilt as in the *yz*-plane. In this case, since we are looking at a plano-convex optical element there is not a need to include the clocking of the two surfaces together (as would be seen in the *xy*-plane).

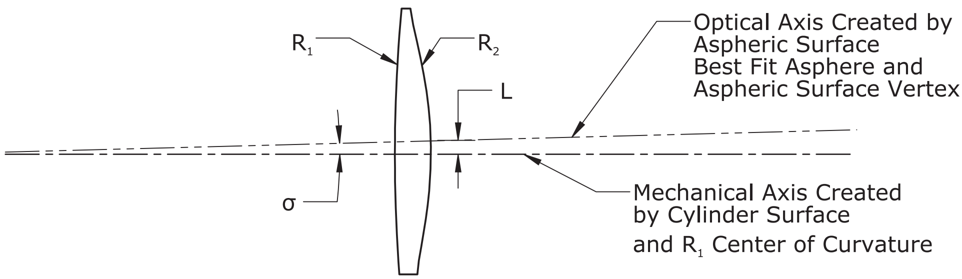


Figure 8.6 Visualization of the right optical surface–aspheric surface (R_2) with a surface tilt and decenter relative to the left optical surface–spherical surface (R_1) and the cylindrical surface datum.

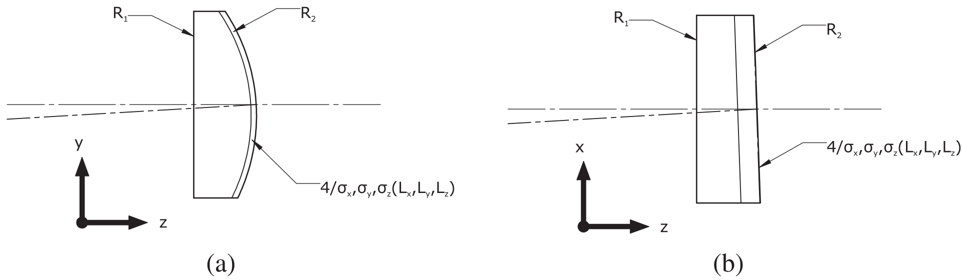


Figure 8.7 Cylinder optical element with tilt present in both the *yz*- and *xz*-orientations.

Freeform optical surfaces may also be designed where the surface definition uses Zernike coefficients, other mathematical functions, or even point cloud data. The more-complex methods create additional challenges for optical centering. These types of surfaces may require multiple surface tilts but a decenter tolerance must also be present. When working with these types of surfaces, the coordinate system in place is critical to ensure the types of decenters and tilts are properly represented. Use of explicit datum systems is encouraged. Implicit centering tolerances may not be enough of a descriptor if their fiducials or datums are not clearly understood. As freeform optical surfaces become more prevalent, their use may not solely be in two-surface optical elements.[8] In these cases, the full Cartesian coordinate including origins is necessary to ensure metrology is possible for each of the centering tolerances.

8.5 Indications on Drawings

The drawing symbol for centering is `4/`. The quantifiers following the drawing symbol are different depending on the centering tolerance method. There are multiple methods to quantify the type of centering tolerance. As such, it is possible to define the datum and centering tolerances of an optical element differently with the same intended result. Subsequently, it might be beneficial to take a tolerance quantifier method and then convert it to a another method based on metrology methods.[9] The quantifier for a centering tolerance also requires an understanding of the associated datum(s). Abbreviated indications are allowed in which the datum(s) are implied in some specific cases. Where two surface tilts are indicated without an explicit datum, the sole datum is the cylindrical edge and the datum axis is the cylindrical axis. Where only one surface tilt is indicated without explicit datums, the datum is the axis connecting the center of curvature and mechanical center of the unspecified surface.

As with other tolerance parameters in ISO 10110, if a dash (–) is present then no tolerance is associated. This is necessary to understand as defining centering tolerance in a tabular format may require no tolerance to achieve the necessary centering method.

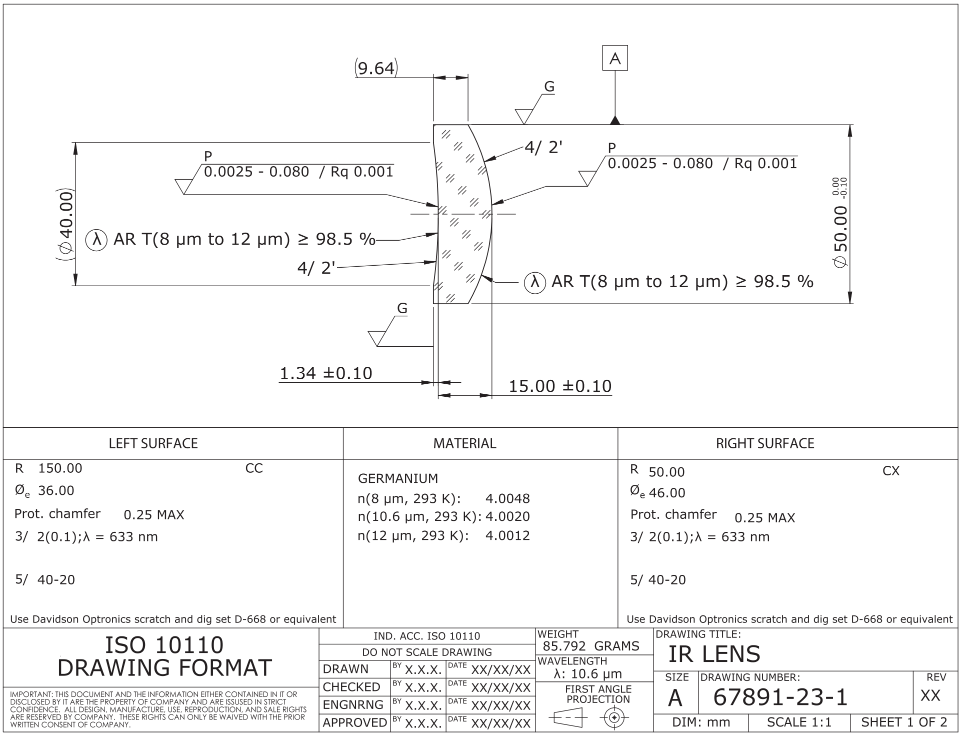


Figure 8.8 ISO 10110 drawing with datum explicitly defined in the drawing field.

An optical element print is shown below in Figs. 8.8 and 8.9, where the centering tolerance is explicitly shown in the drawing field and implicitly defined in the tabular format.

8.5.1 Datum axis

As many optical elements are rotationally symmetric, the outer diameter or cylindrical surface is a simple datum to reference. The cylindrical surface is generally one of the first datums defined. If a centering tolerance will be used for both optical surfaces, the datum would solely be the cylindrical surface [Fig. 8.10(a)]. If a centering tolerance is for a singular optical surface, two datums must be present. In these cases, the other optical surface radius should be the secondary datum [Fig. 8.10(b)]. For either situation, the outer cylindrical diameter is at least one of the datum. This fact is implied in cases where the optical axis and the mechanical axis of the cylindrical surface are coincident.

Along with mechanical centering tolerances, it is possible to define an optical centering tolerance as beam deviation. In this case, the entire optical

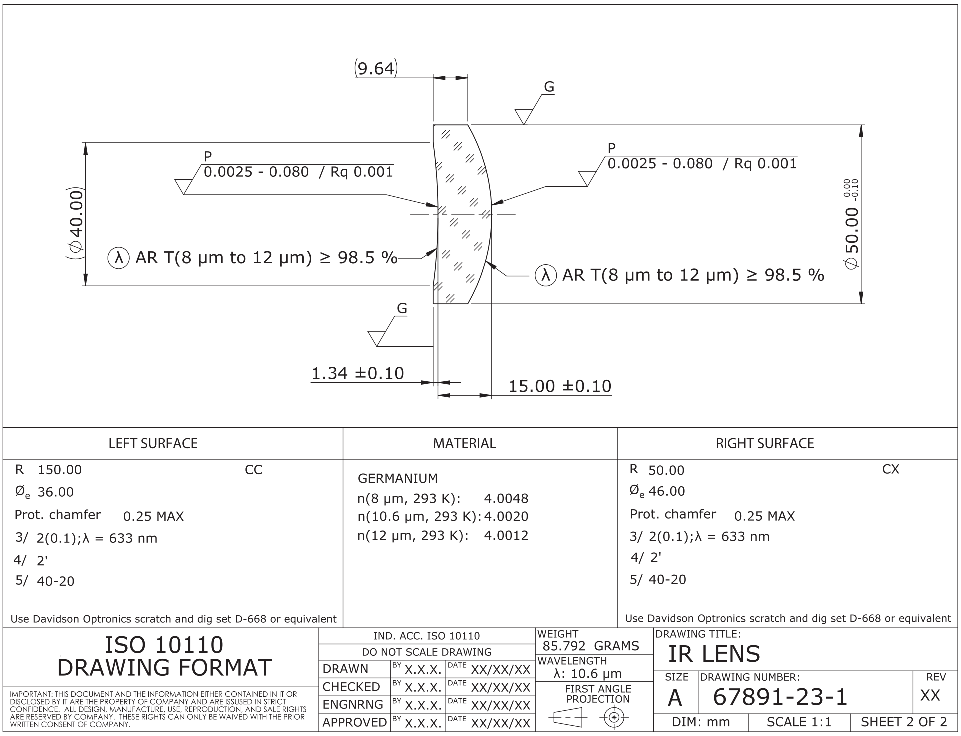
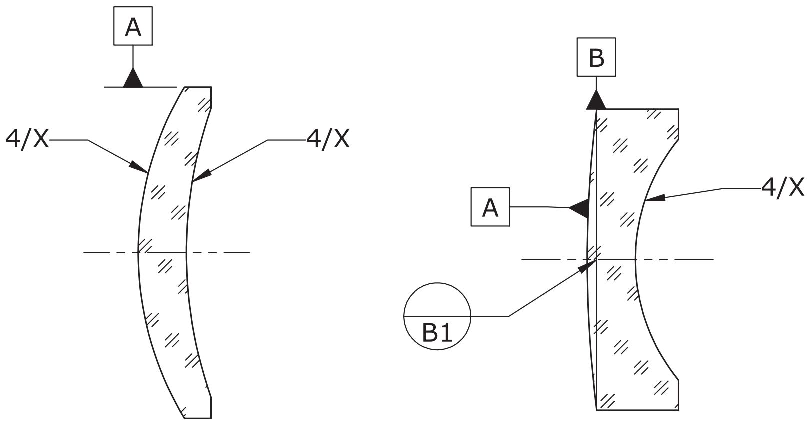


Figure 8.9 ISO 10110 drawing with the datum implicitly defined in the table field.

element has a centering tolerance relative to the optical axis. The datum axis for beam deviation is defined by the cylindrical surface datum as it is a mechanical datum (Fig. 8.11).



(a) (b)

Figure 8.10 Example optical element showing the definition for the outer cylindrical surface (outer diameter) for an optical element with (a) two centering tolerances and (b) two datum and a single centering tolerance.

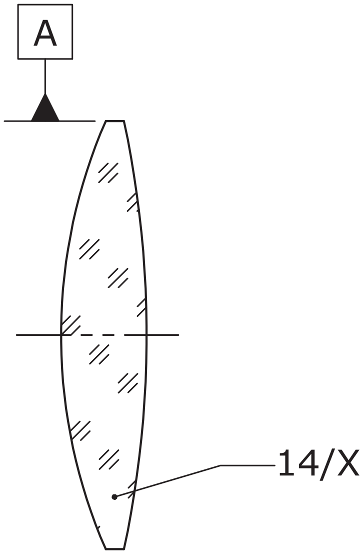


Figure 8.11 Example optical element defined solely by the outer cylindrical surface with a single tolerance of beam deviation.

8.5.2 Datum point

When a datum is not defined sufficiently by a mechanical or physical feature, it is necessary to define the starting datum point. The starting datum point, if not defined, is assumed to be the point where the mechanical axis intersects the left or first surface. For rotationally invariant aspheric surfaces, the datum point is assumed to be the vertex at the axis of rotation of the aspheric surface and need not be shown. The starting datum point must be defined when tolerancing a nonrotationally symmetric freeform optical element where the mechanical reference for the surface is offset in some way (Fig. 8.12). Defining this type of datum point in most cases is not necessary.

8.5.3 Optical centering tolerance

The most basic definition of a centering tolerance is an individual surface tilt. This will be defined by a drawing symbol `4/` followed by the quantifier σ. The quantifier is a tilt in arc minutes, arc seconds, or radians. In definition of the surface tilt, the units must be stated.

In the case that an aspheric optical element is considered, the decenter and the tilt of the optical surface are critical. The aspheric surface, unlike a spherical surface, has a distinct optical axis as well as a vertex that can be displaced. With these surface types, the centering tolerance is defined by the drawing symbol `4/` followed by the quantifier `σ(L)`. The σ for this centering tolerance is still the optical surface tilt in units of arc minutes, arc seconds, or

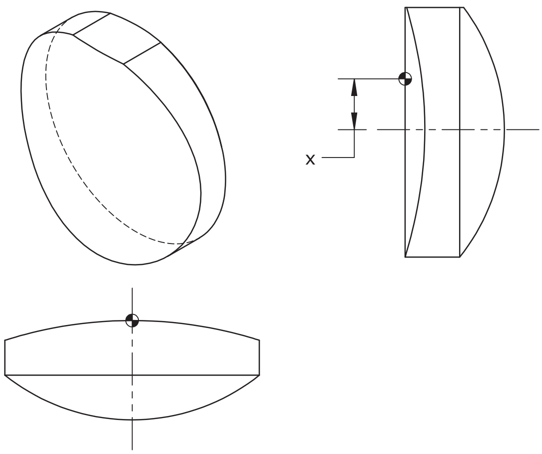


Figure 8.12 An example of a datum point shown on a freeform optical element.

radians. The L is the decenter of the vertex of the aspheric surface with respect to the mechanical center of the surface in units of mm unless otherwise defined. The units for the surface tilt must be stated but the units for the decenter need not be stated.

As tilts and decenters in different directions may impact the performance of a rotationally variant optical surface differently, it is sometimes necessary to define the quantifier in each of the individual directions. These types of surfaces are described by the drawing symbol 4/ followed by the quantifier $\sigma_x, \sigma_y, \sigma_z (L_x, L_y, L_z)$. The σ_x, σ_y, and σ_z are the optical surface tilts about each axis in units of arc minutes, arc seconds, or radians. The units used for the tilt must be stated. The L_x, L_y, and L_z quantifiers are the surface decenters along each axis in each Cartesian direction in units of mm unless otherwise defined. This type of surface may not require tilts or decenters in all directions. If that is the case, the quantifier should be replaced with a dash (–) to state that this term does not have an associated tolerance.

A more mechanical method of specifying a surface tilt is with a runout tolerance for either the optical surfaces or cylindrical surface. If the individual optical surface has its tolerance based on axial runout, a surface tolerance is defined by the drawing symbol 4/ followed by the quantifier a < A, where the quantifier a informs the fabricator that the tolerance is for an axial runout and the A is the value of axial runout in units of mm. For axial runout, the tolerance is applied across the entire clear aperture; but may be measured at another diameter and scaled proportionately.

Similar to axial runout, another mechanical method of tolerancing an optical element's centering is with the circular runout of the cylindrical surface. This type of tolerance is defined by the drawing symbol 4/ followed by the quantifier c < B, where the quantifier c informs the fabricator that the tolerance is for a circular runout and the B is the value of the total circular runout in units of mm. A deeper explanation of mechanical runout is described further in Sec. 8.7.

If the centration of an entire optical element is toleranced instead of a single optical surface, it is possible to specify the allowed beam deviation. In the case that beam deviation is specified, the entire optical element is toleranced with the drawing symbol 14/ followed by the quantifier ρ, where the quantifier ρ is the allowed beam deviation in units of arc minutes, arc seconds, radians, or mm. The units for beam deviation must clearly be stated to avoid confusing angular or linear dimensions. The difference between a beam angle tolerance and a beam height tolerance is discussed more in Sec. 8.7.

Lastly, in the case of an optical subassembly where there are two (or more) optical elements bonded or cemented, the bonding or cemented optical interface may need to be toleranced. For these cases the interface surface tilt is toleranced with a drawing symbol 4/ and a quantifier $\Delta\tau$ where the quantifier τ is the maximum wedge angle between the two optical surfaces in units of arc minutes, arc seconds, or radians. The units must be stated in the case of the surface tilt.

8.5.4 Nonoptical centering tolerance

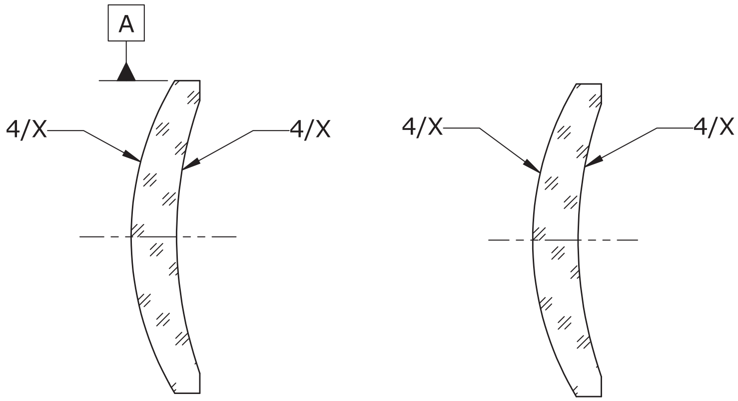
Along with optical surfaces, additional surfaces without optical function may be toleranced using ISO 10110-6. These types of surfaces may be toleranced not only with the notation associated with ISO 10110-6 but also ISO 1101.[10] If the surface is toleranced with notation following ISO 10110-6 they are defined with a drawing symbol 4/ followed by the respective notation used for each tolerance type. In the case where these surfaces are toleranced with ISO 1101 notation, the GD&T symbology is necessary.

8.6 Explicit versus Implicit Indication Examples

The types of different centering tolerances that may be applied to an optical component or optical subassembly may be confusing based on only a description. Examples are shown with the different datums listed and the respective tolerances associated.

In the cases of individual surface tilts with respect to the mechanical axis, the cylindrical edge surface is a necessary datum to define each surface tilt. Figure 8.13 shows that with only the cylindrical surface referenced as a datum each of the two optical surfaces are toleranced with a tilt tolerance.

If the intent is to specify the tilt of a surface with respect to the opposite side, the untoleranced optical surface is a datum along with the center of the surface profile; in this case both datums must be shown in the explicit case along with the surface tolerance [Fig. 8.14(a)]. Where an implicit centering tolerance notation is used, the datum axis is assumed to be the line

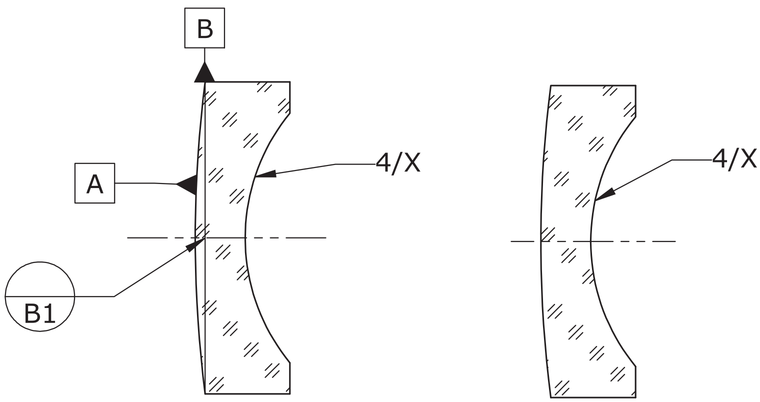


(a) (b)

Figure 8.13 Tolerances on both optical surfaces, referencing the cylindrical surface as the datum (a) explicitly and (b) implicitly.

connecting the center of curvature and the center of the profile of the surface [Fig. 8.14(b)]. The centering tolerance is then the angle of the surface normal at the location where the datum axis intercepts the surface with respect to the datum axis.

When the two datums are used for a single rotationally invariant aspheric optical element, the datum notation is similar to that of a spherical rotationally invariant optical element. For the case of a rotationally invariant aspheric optical element a decenter term is required. The decenter term is in parenthesis after the tilt term as shown in Fig. 8.15. In this case, the decenter tolerance is the distance from the vertex of rotational symmetry to the mechanical center of the specified surface.



(a) (b)

Figure 8.14 Tolerance on the right optical surface, referencing the cylindrical surface and the left optical surface as the datum (a) explicitly and (b) implicitly.

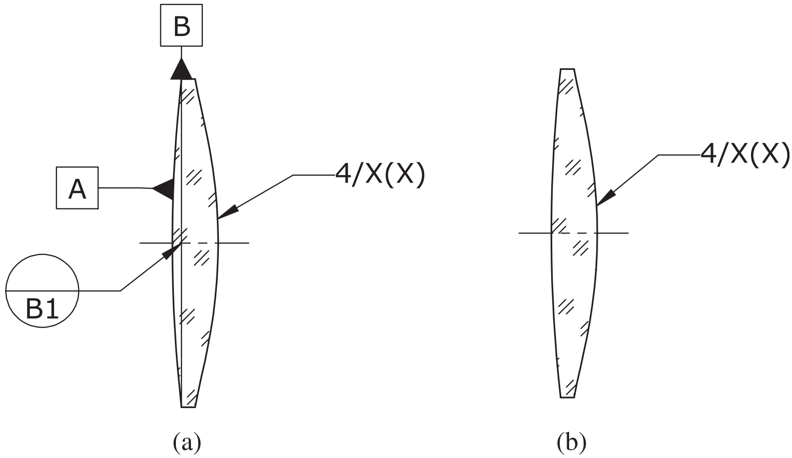


Figure 8.15 Tolerance applied to the right optical surface, an aspheric surface, that references both the left optical surface and the cylindrical surface as the datum (a) explicitly and (b) implicitly.

An optical surface without an axis of symmetry may have decenters in all potential Cartesian coordinate axes. For this situation, the tilt and decenter tolerances can be explicitly noted in their respective directions. To ensure each of these directions is properly represented, a global coordinate notation (Cartesian coordinate notation) should be listed showing the proper directions. Figure 8.16 shows an example where the tilt for both surfaces is evaluated relative to the outer cylindrical surface. The difference in the freeform optical surface relative to a rotationally symmetric optical surface is shown to compare these two types of optical surfaces.

The axial runout tolerance can be used for either one or both optical surfaces. Figure 8.17 shows the right optical surface referenced to the opposite

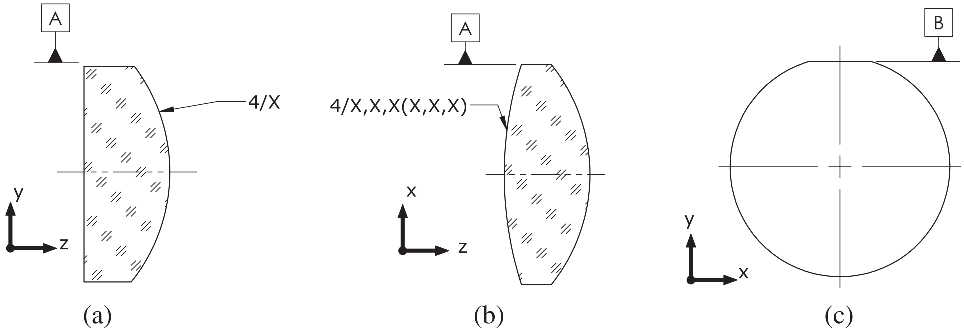


Figure 8.16 Tolerance to a freeform optical element where the left optical surface is a freeform (rotationally varying cylinder surface) and the right optical surface is a standard spherical surface. Both surfaces are referenced to the cylindrical surface (outer diameter) and the flat fiducial. Because of the complexity of this surface, only an explicit tolerance description is shown.

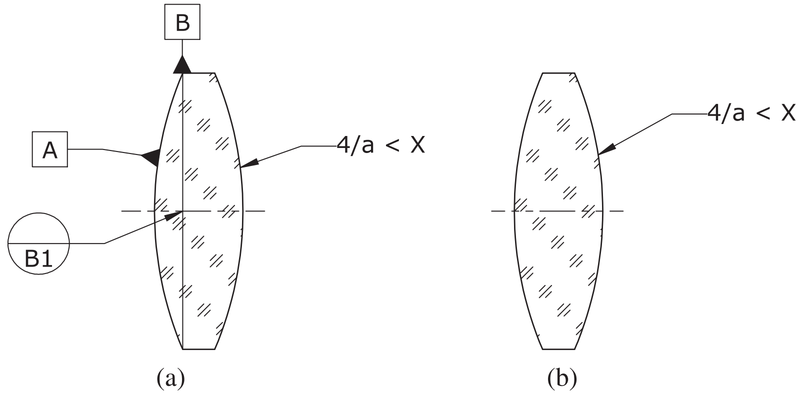


Figure 8.17 Axial runout tolerance on the right optical surface relative to the left surface and the cylindrical surface datum (a) explicitly and (b) implicitly.

optical surface and the cylindrical surface. This tolerance is shown explicitly with the datum to the fixed optical surface and the cylindrical surface. In an implicit notation, without the datum present, the tolerance is assigned only to the referenced optical surface.

In the case that both optical surfaces are specified with an allowed axial runout, the runouts both reference the cylindrical surface as the datum. Both tolerances are then independent from each other (Fig. 8.18).

Beam deviation is a centering tolerance for the entire optical element that will be tested optically instead of mechanically. The tolerance for a beam deviation is shown with the datum reference to the cylindrical surface and the tolerance is shown with assembly notation (`14/`) attributed to the optical element instead of a surface (Fig. 8.19).

If a doublet or optical subassembly has a cemented or bonded optical interface, there are multiple tolerances to consider. The tolerance for the

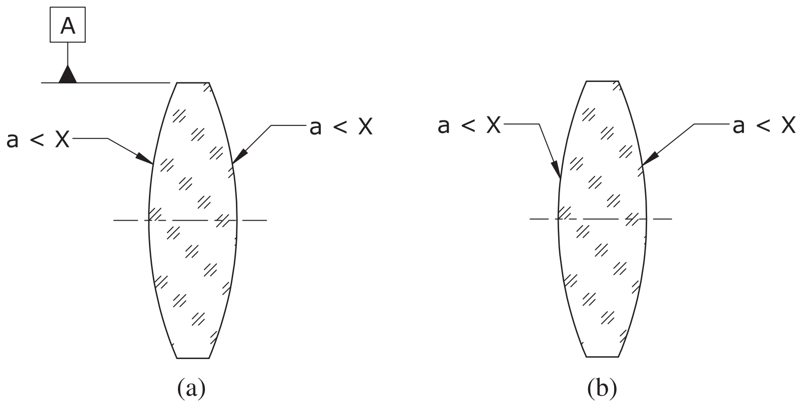


Figure 8.18 Axial runout tolerance on both optical surfaces relative to the cylindrical surface shown (a) explicitly and (b) implicitly.

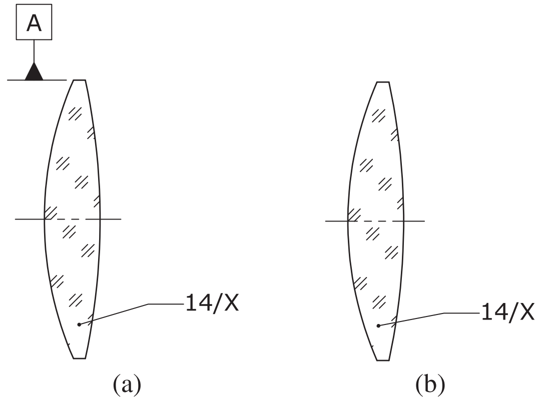


Figure 8.19 Centering tolerance for beam deviation on a spherical lens with both (a) explicit and (b) implicit descriptions.

standard optical surfaces is considered normally by any method previously mentioned. The cemented or bonded optical surface is toleranced differently. This bonded type of interface is toleranced as a wedge angle between the two optical surfaces at the bond interface. There need not be an explicit datum for the bonded surface of the right element. In Fig. 8.20, the datum for the first surface (left surface) and the datum for the edge cylindrical surface are the datum system for the doublet.

Runout tolerances may be defined with either ISO 10110-6 notation or GD&T (as per ISO 1101). For each of the cases with runout specified according to ISO 1101, the datum is explicitly required. When following the notation in ISO 1101, the tolerance indicator has the runout tolerance symbol in the symbol section, the tolerance in the following section, and the datum being referenced at the end [Fig. 8.21(b)]. This same method of tolerancing

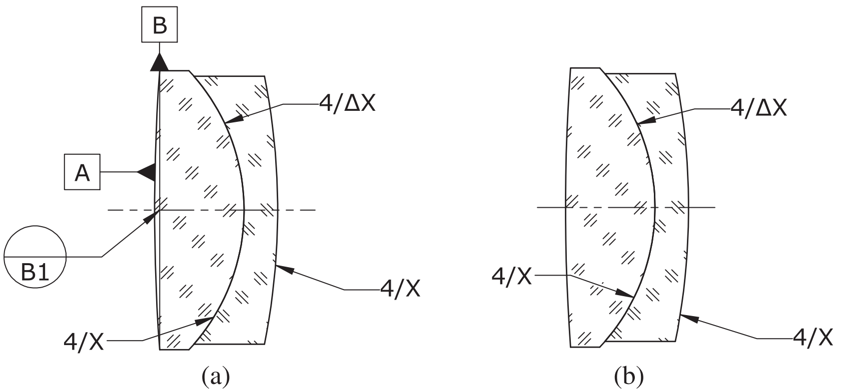
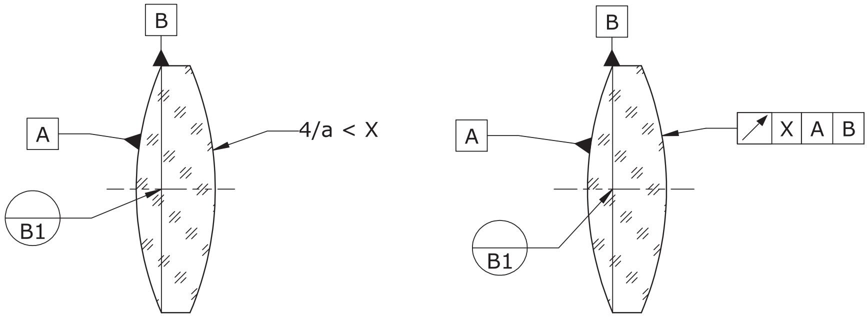


Figure 8.20 Centering tolerance for the first optical element, the cemented surface between the two optical elements and the right surface on the second optical element. This centering tolerance description is shown both (a) explicitly and (b) implicitly.



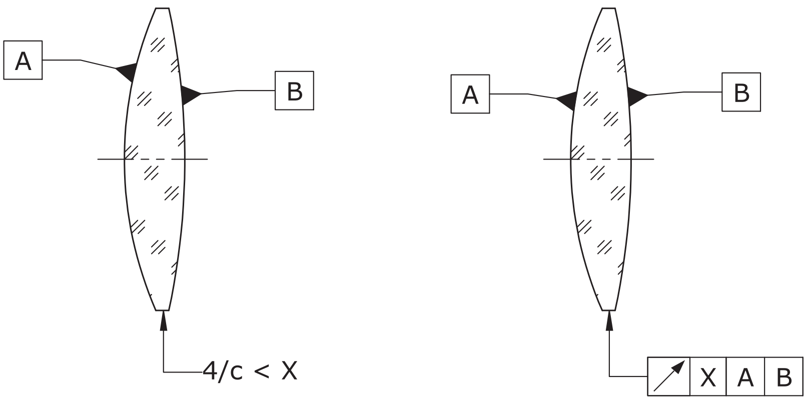
(a) (b)

Figure 8.21 Axial runout explicitly defined for an optical element on the right surface in (a) ISO 1101 notation and (b) ISO 10110-6 notation.

could be used for Fig. 8.18 where both left and right optical surfaces would be toleranced as a runout relative to the cylindrical surface [Fig. 8.20(b)].

The ISO 1101 notation for circular runout on the cylindrical surface must include the datum explicitly referencing the two optical surfaces. This tolerance is listed with a circular runout symbol in the symbol section, the tolerance following, and lastly the datum being referenced [Fig. 8.22(b)].

Along with the optical surfaces, applying a centering tolerance to a nonoptical surface may be done using notation from both ISO 10110-6 and ISO 1101 (Fig. 8.23). When ISO 10110-6 notation is used, if the datum is the axis of the cylindrical surface, it becomes clear using ISO 10110-6 notation where the nonoptical surfaces are referenced and need not be indicated. If there are multiple datums used, specifying the datum in the reference field is critical. When using ISO 1101 notation, the GD&T symbology requires an explicit reference to the datum in the drawing field.



(a) (b)

Figure 8.22 Circular runout explicitly defined for an optical element on the cylindrical surface in (a) ISO 10110-6 notation and (b) an ISO 1101 notation.

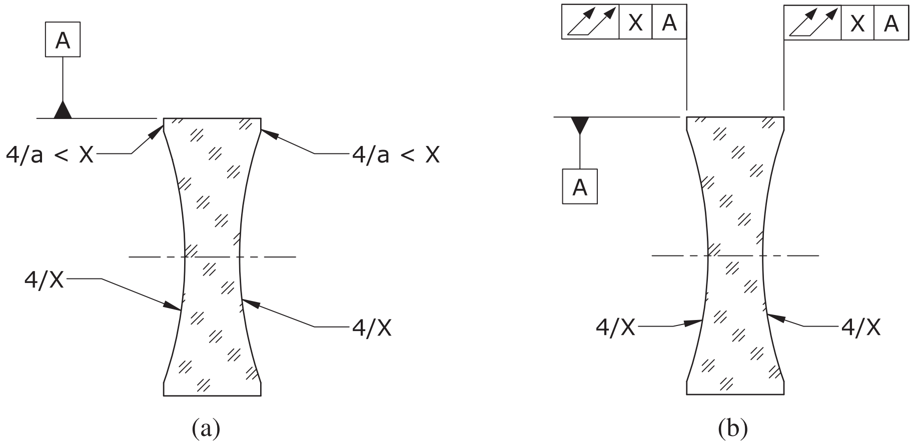


Figure 8.23 Centering tolerance on both optical surfaces and a nonoptical surface referenced to the cylindrical surface. The centering tolerance for both cases are explicitly defined but one is shown with (a) ISO 10110-6 notation and one is shown with the nonoptical surface using (b) ISO 1101 notation.

8.7 Fabrication Explanation

Seeing all the methods of tolerancing the centering of an optical element may be confusing without a clear description of how each of these tolerances would be tested. The figures and descriptions below are simplified explanations for rotationally invariant spherical optical surfaces and methods of measuring centering based on the tolerances assigned.

When looking at a single optical element where the cylindrical surface (outer diameter) is the datum, each optical surface may have its own tilts. With this situation, the mechanical axis determined from the datum is different from the optical axis. Figure 8.24 shows an example of how an optical element with tilts in both optical surfaces creates a difference in the optical axis from the mechanical axis. The mechanical axis, being the reference, is shown as a horizontal line.

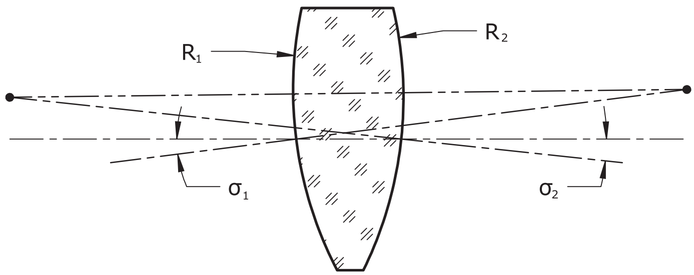


Figure 8.24 Visualization of two optical surface tilts relative to the cylindrical surface datum. This shows the difference between the optical axis and the mechanical axis with each of the tilts.

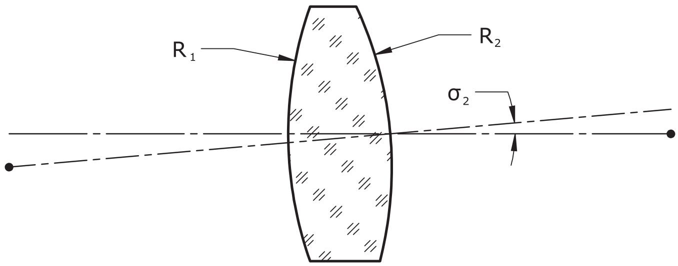


Figure 8.25 Visualization of the right optical surface (R_2) surface tilt relative to the left optical surface (R_1). In this case, the surface normal of the left optical surface at the center of its profile is shown as a horizontal line. The angle of the surface normal of the right surface, as well as the axis of the outer cylindrical surface, could be tilted with respect to this axis.

If we constrain the optical element differently, we can set one of the optical surfaces as a reference. In this type of situation, the optical axis and the mechanical axis will still have a tilt relative to each other. Because of the constraints on this optical element, it is easier to measure the second surface tilt.

With the single optical surface tilt in Fig. 8.25 a decenter for the vertex of the optical surface is present. To calculate the decenter of this optical surface where the tilt is known, Eq. (8.1) should be used:

$$r = R\tan(\sigma) \tag{8.1}$$

where:
σ = tilt of the optical surface,
R = radius of curvature for the tilted optical surface, and
r = decenter of the tilted optical surface.

In the case shown in Fig. 8.26, the two optical surfaces have a tilt with respect to the cylinder axis—which are independent of one another. This leaves the mechanical axis and the optical axis offset and tilted to each other. Testing for this type of error would have the cylindrical surface mounted to a chuck and spun about its axis (mechanical axis). The two optical surfaces would have a runout measured for each other and then computed to a surface tilt.

Using the situation where one optical surface and the cylindrical surface comprise the datum system (as explained in Fig. 8.25), testing the opposite optical surface requires ensuring both the left optical surface and the cylindrical surface are referenced together (Fig. 8.14). For this situation, the left optical surface is rolled about its face until the center of curvature is aligned with the mechanical axis, defined by the cylindrical surface, as in Fig. 8.27. As this optical element is then clamped down, when it is rotated, measuring the tilt of the right optical surface can be achieved with a runout and then calculated to find the tilt. Alternatively, it may be measured optically

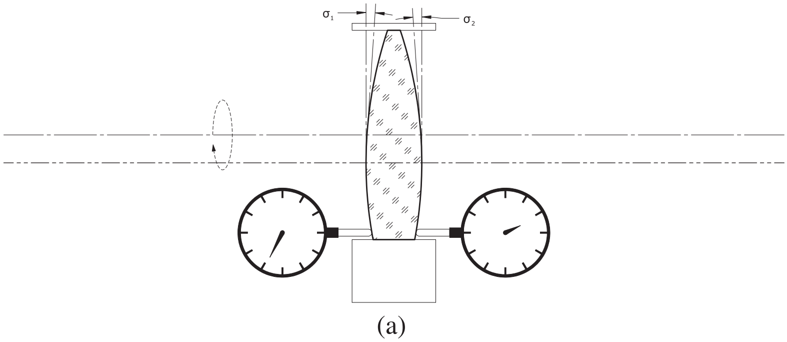


Figure 8.26 Metrology method of testing surface tilt for both optical surfaces. This method is shown with the optical component rotated 180.

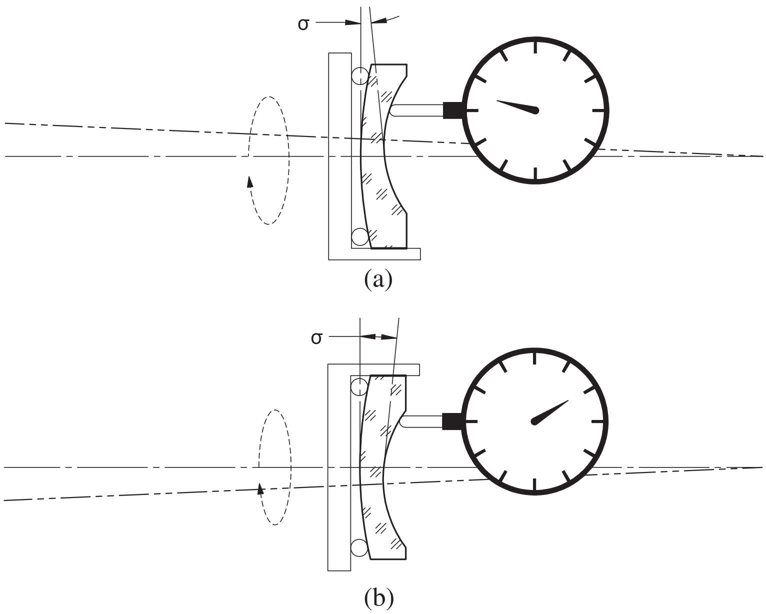


Figure 8.27 Metrology method of testing surface tilt for the right optical surface shown in two positions 180° apart from each other.

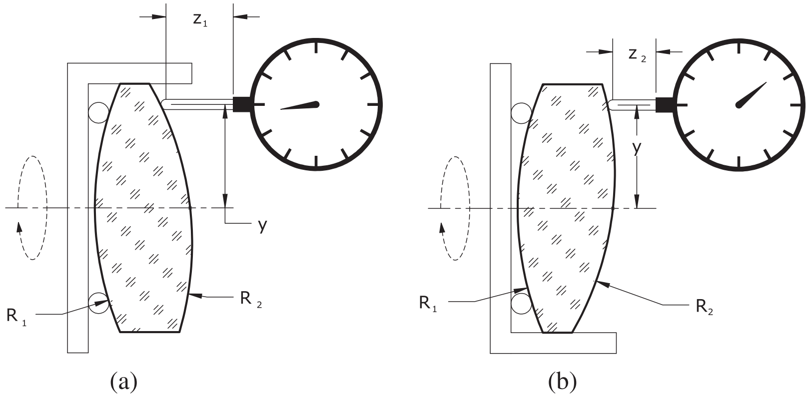


Figure 8.28 Axial runout measurement technique at two different positions showing how the axial runout is measured.

using a centration instrument consisting of a reticle on a microscope that is focused on the center of curvature of the surface opposite the chuck.

For tolerances with an axial runout, the primary datum is the left optical surface and the secondary datum is the cylindrical surface (outer diameter). In this case constraining the optical element in a way that is measurable requires holding this optical element and measuring these differences. Use of a drop gauge that is placed at multiple places radially as the optical element is rotated about the mechanical axis. Figure 8.28 shows how axial runout would be calculated as the difference in z_1 and z_2 at the radial position y.

The setup to measure the circular runout requires the use of a gauge, similar to that of the axial runout measurement setup. In this case the two optical surfaces are mounted in a bell chuck.[11] The clamping from the chuck will set the rotation axis to be the optical axis. As the optical element is rotated about the optical axis the difference in the gauge will give the circular runout measurement. An example of a lens mounted in a bell chuck to measure the circular runout is shown in Fig. 8.29. Note that circular runout is 2× the decenter.

When measuring the centering tolerance of an optical element with optical methods instead of mechanical methods, use of a beam deviation tolerance may be employed. In this case, a collimated light source can be directed through the optical element, thereby measuring the beam deviation angle. Beam deviation can also be expressed as the lateral displacement in position of the focused spot from its ideal location for a perfect optical element. Measuring the beam deviation can be achieved with either angular or linear methods of measurement (Fig. 8.30). The beam deviation then can be calculated back to the optical element wedge angularly with Eq. (8.2). The beam deviation can also be calculated in linear units, based on the optical element or system focal length with Eq. (8.3):

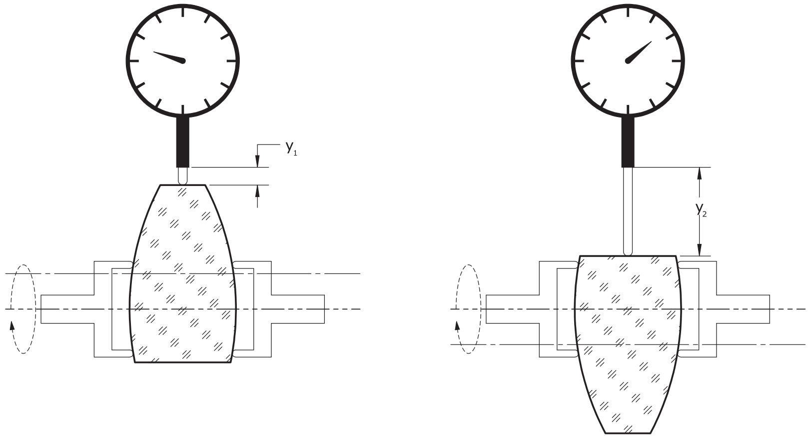


Figure 8.29 Example of a lens in a bell chuck measuring the circular runout.

$$\delta = (n - 1)\alpha \tag{8.2}$$

where:
δ = beam deviation of the optical element focus position in angular space,
n = refractive index of the optical element, and
α = wedge of the optical element; and

$$\delta = \tan^{-1}\left(\frac{\rho}{EFL}\right) \tag{8.3}$$

where:
δ = beam deviation of the optical element focus position,
EFL = focal length of the optical element or system (at the wavelength used in the beam deviation test), and
ρ = lateral beam deviation of the focused spot from the optical element.

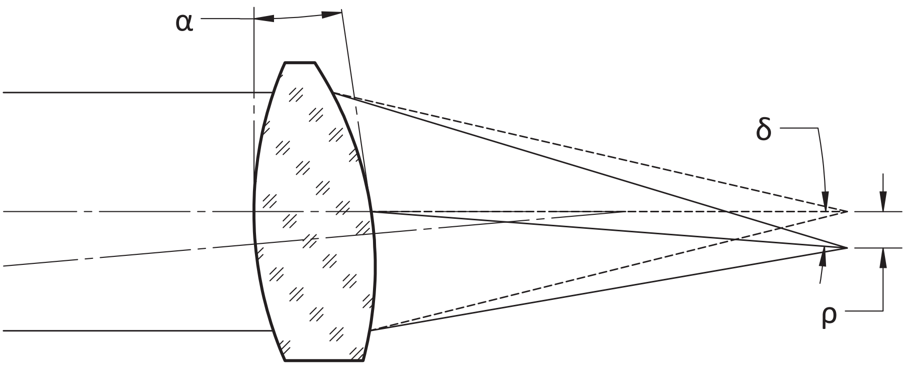


Figure 8.30 Example of beam deviation of an optical element with wedge present, showing the angular (δ) and linear (ρ) tolerances.

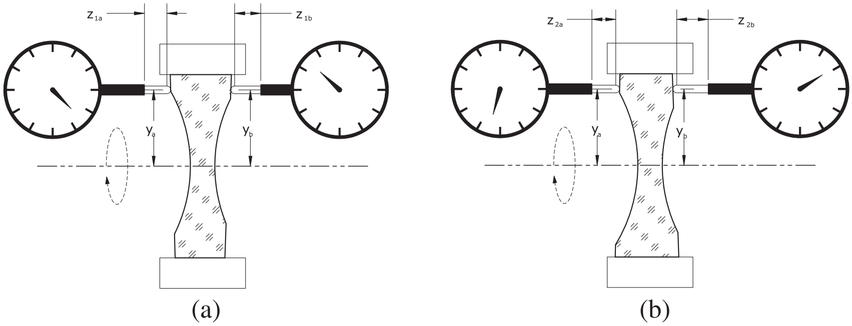


Figure 8.31 Example of an optical element mounted in a chuck to measure the nonoptical annuli surfaces.

When measuring nonoptical surfaces, the same methods of metrology may be used as those discussed above. Similar to Fig. 8.27, measuring the nonoptical surfaces with a runout is achieved with the optical element mounted about the mechanical axis and spun. Measuring the runout of the gauge will show the z displacement at the y-position for both surfaces (Fig. 8.31).

If the use of parallelism is specified for an annulus relative to the opposite annulus, once the surface runout is measured for one nonoptical surface, the optical element is placed on a flat surface (typically a granite surface) and a gauge is used across the entire part to measure the highest and lowest positions (Fig 8.32). The difference in these two z-positions is then reported as the parallelism.

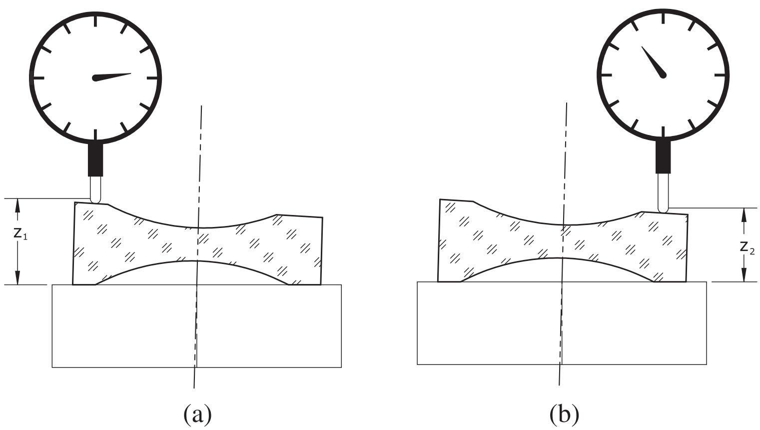


Figure 8.32 Example of an optical element placed on a flat surface to measure the right nonoptical surface and determine the parallelism.

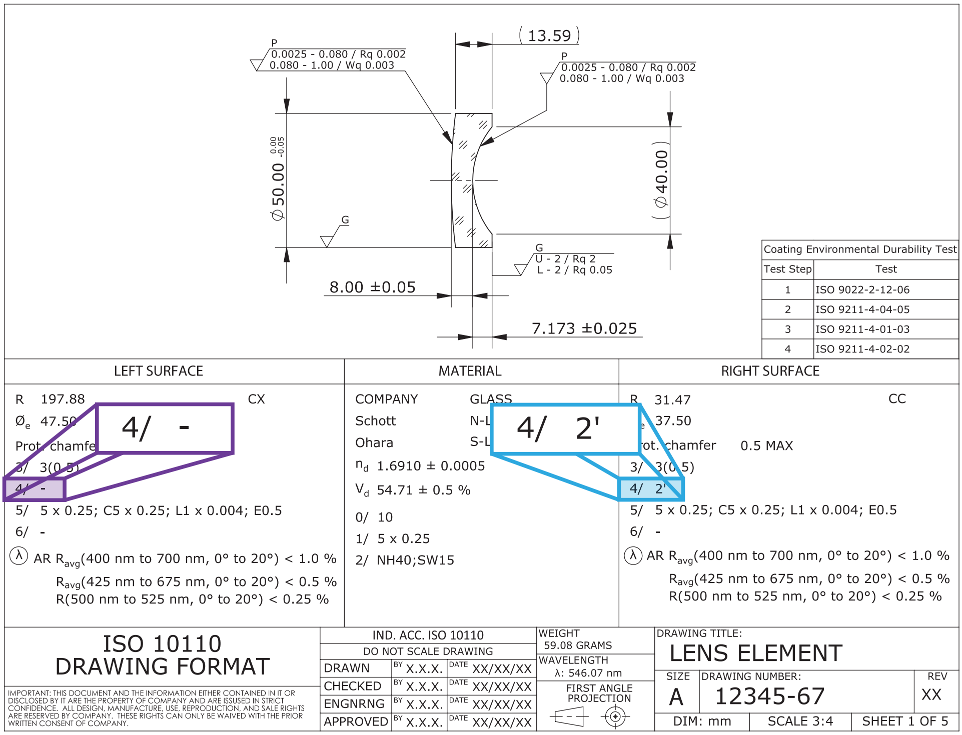


Figure 8.33 Example of spherical optical element ISO 10110 drawing, highlighting centering tolerances.

8.8 Drawing Example

Example prints highlighting the centering tolerances are shown in Figs. 8.33 and 8.34. Both prints show the centering tolerance in the tabular format without explicit datum present. As shown in Sec. 8.6, changing each of these prints to explicitly define the datum is possible. In that case, the datum will be shown in the drawing field and the centering tolerance can also be listed in the drawing field or the table field.

In the first case (Fig. 8.33), the centering tolerance is shown as a tilt tolerance on the right optical surface relative to the datum axis, assumed to be the line connecting the center of curvature and the center of the profile of the left optical surface, as 2 arc minutes.

For the second example print shown, the example is for a rotationally invariant aspheric optical element where there is a surface tilt and surface decenter on the aspheric optical surface. The decenter tolerance is the allowed distance between the vertex of rotational symmetry of the asphere and the mechanical center of the aspheric surface. The tilt of the asphere is relative to the datum axis defined by the center of curvature of the left optical surface (spherical surface) and the center of the surface profile of the left surface. In

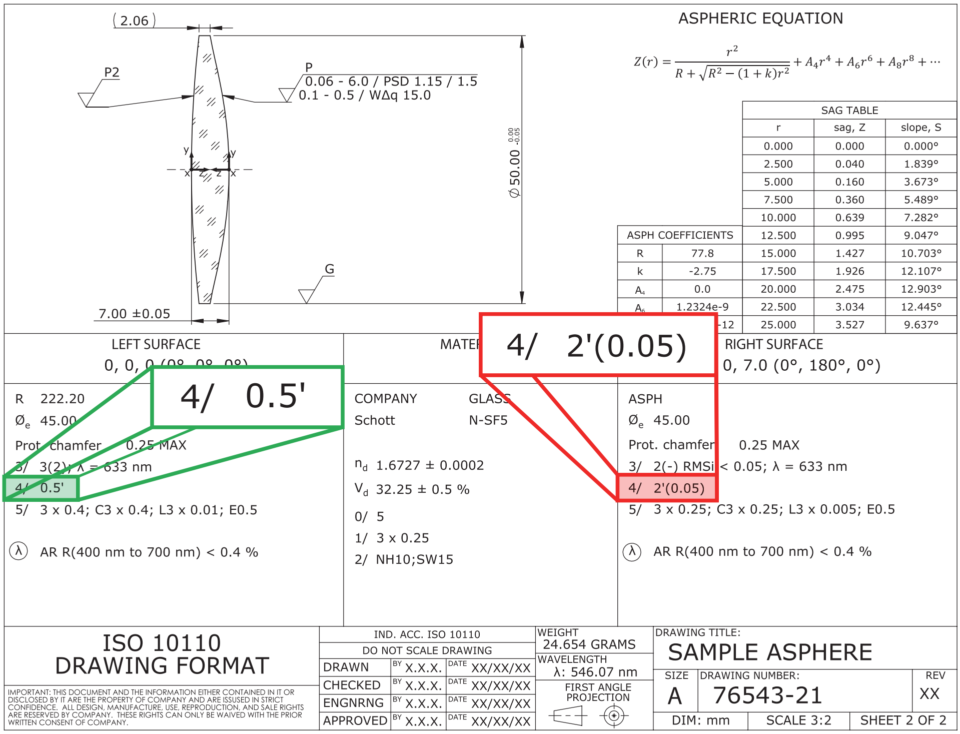


Figure 8.34 Example of aspheric optical element ISO 10110 drawing, highlighting centering tolerances.

this case, the tilt tolerance is 2 arc minutes and the decenter tolerance is 50 μm.

Both drawing print examples are a small subsection of how a centering tolerance may be shown on a print as described in much more detail in Sec. 8.6. These two prints are meant to show an example of how the centering tolerance would be listed in the table field with implicit datum instead of in the drawing field, with either explicit or implicit datum.

References

1. International Organization for Standardization (ISO), *Preparation of Drawings for Optical Elements and Systems—Part 6: Centring Tolerances*, ISO 10110-6:2015 (Geneva: 2015).
2. H. H. Karow, *Fabrication Methods for Precision Optics*, 1st ed., Wiley (Hoboken, New Jersey: 2004).
3. R. Williamson, "Specifying tilts, decenters, and beam deviations using the new edition of ISO 10110-6," *Proc. SPIE* **10377**, 1037708 (2017) [doi: 10.1117/12.2276635].

4. R. Williamson, "Beyond centration: how to create, read, and use a datum system per ISO-10110-6 (2015)," *Proc. SPIE* **11175**, 111751D (2019) [doi: 10.1117/12.2536904].
5. E. Friedman and J. L. Miller, *Photonics Rules of Thumb: Optics, Electro-Optics, Fiber Optics, and Lasers*, 2nd ed., McGraw–Hill (New York: 2004).
6. R. H. Wilson et al., "Considerations for tolerancing aspheric optical components," *Appl. Opt.* **43**(1), 57–66 (2004).
7. S. R. Kiontke, D. M. Aikens, and R. N. Youngworth, "Description and tolerancing of freeform surfaces in standards," *Proc. Int. Optical Design Conf*. 2014, ITh3A.2 (2014).
8. R. N. Youngworth, S. R. Kiontke, and D. M. Aikens, "Implementing ISO standard-compliant freeform component drawings," *Opt. Eng*. **55**(7), 071205 (2016) [doi: 10.1117/1.OE.55.7.071205].
9. R. Williamson, "Centration expressions and conversions," *Proc. SPIE* **10316**, 103160Z (2007) [doi: 10.1117/12.717256].
10. ISO, *Geometrical Product Specifications (GPS)—Geometrical Tolerancing—Tolerances of Form, Orientation, Location and Run-Out*, ISO 1101:2017 (Geneva: 2017).
11. R. Williamson, *Field Guide to Optical Fabrication*, SPIE (Bellingham, Washington: 2011).

Chapter 9
Nonspherical Surfaces

The advantages of nonspherical (that is, rotationally invariant aspheric) profiles for optical surfaces to correct aberrations in optical systems have been known since the early days of geometrical optics and optical design.[1,2] More recently freeform surfaces (that is, rotationally varying surfaces or surfaces with bilateral or no symmetry) have been used in ophthalmic and nonimaging applications.[3] As fabrication and metrology methods have advanced, these surfaces have been increasingly used in imaging applications.[4] Due to the need for standards to cover drawing notations for many different types of aspheres and general surfaces, two parts of ISO 10110, Parts 12 and 19,[5,6] are dedicated to the specification of the theoretical (nominal) aspheric and general surfaces, respectively, in drawings. Additionally, other parts of ISO 10110 have been written (and revised) to support these types of surfaces; notably Parts 5, 6, 8, and 14. In this chapter, we provide background to using ISO 10110-12 and ISO 10110-19 to create drawings for numerous types of nonspherical surfaces. Development and a summary of descriptions in these standards can be studied independently of this book in the article by Schuhmann[4] and the freeform paper by Youngworth et al.[7]

9.1 Background

Optical systems have long incorporated various nonspherical surfaces, including rotationally variant and rotationally invariant surfaces. The critical factor for deciding when to use an asphere (here, we use this term interchangeably with a nonspherical rotationally invariant surface) or freeform (here, we use this term interchangeably with a nonspherical, rotationally varying surface) ultimately is a balance between the functional requirements of the optical assembly and the additional expense incurred. Exceptions are cases for the following:

- Inclusion of the nonspherical profile surface renders the overall system simpler,
- The surface is required for performance with all given constraints, or

- When the manufacturing method is conducive to using general surfaces and/or aspheres.

These manufacturing methods may include diamond turning, injection molding, or glass pressing. Stated differently, in some systems the use of aspheres is quintessential to the application or manufacturing method; in other cases, inclusion of an asphere can improve performance or simplify the design form.

Enabling technologies have ultimately rendered aspheres, and more recently general surfaces including freeforms, key components in the modern lens designer's toolkit. When designed properly, these types of surfaces can provide greater design flexibility than spherical surfaces due to their additional degrees of freedom. Implementation has included low- and high-volume applications, and has covered the span of sizes of optics. A few examples highlight the need for standards to cover specifying these types of surfaces. Aspheric surfaces have long been pervasive in reflective systems such as telescopes, where aberration correction with fewer surfaces is essential. In infrared systems, requirements on limited surfaces and the availability of diamond-turning for surface figuring also make aspheres quite common. Similarly, aspheric surfaces of plastic parts that can be injection molded or glasses that can be press molded are commonly employed.

With a clear need to specify aspheric surfaces, ISO 10110-12 was developed in the 1990s for conicoids, aspheric surfaces, and quadric surfaces with low-order symmetry (such as toroids). Other parts of the standard were written to contain notation within their scope to specify aspheres and ultimately freeforms, as discussed in the next section. Over time these definitions have been improved and revised with a goal of fitting the needs of industry. Most notably, ISO 10110-19 was published in the 2015 to facilitate any general surface definition for any precision optics surface a user may want to define that is not included in ISO 10110 Parts 1 or 12.

9.2 Using ISO Standards for Nonspherical Profile Surfaces

Historically, standard surface descriptions for aspheres were developed in the 1990s and implemented in ISO 10110-12. Notation to support endemic errors to nominal surfaces and components was (arguably) achieved primarily in ISO 10110 Parts 5, 6, 8, and 14. ISO 10110-5 and ISO 10110-14 consider the surface figure and wavefront, ISO 10110-8 considers errors such as waviness and roughness, and ISO 10110-6 considers centration errors. Additional supporting metrology standards such as the ISO 14999 series and ISO 14997 can be applied generally, although perhaps some expansion of metrology standards may benefit the community further.

In the 2010s, as freeform surfaces became more widely used, ISO/TC 172/SC 1 was confronted with expanding the ISO 10110 series to cover a wider variety of precision optical surfaces. The committee fundamentally had to choose to either create new parts for general surfaces, including freeform surfaces; or to expand existing standards to cover these needs. The committee decision was twofold. A new part, ISO 10110-19, was created for the description of general surfaces (including freeforms not already standardized in ISO 10110-12). Simultaneously, existing supporting surface tolerance standards ISO 10110 Parts 5, 6, 8, and 14 were expanded to accommodate the need for more general error specifications.

In terms of surface definitions, this choice led to the current nominal surface specification guidelines as follows. ISO 10110-1 provides the means to fully define optical elements with surfaces that are planar, spherical, and spherical-profile cylinders, as well as providing notation for local and global coordinate systems. ISO 10110-12 defines a default coordinate system, sign conventions, and definition of nominal surfaces for aspheres (e.g., conicoids, polynomials, and Forbes Q-Types)[8] and surfaces with low-order symmetry (such as aspheric profile cylinders and toroids). ISO 10110-19 describes a coordinate system methodology that can be applied for nearly any conceivable type of optical surface. Exceptions to applying ISO 10110-19 are Fresnel surfaces, diffractive optics, micro-optical surfaces, and ophthalmic optics. Authors of the standards intend ISO 10110-19 to be used when there is no defined surface described in ISO 10110 Parts 1 and 12. ISO 10110-19 hence is not simply freeform surfaces but "general surfaces," which covers a wider variety of surfaces such as nonuniform rational B-splines (NURBS), radial basis functions, and point cloud definitions. ISO 10110-19 can thus be applied to define optical surfaces that are rotationally variant or invariant about the element's nominal mechanical axis. In this chapter, we are specifically focusing on how to define nominal surfaces with ISO 10110-12 and ISO 10110-19.

The coordinate system approach used in ISO 10110 is self-consistent as mentioned in Chapters 1 and 8. The coordinate system definition standardized in ISO 10110 is a right-handed and orthogonal coordinate system using the x-, y-, z-axes, as shown in Fig. 9.1. If a surface is on-axis, the z-axis is the optical axis. The sign chosen for a radius of curvature depends on the root of the quadratic solution for a conicoid that is chosen, and the convention used is the de facto standard in all raytracing codes. In the ISO 10110 standard, assuming the z-axis is increasing from the left to the right, the radius of the curvature is positive if the center of curvature is to the right of the vertex, and negative if the center of the curvature is to the left of the vertex. This choice also indicates that the sagitta (i.e., sag) of a point on a surface is positive if the point is to the right of the vertex, and negative if the point is to the left of the vertex. The coordinate system approach described in ISO 10110-19 is more general but is also specified as a right-handed Cartesian system with the z-axis as a reference. If not given on the drawing, the default general surface coordinate

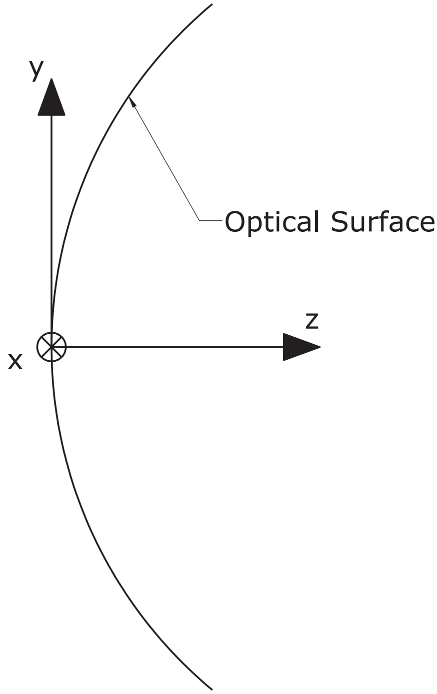


Figure 9.1 Standardized coordinate system for aspheres and general surfaces.

system matches the system provided in ISO 10110-1 Note if the right surface of a lens is aspheric, the default coordinate system used in ISO 10110 may cause confusion because users in some countries (such as the US) define the $+z$-axis into the part (or, in some cases, out of the part) on both sides of the element. Care must be taken in the drawing to ensure that the z-axis used for each surface is understood by all parties.

One last key factor is that ISO 10110 Parts 12 and 19 exclude definition of freeform surfaces for ophthalmic applications. These standards are developed in ISO/TC 172/SC 7; notably for spectacles in the ISO 8980, ISO 9801, and ISO 10322 series.

9.3 Defining Nominal Surfaces with ISO 10110

ISO 10110-12 specifies rules for defining nominal aspheric surfaces with low-order symmetry using an equation. ISO 10110-19 allows users to define surfaces not included in ISO 10110-12, with scope exceptions of “diffractive surfaces, Fresnel surfaces, ophthalmic glasses, and micro-optical surfaces.” In addition to what we describe here, ISO 10110-12 includes informative Annexes summarizing aspheric surface types and details on orthonormal in slope and amplitude aspheres—frequently referred to as Q-type or Forbes aspheres. ISO 10110-19 includes an informative Annex that provides information on mathematical descriptions, data formats, and data reduction for general surfaces. Further discussion of drawing specifics is covered in the next section.

9.3.1 Equations for rotationally invariant aspheric surfaces in ISO 10110-12

The first formula for an aspheric surface in ISO 10110-12 is the standardized version of a conicoid base with a polynomial function of even radial terms of order greater than four added to it. This surface is rotationally invariant about the z-axis and its definition is:

$$h^2 = x^2 + y^2 \tag{9.1}$$

$$z(h) = \frac{h^2}{R\left[1 + \sqrt{1 - (1+\kappa)\left(\frac{h}{R}\right)^2}\right]} + \sum_{i=2}^{n} A_{2i} h^{2i} \tag{9.2}$$

where:
h = a positive root and is the perpendicular distance from the z-axis,
z = sagitta,
R = radius of curvature of the base sphere,
κ = conic constant, and
A_{2i} = aspheric coefficients.

The conicoid for the base surface is an oblate ellipsoid if $\kappa < 0$, a sphere if $\kappa = 0$, a prolate ellipsoid if $-1 < \kappa < 0$, a paraboloid if $\kappa = -1$, and a hyperboloid if $\kappa < -1$.

A second formulation that is expanded from Eq. (9.2) includes even and odd terms in h for the polynomial summation terms starting at $i = 1$. This formula enables, for example, an axicon if only the linear term in h is specified. If surfaces are reversed in a lens, the radius of curvature and all aspheric coefficients change sign. Only the conic constant sign remains the same. In the standard this fact is stated clearly when retaining the orientation of the z-axis.

ISO 10110-12 further includes a family of rotationally invariant aspheric surfaces described by a conic section and orthogonal polynomials. These surfaces are commonly referred to as Forbes aspheres or Q-type aspheres. The standard provides the mathematical descriptions for the surfaces, recursive formulas for calculating to arbitrary order, and enough terms to aid in user implementation. The forms that are included in the standard for Forbes aspheres are orthonormal in slope with a spherical base, orthonormal in slope with a conicoid base, and orthonormal in amplitude with a conicoid base.

9.3.2 Formulas for rotationally variant aspheric surfaces in ISO 10110-12

ISO 10110-12 also provides a set of standardized formulas for some specific surfaces that are rotationally variant. Surfaces included are centered quadrics, polynomial surfaces that are rotationally variant, toroidal surfaces, conical surfaces, and a method to combine surfaces. Inclusion of these forms in ISO 10110-12 can arguably be used to define "freeform" surfaces is a primary reason why ISO 10110-19 is referred to as "general surfaces" which can also include freeform surfaces. If a freeform surface uses a definition in ISO 10110-12 then users should apply that definition. If the formula is not included in Part 12, then Part 19 is used. Consequently, the decision by the standardization committee was to keep definitions intact in ISO 10110-12, and avoid delving into standardizing surface names such as "freeform" with respect to nonspherical profile surfaces.

The first set of definitions are for centered quadrics, starting with implicit functions for such surfaces (including parabolic surface definitions). These equations can be written as the following explicit function for the sag using the standardized coordinate system:

$$z = f(x, y) = \frac{\frac{x^2}{R_x} + \frac{y^2}{R_y}}{1 + \sqrt{1 - (1 + \kappa_x)\left(\frac{x}{R_x}\right)^2 - (1 + \kappa_y)\left(\frac{y}{R_y}\right)^2}} \tag{9.3}$$

where:
R_x = radius of curvature in the xz-plane,
R_y = radius of curvature in the yz-plane, and
$\kappa_{x,y}$ = conic constants.

ISO 10110-12 also provides a formula written in terms of the curvatures. The conic constants have the same meaning as used in Eq. (9.2). The general form given in Eq. (9.3) opens a wide variety of surfaces. For example, cylinders can be defined by having either radius set to infinity. The standard does not make a specific distinction between cylinders that are formed with a spherical, conic, or aspheric profile. Hence the commonly used term "acylinder" is not currently included as part of the vocabulary (but discussions on this topic are expected to continue in the future), as many practicing engineers use the term. Please note, however, that a cylinder can be defined with a spherical profile indicated by `CYL` in accordance with Part 1 without use of Part 12, but such a surface form could also be described by Eq. (9.3) by simply setting the conic constant in the curved direction to zero as well.

ISO 10110-12 also includes a formula for polynomial surfaces that can be rotationally variant. In a general summation form these surfaces are defined by

$$z = f(x, y) = \sum_{i=2}^{n} (A_i|x|^i + B_i|y|^i) \tag{9.4}$$

where the coefficients A_i and B_i weight the polynomial terms included. Note that this expression can include odd terms $i \geq 3$, and the standard also has a nonsummation form that uses a different notation for coefficients.

ISO 10110-12 includes a section stating that any base surface form, $f(x,y)$, can have another function, $f_1(x, y)$, added to it. For example, one can take a centered quadric form and add a polynomial series to it, essentially combining Eqs. (9.3) and (9.4). There are many additional possible variations allowed using any combination of surface descriptions in ISO 10110-12.

9.3.3 General surfaces in ISO 10110-19

ISO 10110-19 provides a framework for any general surface or component nominal surface definition not included in other sections of ISO 10110. Whether a surface is rotationally variant or invariant, this standard can apply. Consequently, ISO 10110-19 can apply when splines such as NURBS, point clouds, radial basis functions, or any conceivable surface is to be defined as an optical surface.

The standard gives definitions for general surfaces, reference axes, and transport formats. The need for each of these is to facilitate data-sharing and avoid confusion since the standard is intentionally written to be extremely general and flexible. Referencing is important for correct definition of any applied mathematical description, reference points, and reference coordinate systems. Additionally, the standard also includes sections that discuss the coordinate system and sign convention to use. It is quite critical for users to have clear specifications for all the factors to provide drawing specifications covered in the standard.

ISO 10110-19 includes an informative Annex containing suggested mathematical descriptions, formats for data transport, and data-reduction methods for general surfaces. Mathematical descriptions can be discrete or analytical, provided all descriptions contain sufficient data for required surface definition accuracy. Cropping or limiting data must not change the surface definition outside of the required accuracy within the free aperture (where the optical surface is defined). The Annex gives further informative guidelines for surface descriptions using point clouds, polygon meshes, and analytical surfaces (such as different polynomial choices and splines). Data transport formats that are discussed in the Annex include TXT, STL, STEP, IGES, and XML format. Data reduction methods are discussed generally and with respect to point clouds, polygon meshes, polynomial surfaces, and spline surfaces (specifically NURBS).

9.4 Indications on Drawings

9.4.1 Per ISO 10110-12

ISO 10110-12 provides a set of rules for specifying nominal (theoretical) aspheric surfaces in drawings. Aspheric surfaces are indicated in a tabulated drawing by replacing the radius of curvature specification with the indication `ASPH`. Conicoids and surfaces of second order are indicated by replacing the radius of curvature specification with the canonical shape (e.g., hyperboloid or toroid). Other normative requirements for all surfaces defined using ISO 10110-12 on the drawing are:

- The formula for the surface,
- The sign for the radius of curvature,
- Any aspheric or other coefficients, and
- The sagitta table.

The inclusion of the sagitta table and the formulas provides significant ability to verify implementation in production and provides a very valuable error-checking mechanism.

Surface form tolerances shall be specified either with ISO 1101 or ISO 10110-5. Centration shall be specified either with ISO 1101 or ISO 10110-6. Surface imperfections and textures, including mid-spatial frequency errors, are specified with ISO 10110 Parts 7 and 8.

9.4.2 Per ISO 10110-19

General surfaces are indicated in a tabulated drawing with a surface designation (in place of the curvature indication) of `GS`. The origin must be specified, and references marked and dimensioned to the general surface's coordinate system.

Users do have the option to represent the surface in an exaggerated manner to render all factors clear, but the part must be marked accordingly if exaggerated. The inclusion of a sag table is always required for general surfaces to enable point-by-point testing of the surface form. ISO 10110-19 states Cartesian and polar coordinates can be used to verify the sagitta and surface slope. Although the standard does not specify testing, the standard does state that the drawing must have sufficient accuracy and inclusion of data to enable a point-by-point testing method. The sagitta table also includes slope as a control variable.

If a mathematical description is used a note is included to indicate the selected form. If a data file is used in the definition, the file must include any pertinent constants, vectors, and coefficients. In the absence of a data file for exchanging between CAD or other systems, then a mathematical description is required. Data files must include an unambiguous file name and extension.

ISO 10110-19 provides three methods of specifying surface form error:

- ISO 1011,
- ISO 10110-5 (`3/`) or ISO 10110-14 (`13/`), or
- In a table with permissible form and slope deviations including sampling and integration length (in addition to the required sagitta table).

In addition, permissible slope deviation is required for the sag surface, tracking the angular deviation of the real surface from the nominal (theoretical) surface.

Centration is specified with either ISO 1101 or ISO 10110-6. Surface imperfections are specified according to ISO 10110-7. Surface texture (including mid-spatial frequency considerations) is specified according to ISO 10110-8.

Uniform transport methods such as XML or STEP shall be used to transfer surface data between software and hardware machines. Due to the possibility of miscommunication of requirements, all general surface data shall be mapped to sufficient precision with clearly stated units. A sagitta table including slope is still required for testing the transferred description and can be in either Cartesian or polar coordinates. If data reduction is necessary, the method used shall be indicated near the filename on the drawing.

9.5 Drawing Examples

Three examples are included to illustrate drawings that use either ISO 10110-12 or ISO 10110-19 to define the nominal (theoretical) nonspherical surfaces. The purpose is primarily to show the required information for the nominal surfaces. Nonetheless, the drawings also provide other indications and tolerances used in ISO 10110 for completeness (e.g., `3/` for figure, and `4/` for centration).

The first example shown in Fig. 9.2 is an asphere that uses ISO 10110-12. The standardized coordinate system (z-axis to the right for sagitta) is employed in the surface definition for the left spherical surface to define the global coordinate system. The aspheric surface is on the right with a new local coordinate system defined. As required, the equation for the surface, all parameters, and a sag table is included for the aspheric surface.

The second example shown in Fig. 9.3 is a rotationally variant lens that uses ISO 10110-12. The standardized coordinate system (z-axis to the right for sagitta) is employed in the surface definition for the right-hand surface that is a cylinder with an aspheric profile. As required, the equation for the surface, all parameters, and a sag table are included for the xz curved profile right-hand surface.

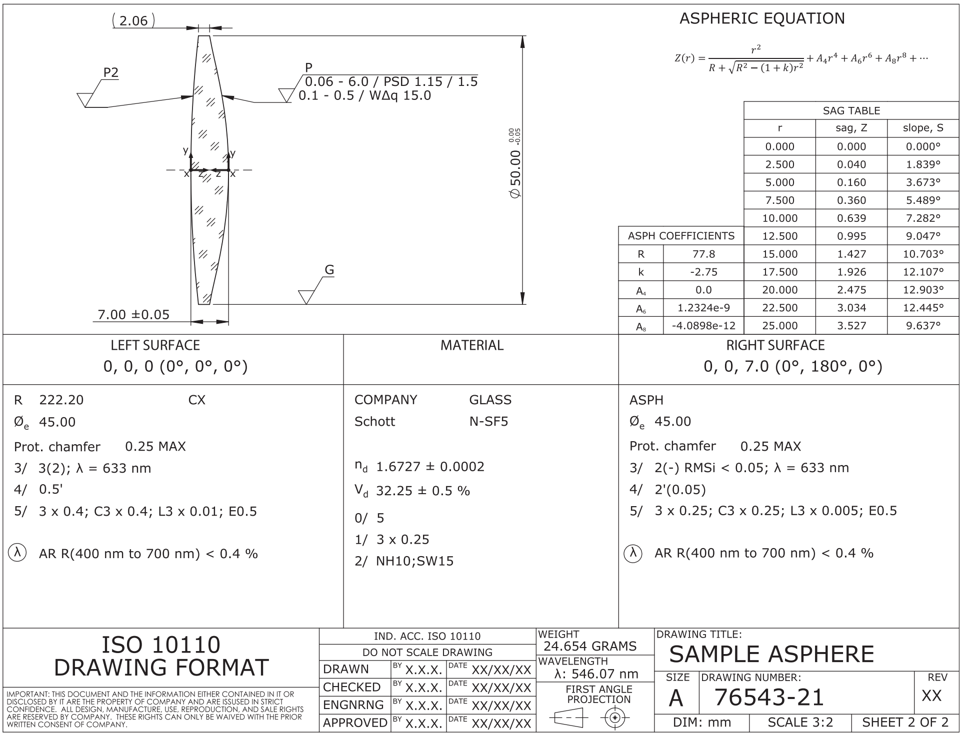


Figure 9.2 Example asphere using ISO 10110-12.

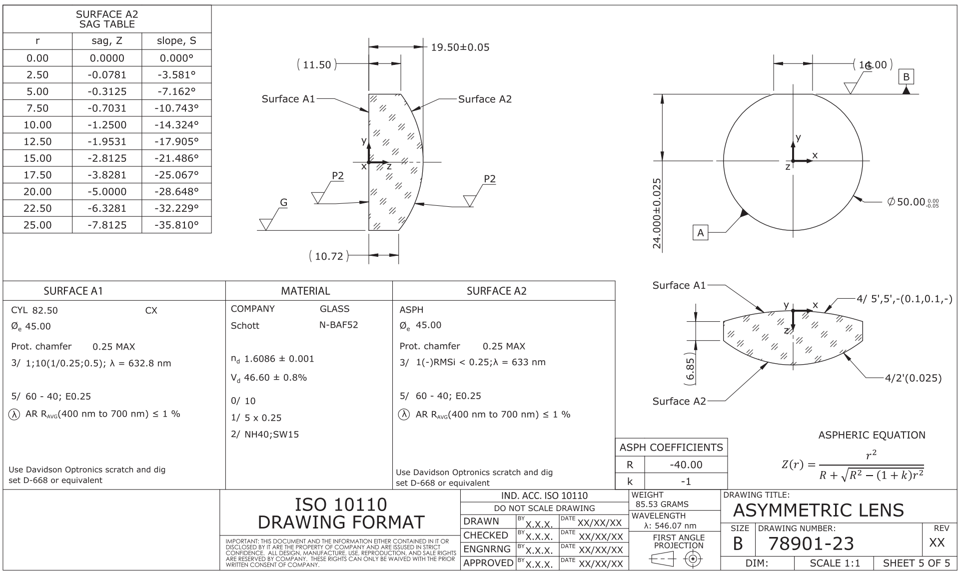


Figure 9.3 Example asymmetric lens requiring ISO 10110-12.

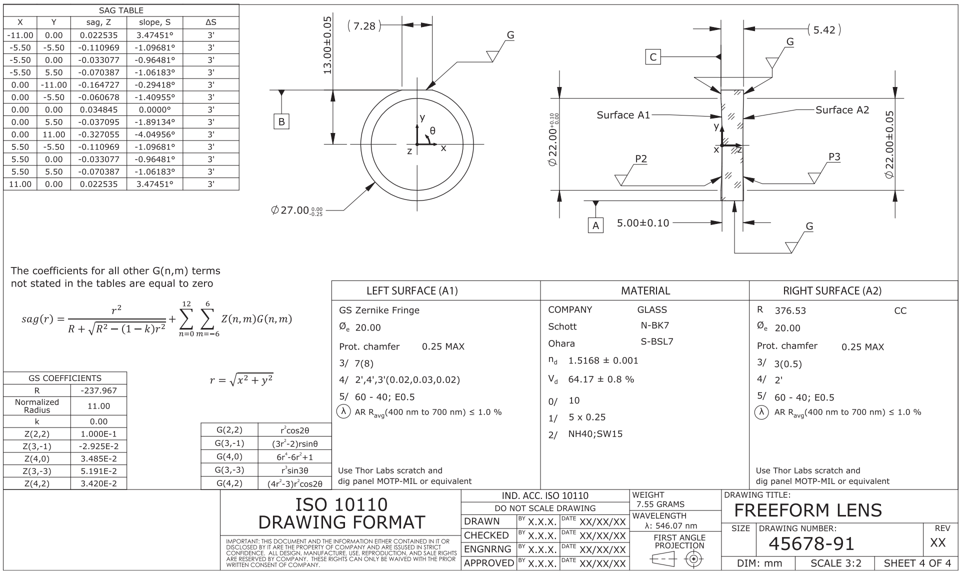


Figure 9.4 Example freeform lens requiring ISO 10110-19.

The third example shown in Fig. 9.4 is a lens with a freeform surface described using ISO 10110-19. The right surface in this example is a concave spherical surface. The left surface is defined with a fringe Zernike polynomial equation, and thus is a general surface that has not been defined in ISO 10110-12. The coordinate system for the surface, required equation, and required sag table and slope are both given to be control variables for the left-hand freeform surface.

References

1. F. Smethwick, *Grinding Optical Glasses*, GB166600149A, United Kingdom (1666).
2. J. Kross, F. W. Oertmann, and R. Schuhmann, "On aspherics in optical systems," *Proc. SPIE* **0655**, 300–309 (1986) [doi: 10.1117/12.938460].
3. H. Ries and J. Muschaweck, "Tailored freeform optical surfaces," *J. Opt. Soc. Am. A* **19**(3), 590–595 (2002).
4. R. Schuhmann, "Description of aspheric surfaces," *Adv. Opt. Technol.* **8**(3–4), 267–278 (2019).
5. International Organization for Standardization (ISO), *Preparation of Drawings for Optical Elements and Systems—Part 12: Aspheric Surfaces*, ISO 10110-12:2019 (Geneva: 2019).
6. ISO, *Preparation of Drawings for Optical Elements and Systems—Part 19: General Description of Surfaces and Components*, ISO 10110-19:2015 (Geneva: 2015).

7. R. N. Youngworth, S. R. Kiontke, and D. M. Aikens, "Implementing ISO standard-compliant freeform component drawings," *Opt. Eng.* **55**(7), 071205 (2016) [doi: 10.1117/1.OE.55.7.071205].
8. G. W. Forbes, "Shape specification for axially symmetric optical surfaces," *Opt. Express* **15**(8), 5218–5226 (2007).

Chapter 10
System Evaluation

Although not covered in the ISO 10110 series, understanding how to evaluate a complete optical system is important when creating component specifications. Such considerations are especially crucial when determining tolerances for an optical design. Optical components may be manufactured for a variety of systems with different applications, and therefore benefit from different requirements and specifications. The same factors are true for system evaluation metrics. Optical system metrics may vary from basic optical properties (such as focal length or field of view) to complex measurements of image quality [such as the modulation transfer function (MTF)]. Along with fundamental optical system metrics, additional aspects of the complete system may drive requirements—such as stray light, eye relief, or system transmittance. Understanding the stray light (unwanted radiation), for example, may drive mechanical aspects of the system design and influence the optical system functionality even more than, for example, surface form or other conventional optical tolerances in some applications.

There are many different ISO standards that can be used to evaluate an optical systems performance. As this book is intended as a guide to the specification and tolerancing of optical components using the ISO 10110 series on drawing notation and not *per se* about system evaluation, only basic descriptions are provided here to introduce some of the key optical system functional metrics and their evaluation.

10.1 Background

The functional evaluation of an optical system is ultimately the final determining factor when verifying the finished optical assembly meets its specification. Specifying functional metrics for optical system performance may result in confusion because of a potential difference between what is expected from the manufacturers, and the manufacturers' understanding or interpretation of how they should perform the functional testing. By having functional testing standards, the communication between fabricators and

Table 10.1 ISO/TC 172 subcommittees for specific types of optical systems. Subcommittees with system requirements discussed in this chapter are in bold.

ISO/TC 172/SC 1	**Fundamental standards**
ISO/TC 172/SC 3	Optical materials and components
ISO/TC 172/SC 4	**Telescopic systems**
ISO/TC 172/SC 5	**Microscopes and endoscopes**
ISO/TC 172/SC 6	**Geodetic and surveying instruments**
ISO/TC 172/SC 7	Ophthalmic optics and instruments
ISO/TC 172/SC 9	**Laser and electro-optical systems**

system designers can be simplified, providing the performance metric being tested and the method of how the optical system should be tested.

Evaluating optical system performance tends to vary by application. The function of a telescope is, for example, different from that of an endoscope. The experts for evaluating one type of system might not be very good at writing standards for other types of systems. With this understanding, ISO technical committee TC 172 is organized into different subcommittees. The standards that we have focused on in this book, namely the parts of ISO 10110, were developed and are maintained by subcommittee SC 1 – the group responsible for fundamental standards and to a lesser extent the materials and coatings standards developed by SC 3. Other subcommittees are in place to develop and publish standards for different types of optical systems.

The subcommittees of TC 172 are shown in Table 10.1. Since SC 7 (Ophthalmic optics and instruments) is as large as the other six subcommittees combined and is not particularly the domain of the authors, we do not discuss standards from this group. SC 4 (Telescopic systems), SC 5 (Microscopes and endoscopes), SC 6 (Geodetic and surveying instruments), and SC 9 (Lasers and electro-optical systems) all have application-specific standards for optical systems.

Many of these subcommittees have developed one or more series of standards for application-specific system requirements and performance testing. In Table 10.2, we have identified some of the more important of these standard series, but this is by no means a complete list. If you are working in one of the specific areas with such a series, your best resource for system performance specifications and testing standards can be found there.

Often, these application-specific optical performance tests are similar in terms of the optical parameter being specified and analyzed, but implementation and notations can be different.

10.2 Table of Generally Useful System Performance Standards

In addition to application-specific standards listed in Table 10.2, there are other more general standards for various system characteristics and testing

Table 10.2 Application-specific standards series related to optical properties of systems.

Application	Standards series	Title	Parts covered
Telescopes	ISO 14490 Parts 1– 9	Test methods for telescopic systems	Test methods for basic characteristics, transmittance, veiling glare, resolution, and field curvature; for astronomical and binocular telescopes and sights
Endoscopes	ISO 8600 Parts 1–7	Medical endoscopes and endoscopic accessories	Various physical and optical properties of endoscopes and various test methods
Microscope Objectives	ISO 19012 Parts 1–3	Designation of microscope objectives	Flatness of field, chromatic correction, and spectral transmission
Microscope Illuminators	ISO 19056 Parts 1–3	Definition and measurement of illumination properties	Bright field brightness, uniformity and color, and fluorescence illumination
Geodetic and Surveying Instruments	ISO 17123 Parts 1–9	Field procedures for testing geodetic and surveying instruments	Theory, levels, theodolites, electro-optical distance meters, total stations, rotating lasers, optical plumbing, global navigation satellite systems (GNSS), and terrestrial lasers
Lasers and Laser Related Equipment	ISO 11146 Parts 1–3	Test methods for laser beam widths, divergence angles, and beam propagation ratios	Test methods for stigmatic and astigmatic beams, and laser beam classification and propagation
Lasers and Laser Related Equipment	ISO 15367 Parts 1 and 2	Test methods for determination of the shape of a laser beam wavefront	Terminology and test methods for wavefront evaluation, including Shack–Hartmann sensors
Microlens Arrays	ISO 14880 Parts 1–4	Microlens arrays	Vocabulary, test methods for geometrical properties, wavefront, and other aberrations

methods. For more general applications of optics and optical systems, SC 1 maintains several system performance test standards. An example of this is ISO 9335 for measuring MTF. Some of these standards are referenced in the application-specific versions of the test methods, while others are not referenced. In addition to ISO/TC 172, the ISO technical committee exclusively for photographic systems (ISO/TC 42) contains multiple standards for optical performance evaluation that are targeted to photographic systems. These tests, however, are often applied more broadly by the user base. We have compiled a list of standards that are generally useful for optical system performance specifications and measurements in Table 10.3.

This list is not meant to be exhaustive, but a good starting point for someone looking to understand different system performance metrics and what standards may be relevant for their application.

Table 10.3 Selection of optical system standards, mostly from ISO/TC 172/SC 1, that are generally relevant to system evaluation metrics for optical assemblies and systems.

ISO Technical Committee/ Subcommittee	ISO Standard Number	Standard Title	Discussed further in
TC 172/SC 1	ISO 10110-14	Optics and Photonics – Preparation of drawings for optical elements and systems—Part 14: Wavefront deformation tolerance	10.3.2
TC 172/SC 1	ISO 9334	Optical transfer function—Definitions and mathematical relationships	10.3.3
TC 172/SC 1	ISO 9335	Optical transfer function—Principles and procedures of measurement	10.3.3
TC 172/SC 1	ISO 9336-1	Optical transfer function—Application—Part 1: Interchangeable lenses for 35 mm still cameras	10.3.3
TC 172/SC 4	ISO 9336-3	Optical transfer function—Application—Part 3: Telescopes	10.3.3
TC 172/SC 1	ISO 11421	Accuracy of optical transfer function (OTF) measurement	10.3.3
TC 172/SC 1	ISO 15529	Principles of measurement of modulation transfer function (MTF) of sampled imaging systems	10.3.3
TC 42	ISO 12233	Photography—Electronic still picture imaging —Resolution and spatial frequency response	10.3.3
TC 172/SC 1	ISO 9039	Determination of distortion	10.3.4
TC 42	ISO 17850	Photography—Digital cameras—Geometric distortion (GD) measurements	10.3.4
TC 172/SC 1	ISO 8478	Measurement of ISO spectral transmittance	10.3.5
TC 172/SC 1	ISO 9358	Veiling glare of image forming systems—Definitions and methods of measurement	10.3.6
TC 172/SC 1	ISO 13653	Measurement of relative irradiance in the image field	10.3.7

10.3 Selection of Specific System-Performance Metrics

Most optical system functional evaluations have individual standards for the respective application. While nuances vary across types of systems, the concept of what is being measured is typically constant for all optical systems. The key metrics of an optical system usually include the fundamental system parameters, (e.g., focal length and field of view) as well as some kind of imaging or resolution metric.

Breaking down each of the evaluation metrics and methods is critical in understanding both how the optical system should be analyzed, but also how it initially should be specified. Many optical systems have multiple system evaluation metrics that must be performed to confirm full system compliance for an application. Each of these tests might be performed multiple times should environmental tests need to be performed, as discussed in chapter 11.

10.3.1 Finding an appropriate standard

When specifying system performance, it can be difficult to decide which standards and specifications to apply. In addition to standards for terminology and theory, there are standards for test methods and setups in general; and then often more standards for performing a test for a specific type of optical system. What to do?

One simple strategy is: if there is an application-specific test standard, try that one first. The TCs and SCs that are focused on a specific application are able to publish and maintain very detailed standards for principles and test methods that are appropriate for that particular type of optical instrument. If there is no test standard for a particular application, it is usually best to refer to the more fundamental standards for specification and testing of optical performance.

On the ISO.org website,[1] it is possible to read the title and scope of all published standards. If the scope seems suitable for a given metric and application, then that may be the best approach. If there is no test standard for your specific application, then determine the parameter that seems most appropriate; for resolution it might be evaluating three-bar targets, measuring encircled or ensquared energy, or MTF. One can read the title and scope of any appropriate general-purpose standard related to that performance metric and determine if that method can work. If the scope seems appropriate, verify that the standard is broadly used by web searching for its number. If so, then you have a good starting point for a system performance test method. It is also important to ensure the standard is active. The ISO.org website allows searching for withdrawn standards. While seeing that some standards for performance metrics were previously written and used, ensuring that the performance metric standard you are looking for is active shows that there is continuous improvement and input across the industry.

By way of example, let us consider the resolution testing of an optical telescope assembly intended to be used on a satellite. While it is tempting to specify the telescope performance based on wavefront error (described in Sec. 10.3.2 below) this can be a difficult test to perform over a range of fields; thus a resolution test is required. Resolution in this application field is most often specified and evaluated using a spatial frequency response (SFR) or modulation transfer function (MTF).

A quick search on ISO.org for "*telescope*" and "*resolution*" yields a hit for ISO 14490-7, "Optics and Photonics—Test Methods for Telescopic Systems—Part 7: Test Methods for Limit of Resolution." This standard seems promising. The scope simply says, "This document specifies the test methods for the determination of the limit of resolution of telescopic systems and observational telescopic systems." That's not exactly what we are looking for; we were hoping for a measure of the imaging performance for a range of spatial frequencies, rather than simply the limit of resolution, and the statement about observational telescopic systems is a little off track for a satellite telescope. On reading that standard we find that the standard uses a visual assessment of a series of four-bar patterns. Since our application is a camera in space and not a user in the field, this test is probably not a good choice.

A search on ISO.org for MTF yields one relevant hit: ISO 15529. However, a search for spatial frequency response yields two more relevant hits: ISO 12233 and ISO 8600-5. Reviewing the title and scope of these standards shows that ISO 8600-5 is specifically for endoscopes, but the other standard seems appropriate for our application. Thus, we have narrowed down the list of possible ISO standards for telescope imaging performance to two relevant standards.

We can refine our list by searching the web for the specific standard numbers and determine if other manufacturers or customers have written any papers or application notes for system performance evaluations in the field of interest. Before long, a clearer picture begins to emerge of the best approach to choosing an appropriate standard.

10.3.2 Wavefront error

One typical method of testing optical system performance is the system transmitted wavefront.[2] This is an important method of system evaluation, especially prior to integration of a source or sensor, and can quickly and easily assess the quality of alignment of a system on-axis and can be used readily in laser systems. The explanation and notation for a transmitted wavefront specification using ISO 10110 Part -14 and the `13/` notation,[2] as well as an introduction to the wavefront test, are provided in chapter 3. This system evaluation method for wavefront may be used across multiple optical system types and applications. It becomes unwieldy, however, when used to evaluate performance across a range of field points.

10.3.3 Resolution

One of the primary methods of determining the optical system performance is by conducting a resolution test. These types of tests may be achieved in different ways to measure different explicit metrics but all target evaluating imaging capability and resolution. Some examples of resolution metrics are Strehl ratio, encircled or ensquared energy, spot size, optical transfer function, and MTF. Various application-specific tests exist for these metrics, but the most-common workhorse for imaging resolution is the optical transfer function (OTF)—which is the autocorrelation of the optical system's pupil function. Figure 10.1 shows the relationship between the optical transfer function and the optical system's pupil function.

The OTF is a complex function that measures the resolution of a system versus the spatial frequencies present in the object. The absolute value of the

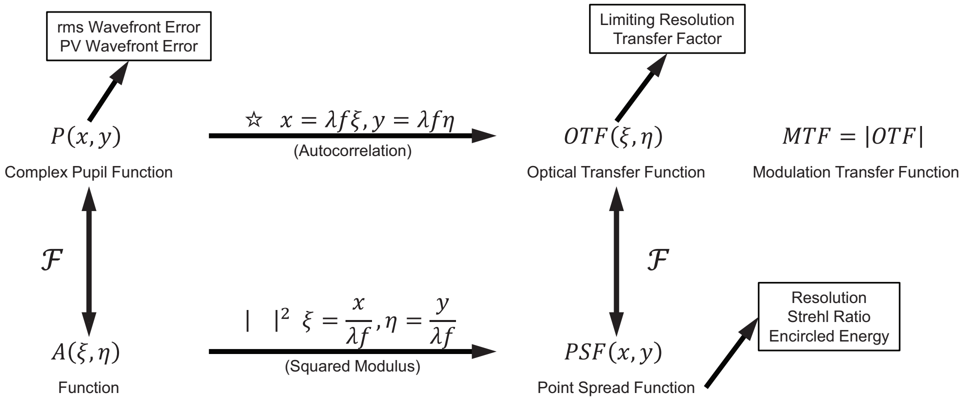


Figure 10.1 Relationship between the optical transfer function, pupil function, point spread function (PSF).[3]

magnitude of the OTF is the MTF. The phase of the OTF is the phase transfer function (PTF). When the effects of the camera and electronics are included in the MTF metric, it is referred to as spatial frequency response (SFR).[4] The resolution is sometimes referred to as the resolving power or resolution of a camera or optical system.

Evaluating the resolution of an optical system has been a topic of interest in optics since the early 19th century.[5] The earliest standard that includes a method of determining the resolving power of a lens or camera of which the authors are aware is MIL-STD-150A, first published in 1951. In that standard, a specific bar pattern was described as a series of horizontal and vertical bars in six groups in a compact spiral arrangement of three layers. The largest two groups, forming the first layer, are located on the outer edges. The smaller layers consist of repeating progressively smaller pairs toward the center. Each group consists of six elements, numbered from 1 to 6. An example of such a high-contrast test target is shown in Fig. 10.2, known ubiquitously as the 1951 USAF Resolution Test Chart.

To determine the resolving power of a system, the test target is placed at the short conjugate of the optical system and the image at the long conjugate is evaluated visually. To evaluate multiple field points, the test chart can be repeated at different locations in the field, as shown in Fig. 10.3. The resolving power of the lens at the tested field point is the smallest element of the smallest group where the three bars can be distinguished from each other.

In a photographic or digital camera system, it may be better to scale the projected resolving power test chart to an appropriate magnification, and place it against a matte background, as shown in Fig. 10.4. Or the resolution test chart can be placed at the focus of a collimator as shown in Fig. 10.5, to project the pattern to infinity, and then evaluate the photographic or digital image for resolution. If the image contrast can be measured in some

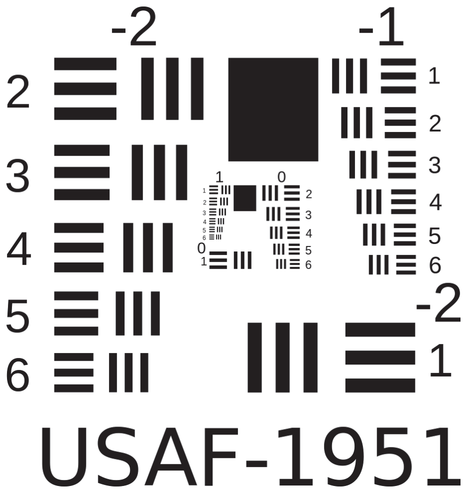


Figure 10.2 1951 USAF Resolution Test Chart.[6]

quantifiable way, it is possible to measure the contrast as a function of the spatial frequency (or spatial frequency response) rather than inspect the image for the resolution limit. With modern digital cameras this can be achieved using an algorithm or image-processing.

The biggest limitation of the 1951 USAF resolution test target is that the spatial frequencies are discrete; measurements can only be made at spatial frequencies present in the bar pattern. However, if the point spread function or line spread function (or more commonly the edge spread function) can be measured precisely, it is also possible to calculate the (2D or 1D) MTF. (Note, however, that measuring the three-bar target and MTF may not give equivalent results, as discussed by Rogers.[7]) With modern computing power, this is precisely what is done to determine the spatial resolving power of a lens over a wide range of spatial frequencies.

There are a host of international standards that help define the OTF, MTF, and SFR; and provide modern test methods for them. Some of the most

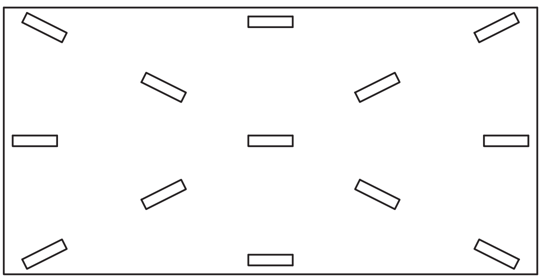


Figure 10.3 Projected resolving power test plate. The resolution test chart is repeated at thirteen locations in the field.[6]

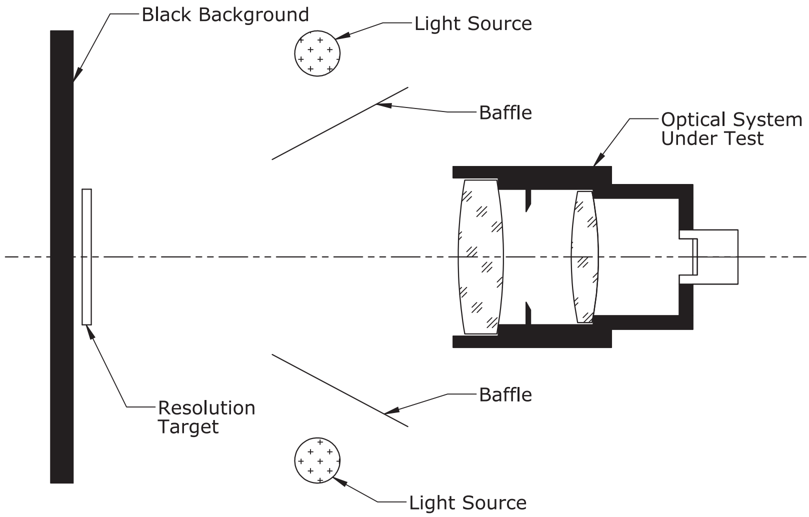


Figure 10.4 Testing a camera with the projected resolving-power test chart.

important are listed in Table 10.3. ISO 9334 provides definitions and mathematical relationships for OTF-related terms, as well as a drawing notation for each term. ISO 9335 provides guidance for test setups and methods for measuring OTF-related functions. ISO 9336-1 and ISO 9336-3

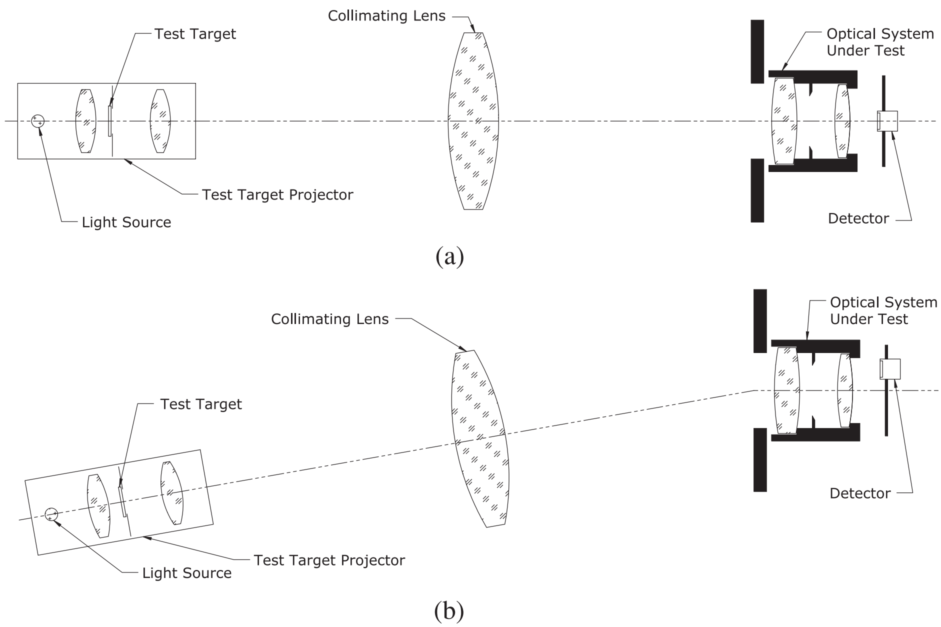


Figure 10.5 Testing a camera for use at infinity using a collimated test target at various field angles both (a) on-axis and (b) off-axis.

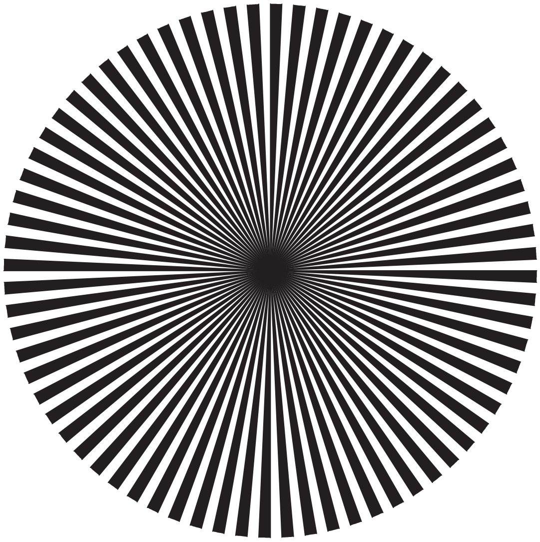


Figure 10.6 SFR test chart, often referred to as the Siemens star.

provide application-specific guidance for measuring the performance of 35-mm still cameras and telescopes, respectively. ISO 11421 provides a framework for evaluating the sources of error in an MTF measurement and recommendations for mitigating them. ISO 15529 provides an algorithm and method for calculating MTF from a line-spread function for sampled imaging systems, such as digital cameras, as well as a methodology for evaluating aliasing. Moreover, ISO 12233, the most commonly cited standard for SFR measurements of camera systems, defines a host of new test patterns designed to work with modern image processing software, as well as algorithms for calculating the SFR from either the edge spread functions or images of patterns such as the Siemens star shown in Fig. 10.6. These ISO 12233 algorithms have been incorporated into many commercial image evaluation software packages, such as IMATEST or QuickMTF, rendering this a powerful solution to specifying and testing system resolution.[8,9]

10.3.4 Distortion

Optical distortion[10,11] is an aberration that can be present within a system.[12] Distortion is typically evaluated independently from the other aberrations. Distortion is a variation of magnification across the field, which causes a mapping error for the optical system wherein the expected position for a rectilinear field is extended or compressed across the field. When magnification decreases from the center to the edge of field, the image of a rectilinear grid bends inward—taking on the appearance of a barrel.[12] If the magnification increases across the field, the image of a rectilinear grid bends

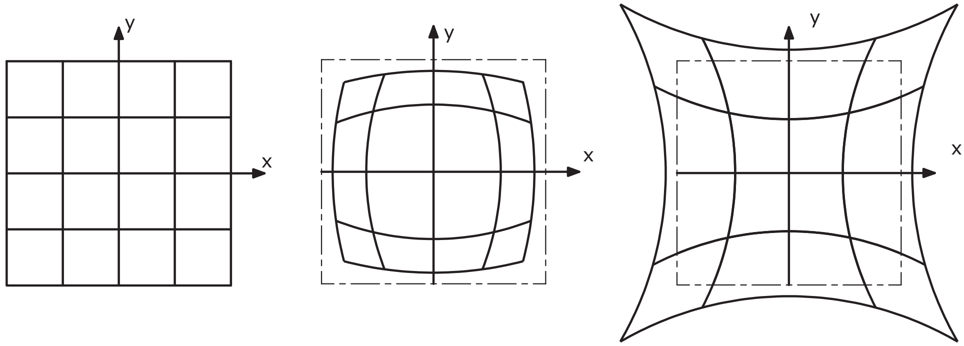


Figure 10.7 Example image out of a grid with (a) no distortion, (b) exaggerated barrel distortion, and (c) exaggerated pincushion distortion.[12]

outward, taking on the shape of a pincushion. An example of barrel distortion and pincushion distortion is shown in Fig. 10.7.

The amount of distortion can be measured in both linear units (mm or μm) for an absolute amount of distortion or a percentage for a relative amount of distortion. Assuming the optical system being evaluated for distortion operates with a finite conjugate object distance and a finite conjugate image distance with a modest field of view, the amount of distortion can be calculated based on Eqs. (10.1) and (10.2). In both instances, the paraxial lateral magnification is necessary for these calculations. The paraxial lateral magnification is based on Eq. (10.3)

$$D_{rel} = \left[\frac{\left(\frac{h'}{h}\right) - m}{m}\right] * 100 \tag{10.1}$$

$$D_{abs} = h' - hm \tag{10.2}$$

$$m = \frac{h'_p}{h_p}, \tag{10.3}$$

where:

h_p = paraxial object height,
h'_p = paraxial image height,
m = paraxial lateral magnification,
D_{abs} = absolute distortion in linear units,
D_{rel} = relative distortion in percent,
h = object height, and
h' = image height.

A few different methods can be used to evaluate system distortion, as there are three different classes of conjugate configurations for an optical system: infinite-to-finite, finite-to-finite as described above, and infinite-to-infinite (afocal). A fourth, finite-to-infinite, is optically equivalent to infinite-to-finite. The optical system configuration determines the method of testing that best suits the distortion evaluation. While all of the tests are similar in some aspects, they are all specific to their own optical system configuration.

For the finite conjugates distortion measurement, as specified in ISO 9039, the optical system is mounted relative to both a light source and a photodetector. To achieve a line illumination a slit should be placed in front of the illumination source. This setup provides a clean target for the optical system to image.

There are two possible setups that can be used when performing this distortion test. In the first setup, the optical system is mounted such that the object source can rotate about the entrance pupil of the optical system. This arrangement enables the source to achieve the angular range of motion necessary. The detector is mounted on a linear device such that the light that reaching the image height can be found in the image plane. An example of this type of test configuration is shown in Fig. 10.8.

In the second mounting configuration—i.e., for testing finite conjugate distortion testing—instead of the object and detector moving across the field, the optical system under test moves. When using this method, it is necessary to still move the optical system about the entrance pupil. As per the standard, this movement is accomplished with a nodal slide. This type of test configuration is shown in Fig. 10.9.

In any of the configurations that include an infinite object or image, use of a collimator is necessary to project the object and/or image to infinity. When the object side is at an infinite conjugate, a collimator lens is placed between the slit and the system under test, one focal length from the object slit. When the image side is an infinite conjugate, a collimator lens is placed between the system under test and the sensor, with the focus of the collimator at the detector in the image plane.

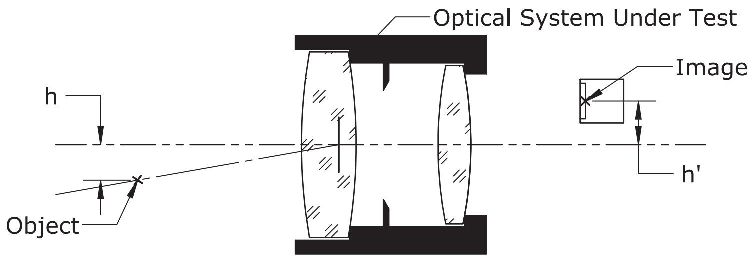


Figure 10.8 Layout for a distortion test on an optical bench with an object coming from a finite object distance to a finite image plane, where the object and image points are rotating.[10]

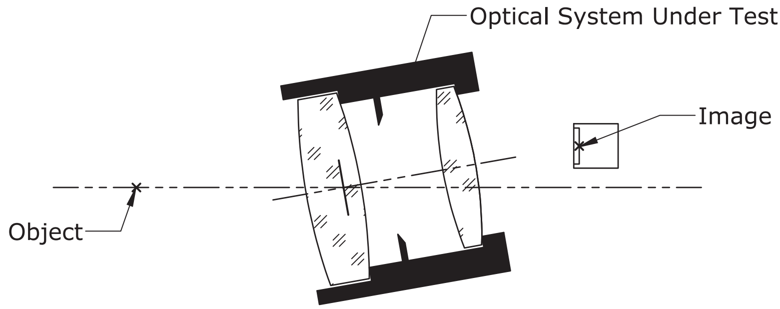


Figure 10.9 Layout for a distortion test on an optical bench with the object coming from a finite distance to a finite image plane, where the optical system is rotating.[10]

Specific to photographic optical systems, a separate standard is in place for measuring the amount of distortion with a single target instead of sampling with an illumination source and a slit. In ISO 17850 a test target is used to image a pattern of dots or lines directly.[11] Figure 10.10 shows an example of these sample test targets.

When performing the direct measurement of distortion, additional considerations must be made for the test setup. These include the illumination of the target and the position of the optical system relative to the test target. The light sources should cause a uniform illumination, to within ±10%.

The image of the test target should fill the field of view for the optical system to ensure that the distortion is properly measured. Once images have been taken of the test targets, post-processing is necessary to calculate the amount of distortion present in the images and thus the optical system.

Note that this entire discussion of system optical distortion has been in the context of a rectilinear system that is intended to have a uniform

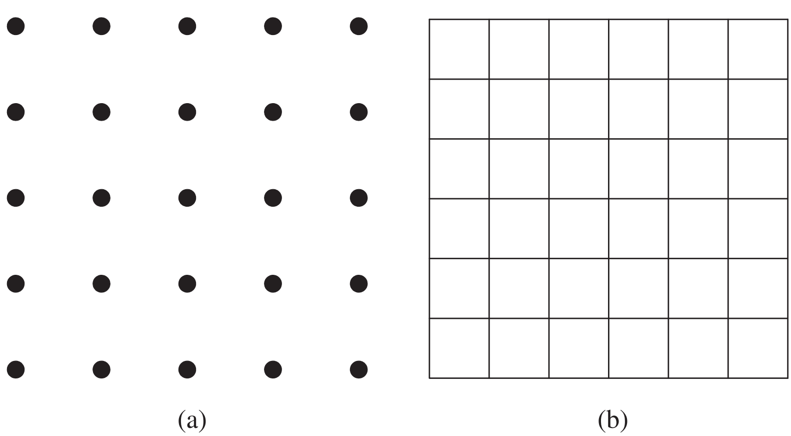


Figure 10.10 Test targets used for a direct distortion test, as per ISO 17850. The test targets are either (a) a grid of dots or (b) a line grid pattern.[11]

magnification across the field of view. For the case where the object is at infinity and the image is at a finite conjugate, and the magnification is uniform across the field, the height of the image is related to the angular position of the object point by trigonometry; that is:

$$h = f\,tan(\theta) \tag{10.4}$$

where:
h = object height,
f = focal length, and
θ = object angle.

A uniform magnification across the field is not always preferred. It is sometimes desirable to have a magnification that is decreasing toward the edge of field, such as with fisheye lenses or scan lenses. In these cases, different field mapping functions may be desired. The most common alternative to the *f*–tan–theta requirement is the so-called *f*–theta requirement, where the image height is instead linear with the field angle.

$$h = f \cdot \theta \tag{10.5}$$

In *f*–theta lenses, distortion is measured against the intended linear mapping with angle, rather than against a rectilinear (trigonometric) mapping.

10.3.5 Transmittance

With optical systems, the ratio of the amount of intended light that reaches the output of the system to the amount of light that enters the optical system, or throughput, is the transmittance.[13–15] The loss of optical system transmittance can be caused by multiple parameters on each optical element and the system itself. The primary property of an optical element that affects the system transmittance is the reflectivity of each of the surfaces in the optical system. This reflectivity, in turn, is usually a function of the optical coating on each surface. For a transmissive optical element, such as a lens or window, the amount of reflection is nominally based on the amount of Fresnel reflection.[12] The Fresnel reflection of a surface is governed by the Fresnel equation, Eq. (10.6), and is a function of the refractive index on either side of the interface. Without an optical coating, assuming elements with an index of 1.5 immersed in air, the amount of reflection from each optical surface is approximately 4%. For an optical system, the amount of transmittance can decrease significantly with multiple optical elements in place, necessitating antireflection coatings.

$$R = \left(\frac{n_1 - n_2}{n_1 + n_2}\right)^2 \tag{10.6}$$

where:
R = amount of reflection for an optical surface,
n_1 = incident refractive index, and
n_2 = refracted refractive index.

Along with the transmittance loss in an optical system based on surface reflections, optical glass has an innate amount of bulk material absorption. The bulk material transmittance is based on the Beer–Lambert law, Eq. (10.7).

$$\tau_{glass} = 1e^{-kd} \tag{10.7}$$

where:
τ_{glass} = bulk transmittance through an optical material;
k = absorption coefficient for the respective optical material, in 1/m and
d = distance or thickness of bulk material light is traversing in m.

The bulk material transmittance is based on the amount of light that is not absorbed within the optical element as light passes through each optical element. Although the amount of bulk material absorption may be small in most cases, it still has an impact on the overall optical system transmittance at extreme wavelengths.

Transmittance for an optical system ultimately impacts the amount of light needed to ensure the optical system has sufficient signal. Tests for measuring the amount of transmittance for an optical system are shown in three different ISO standards: ISO 8478, ISO 14490-5, and ISO 19012-3.[13–15] All three of these standards discuss the basic concept of system transmittance for an optical system for specific applications. The difference between each of the ISO standards relates to how the information should be presented for the different types of optical systems: camera lenses, telescopic systems, and microscopes, respectively.

Wavelength-dependence is typically critical for system transmittance. In the test setup for transmittance of an optical system, either a monochromator or set of optical filters is used to narrow the wavelength band of the light source and provide higher fidelity in testing for the spectral dependence of system transmittance.

A layout of a transmittance test configuration for a generic optical system is shown in Fig. 10.11. In this figure, a light source is used to create a broad spectrum of light. This light then goes through an illumination optical system to ensure spatial features from the light source are mitigated. Along with the illumination source, either a monochromator or an optical filter is used to

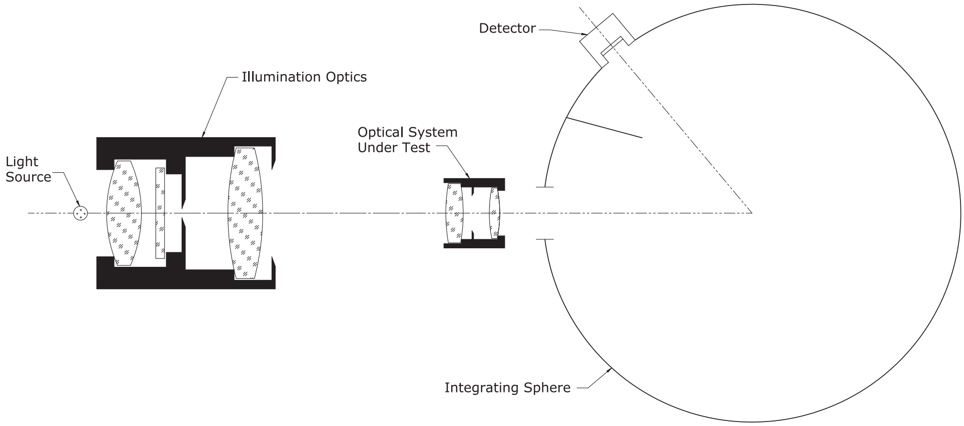


Figure 10.11 Layout for an optical transmittance test.[13,15]

select the wavelength under test. Light then goes through an aperture stop to ensure that the pupil of the optical system under test is just barely filled, limiting stray light from outside the optical system into the test. After the light passes through the optical system in question, the light then goes to either a detector or an integrating sphere. The method for determining how much light is actually transmitted through the optical system requires a two-step process. In the first step, the optical system is removed to establish a calibration measurement or baseline. Then the optical system is put into the test configuration. The system transmittance is calculated as the difference between the amount of light that goes from the illumination source to the detector, with and without the optical system in place.

10.3.6 Stray light

One system quality metric that is often not measured is the amount of stray light,[16,17] or unwanted radiation, present in an optical system. Stray light is effectively unwanted light that reaches the image plane or detector.[18] There are many different potential causes for stray light. These may include ghost reflections from optical elements, scatter from surfaces, or unblocked paths from sources outside the field of view that reach the detector. Stray light mitigation methods for a complete optical system may require specific tolerances or coatings on individual optical elements as well as mechanical controls to block unwanted pathways. Mitigation measures on optical elements may include tight antireflection coatings, minimized surface roughness, and tightened tolerances on surface imperfections and bubbles and inclusions.

Figure 10.12 Example picture showing ghost reflections toward the bottom left of the image.

As an example, light reflecting from the surface of an optical element that does not have the proper coating to mitigate stray light may become a ghost reflection in the system. In these instances, additional spots or light streaks may appear on the image plane. An example of how ghost reflections for light sources may appear is shown in Fig. 10.12.

From a system evaluation standpoint, stray light compliance is often determined by analysis rather than test. When necessary, stray light can be measured in different ways, depending on the application. This test, referred to as the veiling glare test, is described in both ISO 9358 and ISO 14490-6.[16,17] It is important to be aware that the terms "*stray light*", "*ghosts*," and "*veiling glare*" have different meanings for different applications; some practitioners would call this test a stray light test, and have a different method for evaluating "*veiling glare*" or "*ghosts*". In both of these standards, the method of testing for veiling glare is to have the optical system pointed into an integrating sphere and looking at a black object. The black object is chosen to "fill" the nominal aperture and field of the lens, creating a black spot at the image plane because no image forming illumination is present. The result is that any radiation at the image plane during the test is unwanted radiation. As the optical system is looking into an integrating sphere, light from all angles outside of the nominal field of view illuminates the interior of the optical system. Fig. 10.13 shows an approximate layout of the veiling glare test.

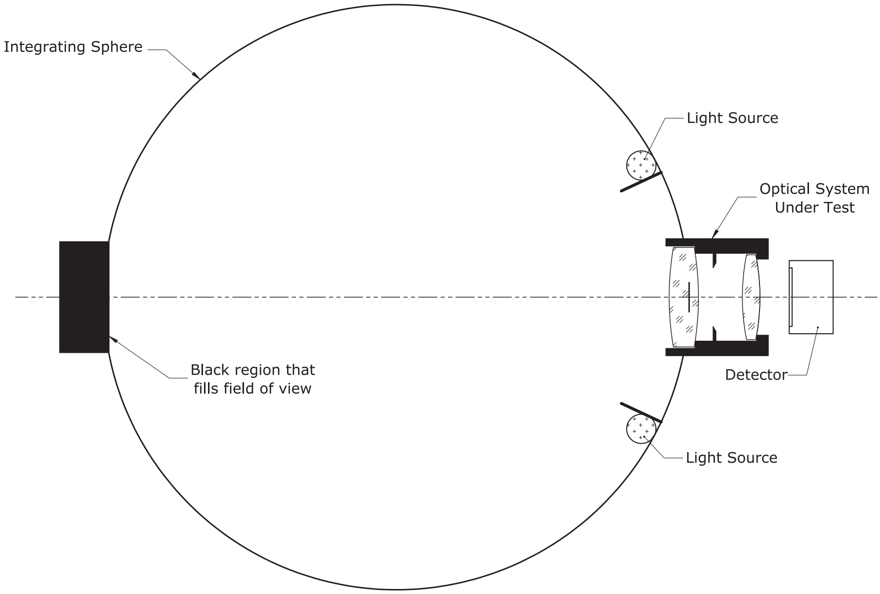


Figure 10.13 Layout of a veiling glare test with an integrating sphere.

To determine how much unwanted radiation is entering the optical system, a detector is placed at the image plane of the optical system. There are multiple quality metrics that can be used to quantify the amount of stray light with a veiling glare test, such as the veiling glare index (VGI) or the glare spread function (GSF), described in ISO 9358. These tests are performed both with the optical system pointed directly into the integrating sphere as well as when the optical system is rotated about the entrance pupil.

10.3.7 Relative illumination

Differences in the radiometric properties across the field of view can induce different illumination properties (such as irradiance) across the field.[19,20] For example, the pupil when viewed from an extreme field point can appear compressed in one axis, typically falling off as the cosine of the field angle. The optical path from the detector to the exit pupil is often longer for higher field positions, further reducing illumination. If the system is not telecentric, an additional compression of the illumination can occur due to the angle of incidence of the principal ray (at the full field the chief ray) at the detector. All of these geometric differences between on- and off-axis field angles can result in a falloff in illumination at a rate of $\cos^3(\theta)$ or even $\cos^4(\theta)$. In addition, effects from optical system metrics such as distortion or pupil aberrations may result in better or worse illumination as the image is formed off-axis relative to the on-axis transmittance. Vignetting, stray light, and field curvature can also

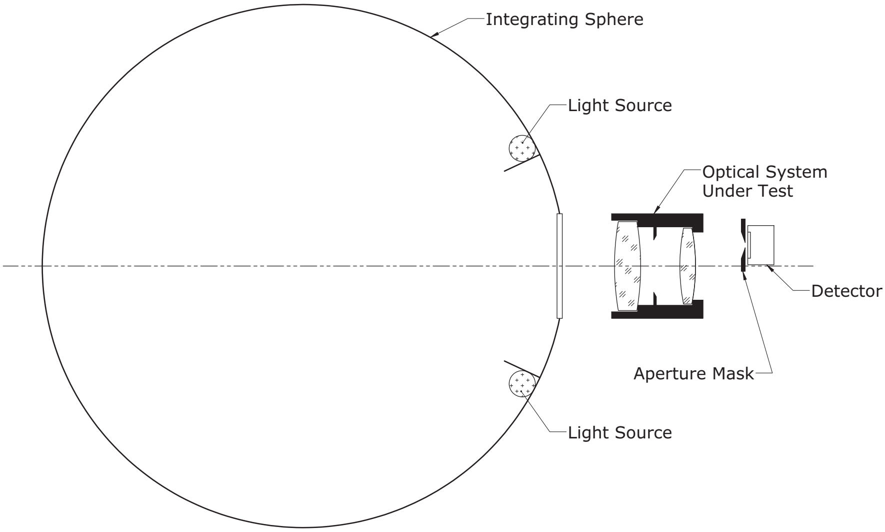


Figure 10.14 Layout of a relative illumination test with a finite conjugate image.

impact relative illumination. Differences in relative illumination results in a dimmer image at some points in the field, degrading signal to noise, contrast, and MTF.

The test method and procedure for performing a relative illumination test is provided in ISO 13653. Performing a relative illumination test can be performed on rotationally invariant and off-axis systems, depending on the mounting conditions of the optical system. The testing methods discussed in ISO 13653 ignore the object space conjugate and assume that the image conjugate is either finite or infinite. Because of this condition, the optical test may change the intended object conjugate and image conjugate. Accordingly, the relative illumination test produces a Lambertian object that fills the entrance pupil of the system. A pinhole is placed at the image position to ensure that the relative illumination achieved is for a discrete image position. This pinhole detector is then traced across the image plane linearly. For an infinite image conjugate, the pinhole is swept across the image field angularly. An example of this test condition is shown in Fig. 10.14 for a finite image conjugate.

After the test is performed for multiple points in a line from the center to the corner of the field, the irradiance or power across the field is calculated relative to the on-axis image. Systems without rotational symmetry about the axis of the entrance pupil—e.g., unobscured reflective systems—requires multiple lines of field points in different orientations to accommodate potential asymmetry in the relative illumination across the field.

10.4 Indications

Unlike ISO 10110, most of the system evaluation metrics do not have a designated drawing notation. As optical systems vary across many different types of applications, as well as different performance metrics, having singular values for an indication would be unwieldy. Often, these metrics are listed in a requirements table in the system specification or an interface control drawing (ICD), rather than on the component drawings themselves. The advantage of using the ISO system performance specifications listed here is that they provide a detailed context for indicating the required performance and how the performance shall be evaluated. In many cases the specifications inform both the customer and manufacturer how each test should be performed, reduce conflict, and improve quality control. Testing setups, methods, evaluation, and final reporting of the optical system performance are generally listed in each respective standard. In most cases, communication between the customer and fabricator is necessary to determine the appropriate tests and metrics for these specifications—but use of international standards provides a framework to streamline those discussions.

References

1. International Organization for Standardization (ISO) s.v., "International Organization for Standardization," accessed September 18, 2021, https://www.iso.org/search.html.
2. ISO, *Preparation of Drawings for Optical Elements and Systems—Part 14: Wavefront Deformation Tolerance*, ISO 10110-14:2018 (Geneva: 2018).
3. J. E. Harvey and A. K. Thompson, "Scattering effects from residual optical fabrication errors," *Proc. SPIE* **2576**, 155–174 (1995) [doi: 10.1117/12.215588].
4. J. D. Gaskill, *Linear Systems, Fourier Transforms & Optics*, University of Arizona (Tucson, Arizona: 2011).
5. G. B. Airy, "On the diffraction of an object-glass with circular aperture," *Trans. Cambridge Philos. Soc.*, **5**, 283–291 (1835).
6. US Department of Defense, *Military Standard: Photographic Lenses*, MIL-STD-150A (Arlington, Virginia: 1959).
7. J. R. Rogers, "Three-bar resolution versus MTF: how different can they be anyway?," *Proc. SPIE* **7071**, 707109 (2008) [doi: 10.1117/12.800403].
8. Imatest, "Imatest Master" (2021).
9. Quick MTF, "Quick MTF" (2019).
10. ISO, *Quality Evalutation of Optical Systems—Determination of Distortion*, ISO 9039:2008 (Geneva: 2008).
11. ISO, *Digital Cameras—Geometric Distortion (GD) Measurements*, ISO 17850:2015 (Geneva: 2015).

12. J. E. Greivenkamp, *Field Guide to Geometrical Optics*, SPIE (Bellingham, Washington: 2003).
13. ISO, *Test Methods for Telescopic Systems—Part 5: Test Methods for Transmittance*, ISO 14490-5:2020 (Geneva: 2020).
14. ISO, *Designation of Microscope Objectives—Part 3: Spectral Transmittance*, ISO 19012-3:2015 (Geneva: 2015).
15. ISO, *Measurement of ISO Spectral Transmittance*, ISO 8478:1996 (Geneva: 1996).
16. ISO, *Test Methods for Telescopic Systems—Part 6: Test Methods for Veiling Glare Index*, ISO 14490-6:2014 (Geneva: 2014).
17. ISO, *Veiling Glare of Image-Forming Systems—Definitions and Methods of Measurement*, ISO 9358:1994 (Geneva: 1994).
18. E. C. Fest, *Stray Light: Analysis and Control*, SPIE (Bellingham, Washington: 2013).
19. ISO, *Measurement of Relative Irradiance in the Image Field*, ISO 13653:2019 (Geneva: 2019).
20. ISO, *Definition and Measurement of Illumination Properties—Part 1: Image Brightness and Uniformity in Bright Field Microscopy*, ISO 19056-1:2015 (Geneva: 2015).

Chapter 11
Environmental Testing

Outside of the functional testing that occurs on an optical system, considerations must be made for the operating conditions of the optical system as well. Environmental conditions such as temperature and humidity can become critical when taking the optical system out from the lab. While not a direct aspect of ISO 10110, listing the environmental conditions for an optical system can be important when specifying a final assembly. This is especially true for systems with dielectric or metallic coatings that are exposed to the environment when used.

The ISO 9022 series is a group of test standards that define the environmental tests and levels at which a system or component will be tested. Most of the standards in the ISO 9022 series are for specific different environmental conditions. There are also three, more rigorous ISO 9022 standards that combine the individual environmental conditions that the optical system may truly experience.

In each part, a series of conditions, termed "*conditioning methods*," are described. These conditioning methods are identified by a numeric "*condition code*." To specify a test, the condition code for that test must be identified as well as the "*severity code*" and the operating state of the instrument being tested. The method of indicating an environmental test on an optics drawing is discussed more in Sec. 11.7.

In addition to the ISO 9022 series of standards, there is another standard (ISO 10109) that provides guidance for environmental conditions an optical system may experience depending on a specific application or global location. This standard lists expected temperature ranges, humidity, pressure, and rain conditions for a series of standard environments. These standard environments are based on global data collected and are meant to guide the user on which parts of ISO 9022 may be applicable for a given situation. These environmental conditions are in place to assist the user if temperature, humidity, and pressure are not known but the type of environment is known. The standard also includes a framework (ISO 10109, Table 9) for identifying and prescribing the environmental testing that may be required for the system, and also serves as a checklist for planning environmental tests.

11.1 Background

When considering an optical system, the functional evaluation metrics considered are typically for the system in a fairly benign indoor environment. This environment is listed in ISO 10109 as standard environment 5, but the temperature range (15°C to 35°C) and relative humidity range (0% to 85%) may seem rather broad compared to a thermally controlled lab. Many optical systems must work outside the lab and sometimes in harsh environmental conditions. Taking the jump from lab evaluation to real-world environments may require acceptance confirmation by additional testing. The ISO technical subcommittee that oversees fundamental standards for optics and photonics (SC1) has a separate working group (WG3) dedicated to environmental test methods. The environments discussed in this working group can range from a temperature- and humidity-controlled indoor location to an environment as extreme as Antarctica.

Understanding the specific environmental conditions that an optical system might experience is critical. These conditions must be conveyed to ensure that the evaluation metric will be met. Each potential environmental test is broken out individually in the ISO 9022 series. Many parts of this series have been consolidated to streamline the standard. Parts −10, −13, −15, −16, and −19 were combined into a new Part −22; while Parts −5, −18, and −21 were combined into a new Part −23. These subsections from the ISO 9022 series are listed in Table 11.1. Parts that are struck through have been withdrawn and replaced as indicated.

The definition of these environmental tests may be known through documented conditions and conveyed as a requirement. Additionally, ISO 10109 is in place to assist the user in defining some environmental conditions if they are not known. ISO 10109 contains an overview of specific environmental conditions that an optical system may experience from an indoor technical environment to various conditions such as "*non-weather protected maritime*" or "*high altitudes of up to 30000 meters*."

11.2 Testing Sequence and Evaluation

During environmental testing, a test procedure must be followed to ensure the validity of the final acceptance criteria. The test sequence, as outlined in ISO 9022-1, is relatively simple. There are three states of operation that an optical system may be under for an environmental test. These vary from the optical system being in storage or shipping conditions (state `-0`), to the system exposed to the environment but without power (state `-1`), to the optical

Table 11.1 List of standards in the ISO 9022 environmental test method series. Parts struck out have been withdrawn.

Standard Number	Standard Name
ISO 9022-1	Part 1: Definitions, extent of testing
ISO 9022-2	Part 2: Cold, heat and humidity
ISO 9022-3	Part 3: Mechanical stress
ISO 9022-4	Part 4: Salt mist
ISO 9022-5	Part 5: Combined low pressure and cold (replaced by ISO 9022-23)
ISO 9022-6	Part 6: Dust
ISO 9022-7	Part 7: Resistance to drip or rain
ISO 9022-8	Part 8: High internal pressure, low internal pressure, immersion
ISO 9022-9	Part 9: Solar radiation and weathering
ISO 9022-10	Part 10: Combined vibration and dry heat or cold (replaced by ISO 9022-22)
ISO 9022-11	Part 11: Mold growth
ISO 9022-12	Part 12: Contamination
ISO 9022-13	Part 13: Combined shock, bump, free fall with dry heat or cold. (replaced by ISO 9022-22)
ISO 9022-14	Part 14: Dew, hoarfrost, ice
ISO 9022-15	Part 15: Combined random vibration with dry heat and cold (replaced by ISO 9022-22)
ISO 9022-16	Part 16: Combined bounce or acceleration with dry heat and cold (replaced by ISO 9022-22)
ISO 9022-17	Part 17: Combined contamination, solar radiation
ISO 9022-18	Part 18: Combined damp heat and low pressure (replaced by ISO 9022-23)
ISO 9022-19	Part 19: Thermal cycles combined with vibration (replaced by ISO 9022-22)
ISO 9022-20	Part 20: Humid atmosphere containing sulfur dioxide or hydrogen sulfide
ISO 9022-21	Part 21: Combined low pressure and ambient temp. or dry heat (replaced by ISO 9022-23)
ISO 9022-22	Part 22: Combined cold, dry heat or temperature change with bump or random vibration
ISO 9022-23	Part 23: Low pressure combined with cold, ambient temperature and dry and damp heat

system in full operation during the environmental test (state −2). Each of the three operational states may present slightly different test sequences due to the assessment and testing required.

A generic outline of the test sequence for an optical system is as follows:

- Prepare the optical system for testing
- Perform a pre-test inspection
- Run the environmental test (often termed "*conditioning*") at the required severity
- Allow the optical system to come to equilibrium
- Test for the acceptance metric and confirm results

The preparation of an optical system includes inspection and cleaning to establish how the system would be received by a customer. Wiping down the system before the testing sequence should be performed as certain testing facilities may not be in a clean room environment.

The other aspects of the test sequence validate the optical system acceptance metric by isolating the environmental test in question. Performing multiple environmental sequences or conditions in a row should include optical system acceptance testing in between each environmental condition. The root cause of a potential failure may not be known if the intermediary testing steps are not taken.

When evaluating an optical system that has gone through an environmental test sequence, a simple visual inspection is performed to determine if major failures have occurred. Major failures may include components falling off the optical system or optical components breaking. Performing a basic evaluation is important in minimizing additional work in the testing process.

The acceptance metric chosen for a given environmental test is the metric or metrics that are expected to be most affected by the environmental condition being tested. This might be total transmission, wavefront error, or some other metric depending on the system, application, and environmental conditions.

11.3 Indications

For optical environmental specifications, a test code designation is necessary. Each code is listed with the designation shown below and is consistent across all types of environmental tests. This understanding allows the requirement to state their test conditions in a single list without ambiguity in their specifications.

The coded notation is different than in an ISO 10110 drawing. Each of the terms for an environmental test is specific to an environmental test condition and severity level. The notation to specify and environmental test is shown below with the explanation of each term:

`Environmental Test ISO 9022 –AA –BB –C`

where:
`AA` = conditioning code,
`BB` = severity level, and
`C` = state of operation during the test.

The conditioning code indicates the specific environmental condition that should be tested. There are 35 different environmental test conditions that can be specified with the ISO 9022 series. Some of the environmental conditions that may be specified are within the same group; e.g., hot and cold environments within the temperature standard. There are an additional 10 environmental test conditions within the ISO 9022 series that are a combination of two or more different types of conditions to be tested concurrently; e.g., pressure and temperature.

To assign the intensity at which a test should occur the degree of severity must be listed with the environmental test condition. These degrees of severity are specific to each type of test and may vary from a fairly benign test to an extreme test.

The state of operation for the optical assembly must be listed. As discussed in Sec. 11.2, these are the state in which the optical system should be kept in during the environmental test. States of operation are `−0` (in its shipping or storage container), `−1` (unpowered but ready for operation), and `−2` (fully operational).

An example for assigning an environmental test condition would be as follows

```
Environmental Test ISO 9022 −10 −01 −2
```

This specification states that the environmental test conditions are for a cold environment with a low degree of severity in an operational state. The specific severity that this optical system would experience is maintaining a

Table 11.2 Description of information required in an environmental specification.

Information Required	Description
Test Code	This is the test code described at the beginning of Sec. 11.5. This information must include the conditioning method, severity level, and state of operation.
Number of Specimens	The list of optical systems being tested, or in some cases, the number of witness samples under test.
Testing or Chamber Information	In many of the environmental test conditions, a requirement is placed on the optical system or witness sample as to where in the test chamber they shall be placed. In these same cases, the associated evaluation equipment location must be understood. As an example, when performing a thermal environmental test, the location for the temperature sensors must be present to clearly understand what occurred during the environmental test.
Preconditioning Method	In nearly all the subsections in the ISO 9022 series, a preconditioning method is described. This information is generally cleaning of the optical system or witness sample and setting the test chamber to a beginning test condition. For instances when the preconditioning method is not described in the relevant environmental test standard, the type of preconditioning used will need to be noted.
Type and Scope of Initial Test	Prior to the environmental test that will be performed, an initial test must be conducted. This is intended to get a baseline of the system test, or visual description in certain cases, before any environmental testing.
Recovery Method	After the environmental test is performed a recovery condition is present to bring either the optical system or witness sample to an ambient condition ready for evaluation. In nearly all the subsections in the ISO 9022 series, a recovery method is described. If the recovery method performed is not what is listed in the relevant ISO 9022 standard, the recovery method must be noted.
Type and Scope of Final Test	Testing that will be performed following the environmental test. This information is critical for the customer to understand the environmental effects on the optical system or witness sample.
Criteria for Evaluation	As described below, this information is a gauge for whether the optical system or witness sample passed the environmental test condition.
Type and Scope of Test Report	The method in which the environmental test must be reported back to the customer. This may be as simple as the notation of the environmental test code, technical requirement, and status after the test.

Table 11.3 Notation for status after test notation.[1]

Status After Test	Comment
A	All performance criteria are satisfied
B	All performance criteria are satisfied. Damage to parts not needed for function or reduced life is possible
C	Not all performance criteria are satisfied. Damage to parts not needed for function or reduced life is possible
D	Device may not operate anymore; damage expected

cold environment at 0°C for a 16 hour period. During that time duration, the optical system will be checked for operation.

For each of the types of environmental testing described in Sec. 11.5 a complete specification must be created that includes the acceptance metric. All conditioning methods have a similar outline for what must be included in this report. There is no guidance in how these reports must be created, but the information they convey is standardized. These reports all must include the information listed in Table 11.2. This information is what must be reported after the environmental test is completed.

When performing an environmental test, the criteria for how the optical system or witness sample must be evaluated are critical. In certain instances, the test may damage the optical system, but this is acceptable. Because the status after the test may vary, a notation is used for each of the environmental test conditions. This is a simplified method of describing what is expected from the environmental test. The notation for these status conditions is listed in Table 11.3.

11.4 Overview of Environments

While considering environmental conditions outside of the lab, understanding the type of environmental tests necessary for an optical system may seem daunting. As there are 11 independent environmental conditions within the ISO 9022 series, breaking down which conditions and what severity to use is not a simple task. There are three additional parts in the ISO 9022 series that have combinations of certain environmental tests, instead of testing each independently. ISO 10109 is in place to try and assist in defining conditions for global environments. The environmental conditions listed in ISO 10109 are suggestions based on historic data evaluated from various regions. In addition to the 35 conditions described in the various parts of ISO 9022, ISO 10109 also references IEC 60129 for precipitation and dust, IEC 60068 for pollution, and place-holders without any test reference for seven other conditions; for a total of 64 conditions to be considered. Looking at this long

Table 11.4 Overview of standard environments listed in ISO 10109. Environments listed in italics are planned for the next revision.[1]

Name	Temperature	Humidity
Indoors	+15°C to +35°C	Up to 85%
Outdoors	−33°C to +55°C	Up to 100%
Outdoors (Limited)	−20°C to + 50°C	Up to 100%
Dry Hot	*−10°C to +55°C*	*Less than 40%*
Humid Hot	*+10°C to +55°C*	*Up to 100%*
Cold	*−45°C to +45°C*	*Up to 100%*
Extreme Cold (Arctic)	−65°C to +35°C	Up to 100%
Coastal	−20°C to +35°C	Up to 100%
High Altitude	−65°C to +55°C	Up to 100%
Ultra-High Altitude	*−145°C to +55°C*	*Up to 100%*

list in ISO 10109's Table 9 can lead one to despair. Obviously, not all systems must undergo all of the various tests; the time and expense of even a fraction of these tests can be exorbitant. However, some applications—such as expensive, one-off space telescopes; or first article testing of a medical instrument or rifle scope to be made in a high volume—may justify more-extensive test regimes than others. A risk analysis can provide some insight into which tests should be conducted and which should not. In many cases only half a dozen conditions must be validated for a given product.

There are six different types of environmental climates described in ISO 10109, with four more planned for the next revision. An overview of these climates is shown in Table 11.4 along with their temperature and humidity ranges. The environmental testing that is performed for these various regions may be necessary for some optical systems.

Each of the environments listed as a standard environment in ISO 10109 will have a different series of tests that can be performed to validate the functional acceptance metric for an optical system. The information presented in ISO 10109 is meant as a guide when selecting environmental tests for an optical system.

11.5 Test Methods

Across the various subsections of the ISO 9022, there are 35 different independent test conditions that may be applied, broken down into various categories by part. There are another 10 test conditions that may be applied that are combinations of some of the independent test conditions. In this section, we describe the various environmental test conditions, more or less in priority order for typical optical systems.

11.5.1 Temperature and humidity[2]

ISO 9022-2 deals with temperature and humidity. Temperature and humidity effects on an optical system may be detrimental to long-term operation. Temperature and humidity requirements are among the most common environmental specifications. Environmental testing for thermal effects may be as simple as testing to see that a system could withstand shipping parameters to operating an optical system over a wide temperature range, or as complex as a system required to operate in extreme heat or cold.

There are differences in the environmental conditions in ISO 9022-2. The first group of thermal tests (conditioning codes −10, −11, −12, and −13) is performed at a steady-state temperature and evaluated over a duration. The second group of thermal tests (conditioning codes −14, −15, and −16) is across a cycled temperature environment at different ramp speeds. Each environmental condition, and its conditioning code, is listed in Table 11.5.

The degree of severity for the four steady-state conditions relate to both the chamber temperature at which the optical system must remain and the exposure time. Setting each steady-state condition requires a ramp-up and ramp-down period for the chamber to achieve the thermal levels necessary. Without a consistent and appropriate ramp period the optical system may experience a shock, potentially breaking the optical system.

The second group of three environmental conditions listed in ISO 9022-2 are for cycling the temperature for an optical system. The cycling severities are

Table 11.5 Conditioning methods for temperature and humidity.

Conditioning Code	Conditioning Method
−10	Cold
−11	Dry Heat
−12	Damp Heat
−13	Condensation
−14	Slow Temperature Change
−15	Temperature Shock
−16	Damp Heat, Cycle

Table 11.6 Sample degrees of severity for conditioning method 16.[2]

Degree of severity	01	02	03	04	05	06	07
Cycling climatic conditions	23 °C ± 2 °C and 80% to 85% r.h. 40 °C ± 2 °C and 90% to 95% r.h. including condensation			23 °C ± 2 °C 55 °C ± 2 °C		23 °C ± 2 °C 70 °C ± 2 °C	
Number of cycles	5	10	20	5	10	5	10
State of operation	0 or 1 or 2			0 or 1			

based on the temperature range, relative humidity, ramp speed, and number of cycles. For these conditions, the ramp period is stated in the condition and may be intended to force the optical system to experience a shock.

An example of a severity table is shown in Table 11.6. In this instance, the conditions for the various severity levels for cycled damp heat are listed. This includes the thermal conditions for each of the cycles as well as the number of cycles and other important details. Within each part of ISO 9022, a detailed description of the test conditions is provided. Other degrees of severity for this particular conditioning method have different thermal conditions for cycling.

11.5.2 Mechanical stress[3,4]

The mechanical stress tests included in ISO 9022-3 are acceleration, drop, and vibration effects (Table 11.7). These types of environmental conditions may cause something to move in an optical system and shake out of place. Along with temperature and humidity, the mechanical stress tests are a common requirement. The testing is based on the type of environment that a system will experience. For example, an environmental condition of free fall would occur when the optical system, in its packaging, falls during shipping.

The testing for conditions `−30`, `−31`, and `−35` are based on a constant change in acceleration. The shock environmental condition (`−30`) is an acceleration of the optical system once at a high acceleration. The bump condition (`−31`) is similar to shock but occurs many times, either 1,000 or 4,000, at less-severe amplitude accelerations. Lastly, the steady-state acceleration (`−35`) occurs when an acceleration is performed, but at a slower amplitude than for shock or bump. All three acceleration conditions occur in the three axes of the optical system.

Conditions `−32` and `−33` are for a fall of the optical system from some height ranging from 25 mm to 1 m, with the system being inspected for damage afterward. Both environmental conditions are intended to occur in shipping where the optical system will fall on either the corners or sides of a box. The drop and topple condition have less-severe heights for the drop. The

Table 11.7 Conditioning methods for mechanical stresses.

Conditioning Code	Conditioning Method
`−30`	Shock
`−31`	Bump
`−32`	Drop and Topple
`−33`	Free Fall
`−34`	Bounce
`−35`	Steady-State Acceleration
`−36`	Sinusoidal Vibration
`−37`	Random Vibration

severities in the free fall condition are based on their fall height and the mass limits of the packaging.

The bounce condition (−34) is a mix of a vibration of the optical system and a drop. The bounce condition occurs at two discrete frequencies, instead of a range. The optical assembly is not tied down to the vibration table in this case. This type of test is meant to assist with a part moving around during shipping.

The last two conditions (−36 and −37) are for vibrations of the optical system that is mounted to a vibration table, commonly referred to as "shake tests". These conditions vibrate the optical system in either a sinusoidal or random pattern for a time duration rather than the number of cycles. In the case of sinusoidal vibration, the vibration will occur to create vibration fatigue based on a time period and frequency band being used for the severity level. Both conditions must be tested in all three axes of the optical system.

Note that in all these conditions the standard draws heavily upon IEC 60068-2, the IEC standard for environmental testing of electronic equipment, which is listed as a normative reference. While this may be appropriate, it may also not be consistent with other shake and shock test methods such as those described in MIL-STD-810H or the ASTM standards; thus, caution is warranted.

11.5.3 Atmospheric pressure and immersion[5]

Three conditioning methods are given in ISO 9022-8. Two atmospheric pressure tests require the optical system to be at a higher or lower pressure relative to the testing chamber. These two conditioning methods are applicable in many environmental conditions. For example, a sealed airborne instrument can be tested against the high internal pressure that might be experienced at takeoff and the low internal pressure experienced on landing after flight. The third conditioning method in ISO 9022-8 tests for the less common condition that an optical system may be completely immersed in water. These three environmental conditions codes are listed in Table 11.8.

For both the high and low internal pressure conditions (−80 and −81), the values for the severity levels are the same; only the direction of pressure differential is reversed. This environmental test will have either dry air or dry nitrogen used in testing the pressure difference. Across all severity levels of

Table 11.8 Conditioning methods for atmospheric pressure.

Conditioning Code	Conditioning Method
−80	High Internal Pressure
−81	Low Internal Pressure
−82	Immersion

internal pressure conditions, the test duration is 10 minutes. The difference in severities is the pressure difference and the pressure drop.

The immersed condition (−82) is when an optical system is immersed in either a water container or in a water pressure chamber. All levels of severity for an immersion have a two hour test duration. The differences in severity are based on the immersion depth to be simulated, ranging from 1 m to 400 m.

11.5.4 Rain[6]

An optical system may experience different levels of rain when operating outside. The amount of rain can vary from a drip to flooding, depending on the global location. To account for the different amounts of rain, ISO 9022-7 has multiple test conditions. These different test condition codes are listed in Table 11.9.

The optical system is mounted to a rotary table that spins during the test during all rain environmental test conditions. Along with the rotary mounting, the temperature within the chamber must be maintained between 15°C and 35°C. The temperature of the water and the temperature of the optical system will not be the same. To account for this temperature difference, the temperature from difference between the optical system and the water must be between 2K and 20K. The normative description of the test setup requires the use of the German Hellmann rain gauge, rather than a US standard rain gauge or any other commercial products, although the informative Annex A says "*Rain gauge according to Hellmann* is an example of a suitable product available commercially. This information is given for the convenience of users of this international standard and does not constitute an endorsement by ISO of this product."

The most benign rain condition is drip rain (−72). This type of environment has uniform amounts of water dripping over an optical system, covering the top exposed area. Levels of severity for a drip rain vary the rate and exposure time for the optical system. The way the drip rain is produced for this test is through a dispenser plate on the top of the test chamber to evenly distribute the water. Samples requiring a drip plate larger than 1 m^2 can be tested in sections.

In the other two rain conditions (−73 and −74) the optical system will be exposed to a shower head that produces a constant water flow. The severity

Table 11.9 Conditioning methods for rain.

Conditioning Code	Conditioning Method
−72	Drip
−73	Steady Rain
−74	Driving Rain

variation of steady rain is the exposure time and the rainfall rate. For a driving rain, the shower head works in conjunction with a wind generator to increase the wind velocity the water is falling. The severity variation for a driving rain is between the wind velocity, exposure time, and rainfall rate.

11.5.5 Dew, hoarfrost, and ice[7]

When there is moisture present in the atmosphere, this moisture may accumulate on an optical system. There are different descriptions of these conditions as moisture forming in the atmosphere may happen at various outside temperatures. Above freezing, the moisture accumulates as dew. Below freezing, the moisture accumulates as frost or ice. Each of these different types of moisture conditioning codes is listed in Table 11.10.

The difference between each of the environmental conditions in ISO 9022-14 are all based on the temperature at which the optical system sits. Dew (−75) is found when both the optical system and the chamber temperature are above 0°C. In this case, the moisture condenses into a liquid on the optical system. For a dew test condition there is a single level of severity, where an optical system is chilled to 10°C and then warmed to 30°C, creating dew.

The next type of environmental test condition is for hoarfrost. Hoarfrost (−76), similar to dew, is created when water vaper condenses on a surface. In the case of hoarfrost, the water vapor condenses on a surface that is below 0°C and the outside temperature is below 0°C. The water vapor then freezes on the surface in question. The severity variation for hoarfrost is the starting temperature (−10°C or −25°C) and whether an additional hold step between cooling and warming is added, where the chamber is held at −5°C, allowing frost to accumulate.

As a note, hoarfrost is different from frost. When considering frost, not hoarfrost, the water vapor condenses on a surface that is below 0°C, but the outside temperature is above 0°C.

The most severe conditioning method for ISO 9022-14 is ice (−77). Ice is formed when rain, rather than water vapor, freezes onto a surface. There can be two types of ice that is formed on a surface. If rain freezes onto a surface immediately upon impact it is called rime ice. Rime ice will appear white and opaque in color. The other ice that can be formed is glazed ice, created when frozen rain hits a surface and freezes onto the surface. Glazed ice will be clear

Table 11.10 Conditioning methods for dew, hoarfrost, and ice.

Conditioning Code	Conditioning Method
−75	Dew
−76	Hoarfrost and Thawing
−77	Ice covering and Thawing

and can be thicker than rime ice. The severity difference for ice is determined by the starting temperature (−15°C or −25°C), the temperature of the hold step (−5°C, −15°C, or −25°C), and the thickness of the accumulated ice (from 2 mm to 75 mm) during the hold step.

In all conditions of ISO 9022-14 the test will have the lower temperature set in the chamber; have the thickness of dew, hoarfrost, or ice built up; and then the chamber will be warmed. Following the test, the residual water vapor may leave a film on the optical system.

11.5.6 Salt mist[8]

A salt mist test described in ISO 9022-4 is intended to ensure that optical systems do not corrode or have mechanical binding issues in a maritime environment. This environmental condition simulates coastal conditions where salt is in the atmosphere. Degradation might occur on an optical surface or coating, as well as the mechanical components in the optical system. This type of environmental test may occur on an either a sample sheet similar to the optical surface or the entire optical system. If testing an optical coating or surface, a sample must be manufactured to mount in the salt mist chamber. The dimensions of a sample sheet are described in the standard.

There is a single conditioning code for salt mist (`−40`) in ISO 9022-4. The difference in severity levels for a salt mist test is the duration the optical system or sample sheet must endure. The temperature in the salt chamber must be held to a constant temperature over the duration of the testing condition. The environmental test for salt mist will have either the optical system or sample exposed to a constant level of salt mist, as listed in ISO 9022-4.

11.5.7 Dust[9]

Dust tests described in ISO 9022-6 expose the optical system to fine SiO_2 dust particles to test seals for a system expected to be exposed to, e.g., desert environment conditions. This test occurs with a variation of dust particle sizes. Similar to a salt mist test, there is only one dust environmental test condition code (`−52`). There are three levels of severity for the dust test. The first is a single, 6 hour exposure to dust in an 8 m/s to 10 m/s velocity at temperatures from 18°C to 28°C. The second severity level requires two, 6 hour exposures: one at 18°C to 28°C and the second at 35°C to 45°C. The third level of severity requires the same exposure from severity one, followed by a 16 hour exposure to low–velocity dust (1 m/s to 3 m/s) at 55°C to 65°C, followed by another six hour exposure to dust at 8 m/s to 10 m/s and 55°C to 65°C.

11.5.8 Acid atmosphere[10]

An acidic atmosphere is when a humid atmosphere contains either sulfur dioxide (SO_2) or hydrogen sulfide (H_2S). These types of gases in the

Table 11.11 Conditioning methods for an acidic atmosphere.

Conditioning Code	Conditioning Method
−41	Humid Atmosphere - Sulfur Dioxide (SO_2)
−42	Humid Atmosphere - Hydrogen Sulfide (H_2S)

atmosphere are found in industrial environments where fuel burning occurs. An atmosphere containing SO_2 or H_2S may have a corrosive effect on an optical system. The conditioning methods for acid atmosphere testing is given in ISO 9022-20 (Table 11.11). Note that these tests are not appropriate for most optical systems and require a specialized test facility.

For both acidic atmospheres containing SO_2 (−41) or H_2S (−42), the environmental conditioning will occur in a typical outside temperature range, between +25°C and +35°C. The severity levels of an acid atmosphere vary the amount of SO_2 (1 cm^3/m^3 to 15 cm^3/m^3) or H_2S (0.5 cm^3/m^3 to 6 cm^3/m^3), in the chamber and the testing durations.

11.5.9 Solar weathering[11]

Solar weathering conditioning verifies resistance to weathering effects. Unlike most other sections in the ISO 9022 series, the two environmental test conditions in ISO 9022-9 are used to evaluate the effects of weathering using two entirely different methods. When discussing each of the test conditions, they must be discussed independently. The conditioning codes for the solar weathering test conditions are listed in Table 11.12.

The solar radiation environmental test condition (−20) simulates an optical system exposed to direct sunlight. The spectral distribution has varying irradiances from the ultraviolet to infrared when considering this type of environmental condition. Testing for solar radiation occurs in either a cyclical method with lower radiant exposure levels that simulate environmental exposure, or with a single continuous cycle with high radiant exposure levels to test for possible ageing effects. The temperature will vary from room temperature up to at most 55°C in all variations of solar radiation testing. This type of environmental test is intended for an optical system ready for operation. Lower levels of severity have the option of testing the optical

Table 11.12 Conditioning methods for solar radiation.

Conditioning Code	Conditioning Method
−20	Solar Radiation
−21	Laboratory Weathering

system in operation, but higher levels of severity are in place solely for ready-for-operation conditions.

Evaluating weathering in a lab environment (−21) utilizes a xenon-arc lamp to speed up lifetime testing. This condition is measured and controlled with either a black-panel or black-standard thermometer. The two thermometers require slightly different measured temperatures throughout the test. Differences in the levels of severity are partially determined by the type of cycle, either behind a window or not behind a window. The exposure parameters for this type of environmental test condition are described more in depth in ISO 4892, *Plastics - Methods of exposure to laboratory light sources*.[12,13]

11.5.10 Mold[14]

Another uncommon environmental test is resistance to mold growth, described in ISO 9022-11. Unlike most parts of 9022, this part is usually intended to be performed on a test piece, mounted optic, or surface treatment; to aid in the selection of materials and not as a production control evaluation. Mold may develop on an optical surface and potentially corrode an optical coating or surface. There is only one environmental conditioning code for mold growth (−85). There are 10 listed species that are used to test for mold growth. When the test is performed, there is a process where the fungi spores are mixed with water and subsequently a mineral salt solution. The optical system is then subjected to the fungus solution. For fungus exposure of complete systems, the test is always performed with the optical system in a state ready for operating.

There are two severity levels that may be assigned in the mold environmental condition. The difference between the severity levels is the amount of exposure time.

11.5.11 Contamination[15]

Contamination occurs when potential chemicals in the environment have an adverse effect, often cosmetic, on an optical system. As with mold growth, contamination susceptibility is intended to be performed on representative samples to aid in the selection of materials likely to suffer contamination during surface life, rather than for regular production control. The different conditioning methods for contamination given in ISO 9022-12 are shown in Table 11.13.

All contamination condition tests are performed with a felt pad soaked with relevant test agents. As some of the environmental test durations for contamination are over longer periods of time with solvents, the felt pad must be constantly soaked as to not dry out.

The differences in the environmental conditions in Table 11.13 have variations in the test agents used and, generally, the duration of the test. Testing for laboratory agents (−86) is the only environmental condition in

Table 11.13 Conditioning methods for contamination.

Conditioning Code	Conditioning Method
−86	Basic cosmetic substances and artificial hand sweat
−87	Laboratory agents
−88	Production plant resources
−89	Fuels and resources for aircraft, naval vessels, and land vehicles

ISO 9022-12 with two durations. This test condition has the system or witness sample subjected to two different types of test agents over a short duration. The other three environmental conditions have the optical system exposed for various durations in days, depending on the severity level.

As is obvious from their names, the different environmental conditions for contamination have different types of test agents used. The mixture of test agents is what one would expect to find for the corresponding types of locations. An example is for the laboratory agent environmental condition. In this case, the component or surface will be exposed to various acids, ethanol, and acetone.

Following the exposure to the test agents an evaluation is performed. The evaluation is performed by visually examining the optical system surfaces for degradation and corrosion.

11.5.12 Combined environments

It may be beneficial to expose an optical system to multiple conditions concurrently for certain environmental tests. This is an extension of realistic global environments. These combined environmental conditions are based on a mixture of the individual environmental test conditions listed above. The severity levels for these combined environmental conditions are a mixture of the individual conditioning method severities to best simulate global environments.

Temperature, Humidity, and Pressure[16]

The simplest type of combined environmental condition is a mixture of temperature and pressure described in ISO 9022-23. These specific cases are a mixture from thermal conditions (−10, −11, and −12); a low ambient pressure; and in −50 and −51, the presence of hoarfrost. They are intended specifically for optical systems designed for operation and/or transport in high mountainous areas or onboard aircraft or missiles. All combined environmental condition codes for temperature and pressure are listed in Table 11.14.

The various conditioning methods in ISO 9022-23 can be grouped into three categories. The first category is for conditioning methods −45 and −46. These two conditioning methods have the test chamber set to a temperature,

Table 11.14 Conditioning methods for temperature and pressure combined.

Conditioning Code	Conditioning Method
−45	Low ambient pressure combined with ambient temperature
−46	Low ambient pressure combined with dry heat
−47	Low ambient pressure combined with damp heat; pressure difference low
−48	Low ambient pressure combined with damp heat; pressure difference medium
−49	Low ambient pressure combined with damp heat; pressure difference high
−50	Low ambient pressure combined with cold, including hoarfrost and dew
−51	Low ambient pressure combined with cold, without hoarfrost and dew

as listed in the severity level, and then pump the pressure down. Both test conditions are then set for a steady-state duration and pumped back to ambient conditions.

The second category of conditioning methods (−47, −48, and −49) are for low ambient pressure with a damp heat at various pressure differences. For the first two of these three test conditions a cycling method is described in the severity level. This cycling method creates the required pressure differences. Because of this cycling, determining the numerical pressure difference from ISO 9022-8, conditioning method −81 is not necessary. In the third condition of a high pressure differential, a single cycle is performed at a more extreme low pressure.

The last two categories of conditioning methods are for low ambient pressures with cold (−50 and −51). These two conditioning methods have a chamber pumped down with a starting pressure difference and then the temperature is reduced. In conditioning method −50, moisture is added to generate hoarfrost or dew; while in conditioning method −51, formation of condensation is prevented. The difference in severity levels is discussed for each section of the conditioning methods.

Temperature, Humidity, and Vibe[17]

Another useful combined environmental condition is a combination of thermal conditions and mechanical bumps or vibrations. For this combined environmental condition, there is one environmental condition code (−22). This is different from other environmental condition codes. With the mix-and-match nature of ISO 9022-22, the notation used to describe this type of environmental condition contains a slash between the two individual aspects: temperature and mechanical stress.

When evaluating the thermal and humidity aspect of this standard, the conditioning codes from ISO 9022-2 that are applicable are cold (−10), dry heat (−11), and slow temperature change (−14). These three types of environmental conditions are good indicators of global environments.

Instead of all severity levels for each environmental condition ISO 9022-22 only uses three severity levels from each type of thermal conditioning code.

All three severity levels provide a broad range or thermal conditions that an optical system may experience outdoors.

The mechanical environmental conditions that may be used for ISO 9022-22 are bumps or random vibrations. The there is only one bump (−31) severity level that is considered in ISO 9022-22. In this case, the most benign condition is used for testing. There are three severity levels that could be tested for in random vibration (−37). The difference in the severity levels for random vibration are the maximum frequency on the test profile. The three severity levels for random vibration are across the entire severity spectrum in this environmental condition. These are meant to give a good indication for the possibilities found in the field.

As stated above, the notation for ISO 9022-22 is different from all other environmental conditions, which are described in detail in Sec. 11.3. For this case, the notation used is similar to each individual conditioning code and severity level, but with a slash between them. An example of this is

ISO 9022-22 −10 −05/31 −01 −1.

This conditioning code states that the thermal conditions are for a cold environment with a severity level of 05, the temperature set to −25°C, and the mechanical condition for a bump. As used for all environmental notation, the final code for this specification is the operation code. In this case, the operation is for a −1 state, where the optical system is ready for operation but not functioning through the test.

Contamination and Solar Radiation[18]

The final combined environmental condition is mixing contamination and solar radiation. As with mold growth and contamination, these conditions are meant for selection of materials and components—for instruments likely to be subjected to combined contamination and solar radiation during service life, rather than for regular production control—and complete instruments are not to be tested under these conditions except for special circumstances.

These combined environments simulate the environment when an optical system is outside in the sun and is encountering either hand sweat or vehicle fuel. There are two conditioning methods available; they are listed in Table 11.15.

Table 11.15 Conditioning method for contamination and solar radiation combined.

Conditioning Code	Conditioning Method
−90	Basic cosmetic substances and artificial hand sweat, combined with solar radiation
−91	Fuels and other resources for aircraft, naval vessels, and land vehicles, combined with solar radiation

Both combined environments in ISO 9022-17 are extremes for the contamination requirement. The hand sweat and solar radiation conditioning method (−90) simulates a system operating outdoors. The fuel and solar radiation conditioning method (−91) simulates an optical system operating in a rough environment outdoors with vehicles present. Both severity levels for these combined environments are very similar to their individual conditioning methods. There are a few differences in the severity levels, associated with the humidity requirement, testing duration, and type of operation.

11.6 Testing Equipment

Performing each of the environmental tests listed above requires different types of equipment or chambers to create the necessary conditions. Some of these tests can be conducted using a standard commercially available environmental test chamber. However, because of the diversity in types of environmental tests there is no single piece of equipment that works for all types of environmental test conditions. For most of the environmental tests a type of test chamber is needed to create a sealed off environment. Within the ISO 9022 series the chamber or test equipment necessary may not be specified. There are aspects in the ISO 9022 series that provide information to how the test chamber should be created and operate. Table 11.16 lists the parts of ISO 9022 that discuss the test chamber.

Along with the test chambers, some of the sections in the ISO 9022 series describe the type of witness or sample sheet that should be used; the sample sheet is described in the salt mist, mold, and contamination subsections of ISO 9022.

Table 11.16 ISO 9022 standards that provide a description of the test chamber to be used.

ISO Standard	Type of Test	Chamber Description
ISO 9022-4	Salt Mist	A schematic of the test chamber is provided.
ISO 9022-6	Dust	A schematic of the test chamber is provided.
ISO 9022-7	Rain	A schematic of the test chamber is provided as an overview.The additional attachments for the different environmental test conditions in ISO 9022-7 are shown.
ISO 9022-8	Atmospheric Pressure	A diagram of the test chamber is provided.A test connector drawing is shown along with thread information for connecting the valve.
ISO 9022-11	Mold	A description of the incubation chamber for the mold is provided
ISO 9022-12	Contamination	The test chamber is described in cursory detail; but the method of producing the contamination, through a test pad, is shown with a diagram.
ISO 9022-20	Acid Atmosphere	A schematic of the test chamber is provided.

Table 11.17 Example of a section of environmental tests based on ISO 10109 documentation.[1]

				Area of application: Indoors			
				Technical climate in weather-protected locations			
Test number	**Test description**	**In accordance with standard**	**Conditioning method**	**Degree of severity**	**State of operation**	**Technical requirement**	**Status after test**
1	Climate						
1.1	Dry heat	ISO 9022-2	11	01	2	+10°C, Relative humidity < 40%, 16 hours	A
1.2	Damp heat	ISO 9022-2	12	01	2	+40°C, Relative humidity 90% to 95%, 24 hours	A
2	Atmosphere						
2.1	High internal pressure	ISO 9022-8	80	09	2	+40°C, Pressure difference: 40 kPa, Pressure drop: 10 kPa, 10 min	A
2.2	Low internal pressure	ISO 9022-8	81	09	2	+40°C, Pressure difference: 40 kPa, Pressure rise: 10 kPa, 10 min	A
2.3	Solar radiation	ISO 9022-9	20	03	1	+40°C, 1 kW/m^2	A
3	Mechanical						
3.1	Drop and topple	ISO 9022-3	32	01	0	Height of drop: 25 mm	A

11.7 Notation of Environmental Test Parameters

When actually specifying the environmental test conditions for an optical system, the indication described in Sec. 11.3 can be used individually. When using conditions from a standard environment, as per ISO 10109, the notation should be described in a table, as shown in Table 11.17. The indications in Table 11.17 are the same as would be found independently. In some cases, the severity level may not match up exactly with the standard environment. In these cases, the user must decide which severity level best suits their application.

For the example shown in Table 11.17, the environmental conditions are based on an indoors environment. The example optical system would be tested for operation in the specified environment.

References

1. International Organization for Standardization (ISO), *Guidance for the Selection of Environmental Tests*, ISO 10109:2015 (Geneva: 2015).
2. ISO, *Environmental test methods—Part 2: Cold, Heat and Humidity*, ISO 9022-2:2015 (Geneva: 2015).
3. ISO, *Environmental Test Methods—Part 3: Mechanical Stress*, ISO 9022-3:2015 ISO 9022-3:2015 (Geneva: 2015).
4. ISO, *Environmental Test Methods—Part 3: Mechanical Stress—Amendment 1*, ISO 9022-3:2020 (Geneva: 2020).
5. ISO, *Environmental Test Methods—Part 8: High Internal Pressure, Low Internal Pressure, Immersion*, ISO 9022-8:2015 (Geneva: 2015).
6. ISO, *Environmental Test Methods—Part 7: Resistance to Drip or Rain*, ISO 9022-7:2015 (Geneva: 2015).
7. ISO, *Environmental Test Methods—Part 14: Dew, Hoarfrost, Ice*, ISO 9022-14:2015 (Geneva: 2015).
8. ISO, *Optics and Photonics—Environmental Test Methods—Part 4: Salt Mist*, ISO 9022-4:2014 (Geneva: 2014).
9. ISO, *Environmental Test Methods—Part 6: Dust*, ISO 9022-6:2015 (Geneva: 2015).
10. ISO, *Environmental Test Methods—Part 20: Humid Atmosphere Containing Sulfur Dioxide or Hydrogen Sulfide*, ISO 9022-20:2015 (Geneva: 2015).
11. ISO, *Optics and Photonics—Environmental Test Methods—Part 9: Solar Radiation and Weathering*, ISO 9022-9:2016 (Geneva: 2016).
12. ISO, *Methods of Exposure to Laboratory Light Sources—Part 1: General Guidance*, ISO 4892-1:2016 (Geneva: 2016).
13. ISO, *Methods of Exposure to Laboratory Light Sources—Part 2: Xenon-Arc Lamps*, ISO 4892-2:2013 (Geneva: 2013).

14. ISO, *Environmental Test Methods—Part 11: Mould Growth*, ISO 9022-11:2015 (Geneva: 2015).
15. ISO, *Environmental Test Methods—Part 12: Contamination*, ISO 9022-12:2015 (Geneva: 2015).
16. ISO, *Optics and Photonics—Environmental Test Methods—Part 23: Low Pressure Combined with Cold, Ambient Temperature and Dry and Damp Heat*, ISO 9022-23:2016 (Geneva: 2016).
17. ISO, *Environmental Test Methods—Part 22: Combined Cold, Dry Heat or Temperature Change with Bump or Random Vibration*, ISO 9022-22:2012 (Geneva: 2012).
18. ISO, *Environmental Test Methods—Part 17: Combined Contamination, Solar Radiation*, ISO 9022-17:2015 (Geneva: 2015).

Chapter 12
Standards in Practice

We have shown in previous chapters that ISO 10110 can be a very powerful method for conveying information between the optical designer and the fabricator. The coded notation may seem confusing at first, but once understood, it obviates the need for significant additional communication throughout the fabrication process. By instituting the ISO 10110 drawing format, it is also possible to define an entire optical system specification and acceptance test. Starting the entire project with ISO 10110 and the associated ISO performance standards in mind allows the designer to have a concise set of requirements defined that can also be easily documented and communicated.

Users should develop an understanding of how to use standards holistically; in other words, how to specify full optical systems and assemblies. In order to better understand how to use standards successfully for both tolerances and for performance, in this chapter we provide a complete example. The goal is to help users understand how to take advantage of the ISO 10110 drawing notation from the beginning of an optical design, through tolerancing, and result in finished drawings.

Along with component drawings made with the ISO 10110 drawing format, the finished optical system may require overall performance verification. As discussed in chapter 10, there are many different types of performance metrics that can be used. In this chapter, we discuss specifications for an optical system that must be tested for multiple parameters.

12.1 System Parameters

To begin the optical system example, basic parameters must be established for the optical system itself. The requirements below are assumed to be reasonable for the example optical design to avoid needless complexity.

- Wavelength: Visible (400 nm to 700 nm)
- Effective focal length: 100 mm
- F/#: F/5

- Field of view: ±5°
- Object distance: infinity
- Field-averaged RMS wavefront error: ≤0.1 waves polychromatic (*F*, *d*, and *C*; evenly weighted)
- System transmittance: >85% at *d* (no requirement on *F* and *C* wavelengths)
- Environment: must operate in a non-weather protected location with restricted limiting values as per standard environment six in ISO 10109.[1]
- Testing requirement: only axial performance to be tested.
- Other requirements: no requirements on distortion, relative illumination, weight, or size

The example lens we show for these specifications has four elements. This lens allows us to produce ISO 10110 drawings with some variety to better indicate parameters and notation on ISO 10110 drawings.

Additionally, this system must perform outside of a climate-controlled environment. Consequently, we must state requirements for the optical system to work in a relatively benign outdoor environment, such as what might be seen by a commercial digital camera. We take this to be the standard environment six in ISO 10109. That is, the optical system must perform in different temperatures, humidity levels, pressures, and weather that include rain.

12.2 Optical Design and Tolerance Analysis

The optical design form from the design stage is a modified Petzval type, using the Gehrke PN-P-N form.[2,3] This layout is a Petzval type, but the rear element is an air-spaced doublet to flatten the field. Note also that performance of the system has been further enhanced by the inclusion of an aspheric surface on the first surface of the third lens. While an asphere is not necessarily required in practice, we include one here to broaden the variety of ISO 10110 drawings required in the example.

12.2.1 Optical design

The resulting layout and prescription for this optical design is shown in Fig. 12.1 and Table 12.1, respectively. In this example, the system figure of merit is required. We use an RMS wavefront error metric that is specified on the assembly drawing. From the optical design code, we have a nominal RMS wavefront error of 0.061 waves polychromatic (*F*, *d*, and *C*; evenly weighted).

Along with the optical system RMS wavefront error performance required for our final test, there is an additional requirement on the system

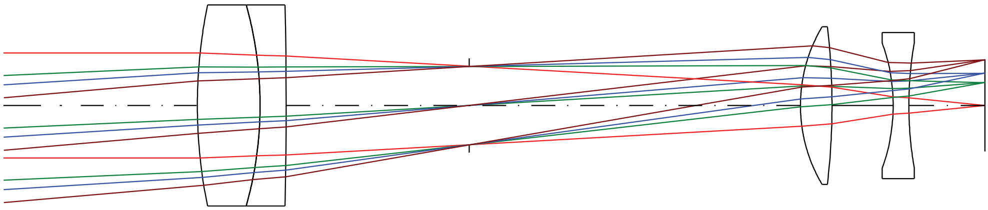


Figure 12.1 Modified Petzval lens design layout.

Table 12.1 Modified Petzval lens design prescription.

Surface	Radius (mm)	Thickness (mm)	Material Code	Semi-aperture (mm)
1	96.00	12.00	618634	18.00
2	−72.00	5.00	750353	18.00
3	−833.00	0.00	AIR	18.00
4	Infinity	35.00	AIR	
STOP	Infinity	61.36	AIR	7.448
5 (ASPH)	30.00	6.00	618634	14.00
		A: 1.5658E-8 B: 4.5519E-9 C: -1.4278E-11		
6	−130.00	0.00	AIR	14.00
7	Infinity	11.71	AIR	
8	−35.00	3.00	789284	12.00
9	70.00	0.00	AIR	12.00
10	Infinity	14.50	AIR	AIR
Image	Infinity	0.00	AIR	

transmittance. While transmittance is not a figure of merit with its own notation in ISO 10110, it is important for some applications; thus having the transmittance on our drawing is beneficial.

12.2.2 Optical tolerances

From past experience and what has been widely published,[4,5] a set of tolerance values are assigned based on computer-aided tolerance analysis. The assigned tolerances are listed Table 12.2.

After tolerance assignment, a Monte Carlo simulation of 1,000 cases shows that we would expect better than 0.099 waves of field-averaged RMS wavefront error in 95% of the as-built simulations.

When going through the tolerance run, three compensators are used. We allow three airspace thicknesses to vary: from the aperture stop to element 3 (aspheric lens), from element 3 to element 4, and the back-focus distance. The limits of these distances are converted into thickness tolerances in the optical system drawing.

Most of these tolerances listed would be considered precision based on published information. If this design were to be produced, we likely would undertake more work to loosen tolerances and tune the design while concurrently working on the mechanical design. Our purpose here, however,

Table 12.2 Optical tolerances assigned for modified Petzval lens.

Tolerance Description	Value
Power	±1 Fringe @633 nm
Irregularity	±0.25 Fringes @633 nm
Refractive Index	±0.0005
Abbe Number	±0.5%
Homogeneity	10×10^{-6} or NH10
Center Thickness	±50 μm
Annulus Sag	±0.025 μm
Diameter	+0.0 mm / −0.10 mm
Wedge	0.010 mm total indicator runout – across entire diameter
Coating	Antireflective across entire waveband; $R \leq 0.5\%$
Element Tilt	±1′
Element Decenter	±12.5 μm

is to show the use of the ISO 10110 standard notations and this tolerance set is sufficient for that purpose.

12.3 Optical Element Drawings

As each of the elements within the optical system are different, we make distinct types of drawings that follow different schema described in ISO 10110 to provide a better example.

Some parameters on each of these drawings are consistent for all elements. All optical elements have the same tolerances on the material properties. One set of these material properties are shown in Fig. 12.2. All three optical elements have a tolerance of `±0.0005` for the index and `±0.5%` for the Abbe

MATERIAL	
COMPANY	GLASS
Schott	N-PSK53A
Ohara	S-PHM52
Hoya	PCD4
n_d 1.61800 ± 0.0005	
V_d 63.33 ± 0.5%	
0/ 10	
1/ 1 x 0.16	
2/ NH10;SW10	

Figure 12.2 Sample material properties table.

(λ) AR R(0.4 μm to 0.7 μm) ≤ 0.5 %

Figure 12.3 Sample notation for optical coating.

number. The material manufacturer and type, index, dispersion, and the index and dispersion tolerance are all shown in accordance with ISO 10110-1.[6] The stress birefringence (0/) is 10 nm/cm. The number of allowable bubbles and inclusions (1/) must be less than a quantity of 1 with a grade of 0.16. The homogeneity and striae (2/) have a specification of NH10 and SW10, respectively. The NH10 requires the inhomogeneity to be less than 10×10^{-6}. The SW10 requires any striae be less than 10 nm per 50 mm glass path length. The 0/, 1/, and 2/ notations are in accordance with ISO 10110-18.[7]

All the optical elements also must have antireflective (AR) coatings. The specification for the coatings operates across the entire visible waveband, 0.4 μm to 0.7 μm, and have reflections less than or equal to 0.5% (Fig. 12.3). These optical coating requirements are given in accordance with ISO 10110-9, but could be shown graphically in accordance with ISO 9211-2 instead.[8,9]

Along with the optical coating specification, environmental durability tests must be performed on the coatings. These tests are shown in Fig. 12.4. The three tests listed are for damp heat, or humidity (step 1); abrasion (step 2); and adhesion (step 3). The notation for each of these tests indicate the required test severity level. For example, the abrasion test uses cheesecloth rubbed against the coated surface 100 times with a force of 5 N. The coating durability test requirements are shown in accordance with ISO 9211-3.[10]

12.3.1 Doublet optical element

For the first drawing in our optical system (Fig. 12.5), we have a doublet lens. We could make individual drawings for each of the optical elements, but we choose instead to make a single print with ISO 10110 for the entire doublet lens, which is allowed according to the latest revision of ISO 10110-1. By employing the doublet print method, we can convey our centering tolerances without ambiguity on each individual print.

The dimensional tolerances that are listed with direct tolerances are solely for the center thicknesses of the two optical elements and the outer diameter of

Coating Environmental Durability Test	
Test Step	Test
1	ISO 9022-2-03-07
2	ISO 9211-4-01-02
3	ISO 9211-4-02-02

Figure 12.4 Sample notation for optical coating environmental durability test specification.

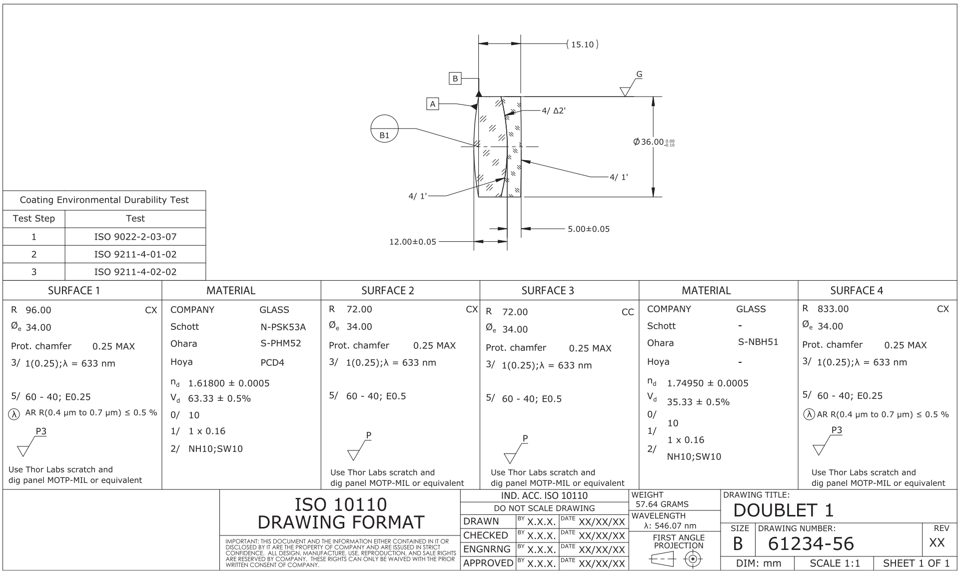


Figure 12.5 Optical system sample drawing for elements 1 and 2 using the ISO 10110 doublet drawing format.

the entire doublet. The radii of all surfaces are listed in the tabular field. These are all given in accordance with ISO 10110-1. The power and irregularity on all four optical surfaces are `1` fringe of power and `0.25` fringes of irregularity at a wavelength of `633 nm`, and are shown in accordance with ISO 10110-5.[11] The surface imperfections are shown with the visibility method, a scratch and dig requirement of `60-40` in accordance with ISO 10110-7.[12] Along with the specification for the scratch and dig, we listed the comparison standard that would be expected to verify this specification. In this instance, we expect listed the comparison standard as the `Thor Labs MOTP-MIL scratch and dig panel`. We also allow other comparison standards to be used, as long as they are equivalent to the Thor Labs panel. Additionally, edge chips that extend less than `0.25 mm` in from the outer diameter are allowed.

The surface roughness for the exterior surface for this optical element has a polishing grade of `P3`, `Rq <2 nm` over the default spatial band of `0.002 mm` to `1.00 mm`. The buried surface that is contacted between the two optical elements does not have a specified roughness or waviness requirement. The specification for the buried surfaces is solely a polish. The outer diameter or cylindrical surface is specified to be a ground surface without a specific requirement on the roughness over a spatial band. All these tolerances are given in accordance with ISO 10110-8.[13]

The centering tolerances for this doublet are shown in accordance with ISO 10110-6.[14] The left optical element (element 1) uses the center of curvature of the left optical surface (surface 1) and the cylindrical datum target B1 to create a datum axis for all the centering tolerances. The tilt error for the right optical surface (surface 2) must be within 1′ of this defined axis. The contact surface for the right optical element (element 2) has a 2′ tilt tolerance relative to the same datum axis. The right surface (surface 4) for the right optical element has a tilt tolerance of 1′ relative to the same datum axis. Notice that these centering tolerances are shown in the drawing field for clarity rather than the table field.

12.3.2 Aspheric optical element

The second drawing in the optical system (Fig. 12.6) shows the next element, which makes use of a different format since it has an aspheric surface. Using the drawing format for an asphere described in ISO 10110-12,[15] we list the necessary aspheric terms in a separate table from the standard table field as well as specifying that the left surface is an aspheric surface with ASPH in the tabular field. Note that the vertex radius for this surface is not shown in the left surface table field; the vertex radius R is now a coefficient (analogous to k or A_4) that describes the aspheric surface and is tabulated accordingly. A sag table is also included, as required. The equation is listed on the drawing,

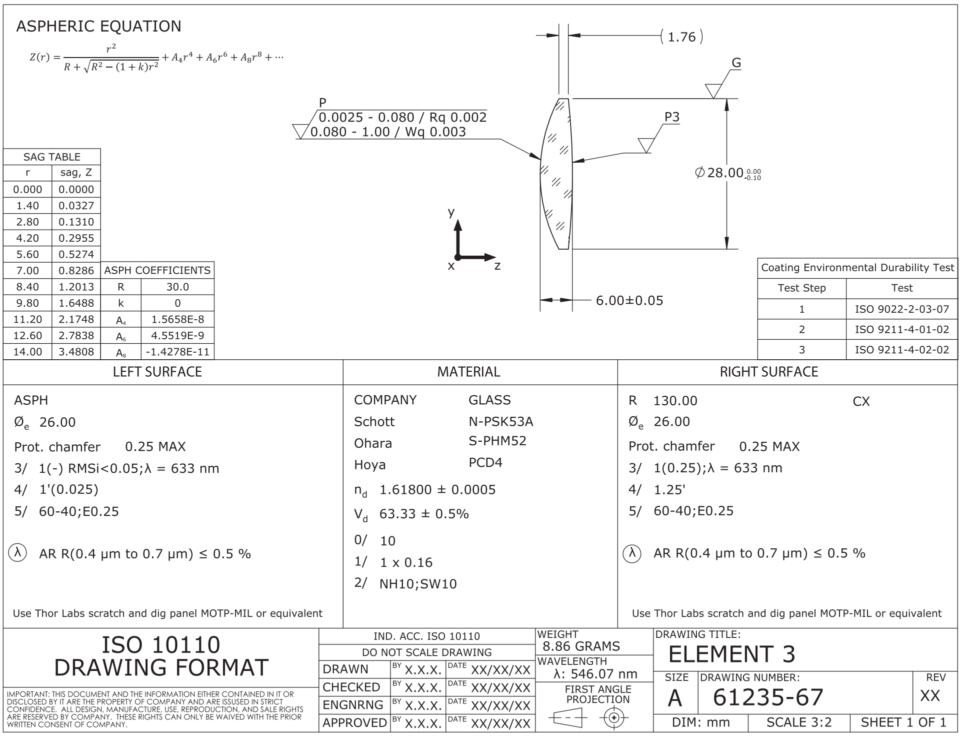


Figure 12.6 Optical system sample drawing for element 3, an aspheric element, using the ISO 10110 drawing format.

normatively required by ISO 10110-12 to avoid confusion on what definition is used. Additionally, a coordinate system is explicitly defined on the drawing to show how the aspheric surface should be interpreted unambiguously. In this case, we have chosen to use the default coordinate system, a right-handed coordinate system with the *z*-axis pointed to the right, in which the origin of the coordinate system is the vertex of the first surface. The spherical surface has the radius listed in the tabular field, as would be expected. The other dimensional parameters, such as center thickness and diameter, are listed with their respective tolerances.

The power tolerance for both surfaces is defined as `1` fringe with a wavelength of `633 nm`. The irregularity tolerance for the spherical surface is `0.25` fringes at a wavelength of `633 nm`. For the aspheric surface, the irregularity tolerance is specified as an RMS irregularity less than `0.05` fringes at `633 nm`. The surface imperfection tolerances are the same as for the doublet in Fig. 12.5.

The method for defining the wedge tolerance for this optical element is different from the previous method as the outer diameter or cylindrical surface is used as the datum to define a mechanical axis. Both optical surfaces have a tilt relative to that datum axis. As this is an aspheric surface, there is a tolerance on the decenter of the best fit asphere vertex from the datum as well as the surface tilt. Because the datums for this case are defined in ISO 10110-6 they are implicit, and need not be shown on the drawing.

The surface roughness for the spherical surface has a polishing grade of `P3`; that is, `Rq <2 nm` over the default spatial band of `0.002 mm` to `1.00 mm`, as before. The aspheric surface has a more well-defined tolerance for the surface texture. Evaluated over the spatial band `0.0025 mm` to `0.080 mm` the RMS roughness must be less than `0.002 µm`. The rms waviness tolerance of the surface is `0.003 µm`, evaluated over the second spatial band from `0.080 mm` to `1.00 mm`. The outer diameter or cylindrical surface must be a ground surface without a specific roughness requirement. These notations are all in accordance with ISO 10110-8.

12.3.3 Spherical optical element

The drawing of the last optical element in the optical system (Fig. 12.7) is the bi-concave singlet. This optical element drawing is simpler than the previous two drawings. Most of the dimensions and tolerances are shown as in the previous examples. The only mechanical feature in this drawing not present in the others is the inclusion of face-flat annuli surfaces on both sides of the lens. These two features are defined and toleranced based on the sag from the radii of each side of the optical element, and shown in the drawing field.

Unlike the other two optical elements, the surface imperfections on this lens are defined with the dimensional method instead of the visibility method.

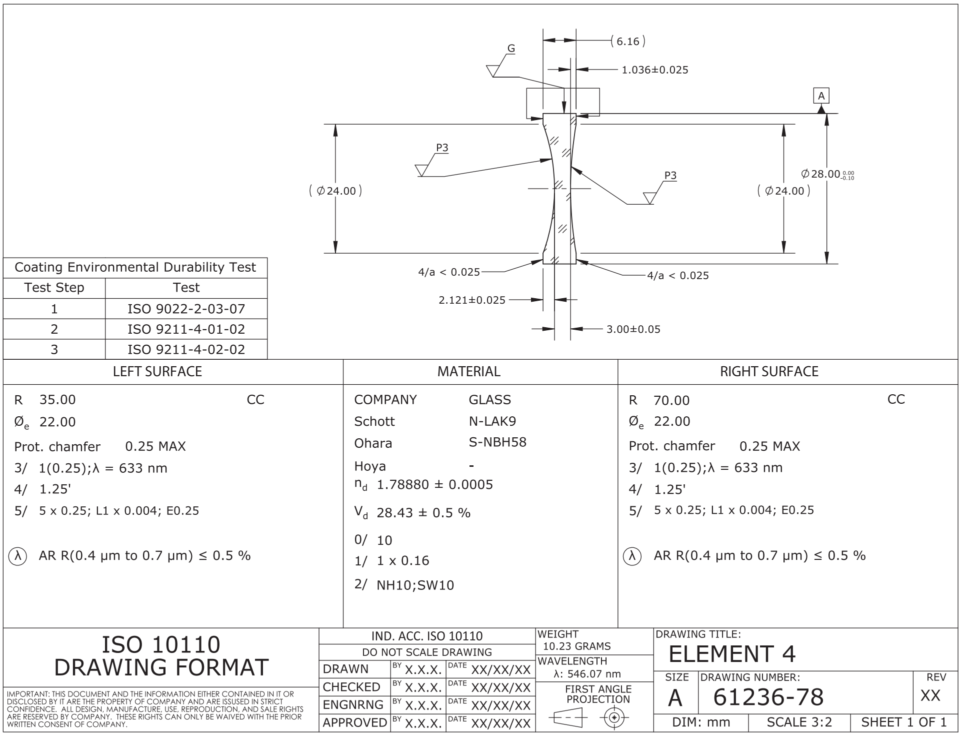


Figure 12.7 Optical system sample drawing for element 4 using the ISO 10110 drawing format.

It is common for field lenses or elements close to a focal plane to be specified dimensionally rather than by visibility to provide a more-objective evaluation method, because of the relative size of the light bundles and the imperfections. However, in this case we have merely chosen this specification method to provide a different surface imperfection format example. For this lens element, both surfaces are allowed a maximum of `5` surface imperfections of grade `0.25`. As discussed in chapter 5, there is also the possibility of accumulations for these types of surface imperfections, provided the sum of the area of all the imperfections is less than 5×0.25^2, or 0.3125 mm^2. One long surface scratch with a width of `4 µm` is allowed, regardless of length. Thinner scratches may be allowed provided the sum of all scratch widths is less than `4 µm`. Additionally, edge chips that extend less than `0.25 mm` in from the outer diameter are allowed. The wedge for this optical element has a tolerance of each optical surface tilt of `1.25`′ from the datum axis created by the outer diameter or cylindrical surface. Additionally, the two face-flat annuli are toleranced with a maximum axial runout relative to the cylindrical datum axis in accordance with ISO 10110-6. Note that axial runout is specified in mm and not degrees.

12.4 Optical Assembly Drawing

The last drawing in this example is the assembly drawing (Fig. 12.8). As this is an optical assembly without component manufacturing tolerances that must be defined, the tabular field is removed.

The full optical layout is shown in the drawing field with the axial spacings defined and toleranced accordingly. Each of the three optical elements are identified by an element number. A table in the top right of the drawing associates these element numbers with their part number and description, as well as the required tilt and decenter tolerances for each. The spacing from the aperture stop to element 3, and between elements 3 and 4, is identified as variable axial spacings. These spacings are used in the tolerance analysis to maintain a one-time, factory-set internal focus compensation adjustment. This choice is noted with the reason for this variation in the bottom right corner. Additionally, the back-focus distance is specified as an adjustable axial separation. This spacing is used in the tolerance analysis to achieve best focus, as stated in the bottom right corner of the print. The placement and diameter of the aperture stop between doublet 1 and element 3 is shown with tolerances. There is also a physical aperture defined at the front of the optical system for testing purposes. This aperture is not critical to the functionality of the optical system but is necessary for the required testing of the transmitted wavefront error and transmission.

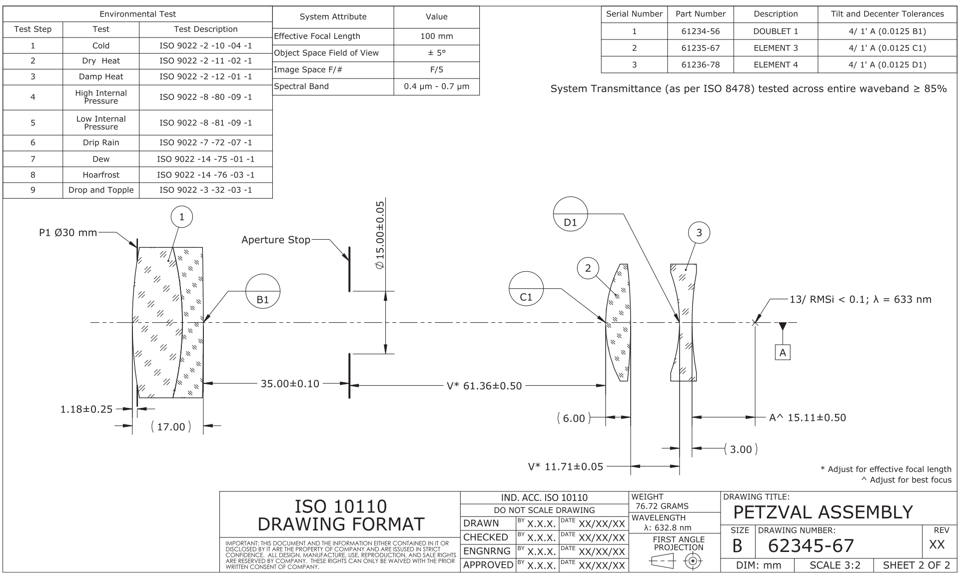


Figure 12.8 Optical system sample drawing for the entire optical assembly using the ISO 10110 drawing format.

The transmitted wavefront error, specified with the `13/` notation in accordance with ISO 10110-14, is a functional requirement for this optical system.[16] The specification requires the measured RMS wavefront with power removed to be less than `0.1` waves single-pass with a test wavelength of `633 nm`. According to our Monte Carlo analysis that is discussed in Sec. 12.2.2, the first-pass yield for this requirement should be 95% or greater. In addition to the transmitted wavefront error requirement the measured system transmittance must be greater than or equal to `85%`. As shown with a note in the upper right corner a transmittance test is necessary for acceptance of the optical system. Analysis shows that this requirement will be met if the AR coatings on the components meet their requirements. The transmittance test is described in detail in Sec. 10.3.5.

As the optical system is intended to be used in an outdoor environment, environmental tests are defined in the top left corner of the drawing. There is a sequence of eight tests to be performed. Between each of these environmental test steps, the performance for both the transmitted wavefront and system transmittance should be evaluated. The final acceptance of the optical system is completed after these environmental tests are performed and the optical system is verified to meet requirements. Note that often in production, these environmental tests are conducted on the first article only, perhaps on a witness sample periodically, and not on each delivered telescope.

12.5 Expected Deliverable

Unlike tolerances for individual optical elements, system-level tolerances and tests can be harder to communicate. These types of results must be reported, but their presentation can vary significantly. As shown on Fig. 12.8, there can be multiple tests required for acceptance of an optical system. The required information and basic presentation of each system level result is typically shown in their respective ISO standards. All standards require a different amount of information to be conveyed. Since the presentation of system-level test results can vary, it is up to the vendor to provide their own method of reporting results of acceptance tests of an optical system. Nevertheless, there are some good practices that should be observed that we discuss in the context of our example.

The following pieces of information are requested in nearly all system evaluation standards and should always be included:

- Title of test performed,
- Name of vendor/operator who performed the test,
- Identification number for the test report,
- Name of the client,

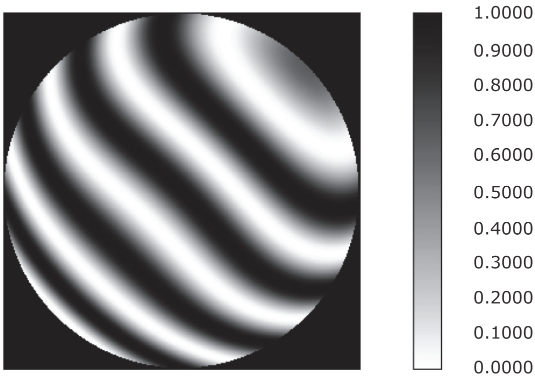


Figure 12.9 Sample interferogram for a transmitted wavefront of the optical system.

- Value or plot of optical system test results, and
- Sign-off from the operator that the optical system met the requirement.

12.5.1 Transmitted wavefront test report

The result from the axial transmitted wavefront test required for the sample optical system (13/ RMSi < 0.1; λ = 633 nm) consists of a single output value: the RMS wavefront error. As the transmitted wavefront is more complex than a single value, an interferogram should be a deliverable to the customer, presenting the transmitted wavefront and verifying the required transmitted wavefront value (Fig. 12.9). Using the interferogram, a customer can analyze the results themselves and verify the acceptance of the optical system.

12.5.2 Optical system transmittance test report

The system transmittance requirement in our example applies to all wavelengths in the spectrum; thus, a single value is not sufficient. A curve showing the system transmittance is required to provide acceptance. This situation is common. As shown in chapter 10, the type of setup required for a transmittance test provides results for the conditions needed. An example of a transmittance curve is shown in Fig. 12.10. This transmittance curve is an example of what may be found from testing the optical system from Fig. 12.8.

Along with the deliverable of the system transmittance curve, the conditions in which the transmittance was performed should be shown. This includes the way in which the optical system was mounted and the test and measurement parameters. This report can include the placement of the optical system relative to the integrating sphere, the focal length of the collimator, the light source used, and the aperture size of the optical system under test.

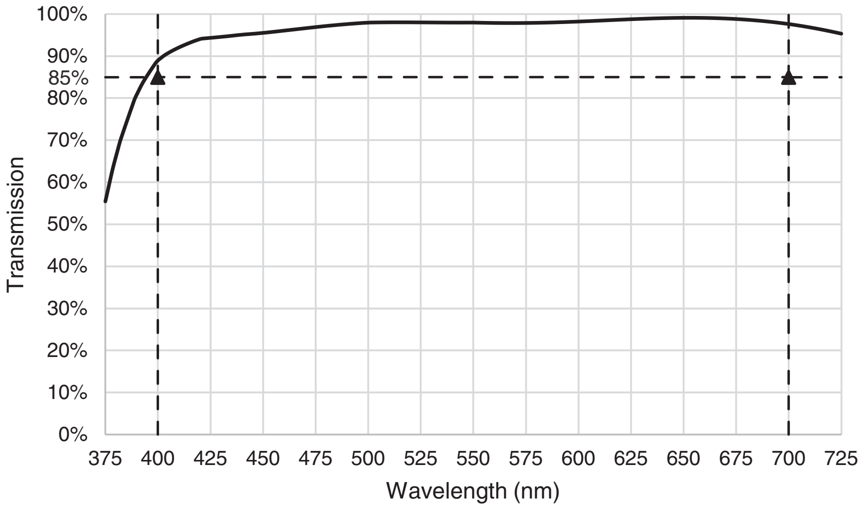


Figure 12.10 Sample transmittance results for an optical system (coating data for figure courtesy of Edmund Optics, Inc.).

12.5.3 Environmental test report

Each individual environmental test condition lists their own required amount of information for acceptance. In general, data sets include the as-performed test conditions, testing devices used, calibration dates, and resulting test cycle profiles. Since the environmental test is not the functional test, a simple method of acceptance for environmental tests is shown in Table 12.3.

Since environmental tests on their own do not include a functional optical test it may be acceptable to separately present the independent results from an environmental test and the optical system evaluation. A full example of this is shown in ISO 10109. Table 12.4 shows a possible example of these test results for temperature requirements.

Ultimately, the details of the system performance tests, environmental tests, and other documentation of the compliance (or noncompliance) of every aspect of an optical system is the subject of extensive discussion and agreement between the manufacturer and the customer. However, having

Table 12.3 Test notation used in ISO 10109 for status after an environmental test.

Test Notation	Description
A	Performance is satisfied
B	Performance is satisfied but there is damage to the optical system. The damage is not needed for functionality but may affect lifetime
C	Not all performance is satisfied and there is damage to the optical system. The damage is not needed for functionality but may affect lifetime
D	The optical system may not function anymore

Table 12.4 Example table to present environmental test results when following ISO 10109.

				Area of application: Outdoor			
				Non-weather-protected locations in temperate climate			
Test number	Test description	In accordance with standard	Conditioning method	Degree of severity	State of operation	Technical requirement	Status after test
1	Climate						
1.1	Cold	ISO 9022-2	10	04	1	−20°C, 16 h	A
1.2	Dry heat	ISO 9022-2	11	02	1	+40°C, r.h. < 40% and 16 h	A
1.3	Damp heat	ISO 9022-2	12	01	1	+40°C, r.h. 90% to 95% and 24 h	A

those discussions within the context and framework of the ISO standards on the subject greatly facilitates the path to a successful outcome.

References

1. International Organization for Standardization (ISO), *Guidance for the Selection of Environmental Tests*, ISO 10109:2015 (Geneva: 2015).
2. F. W. Gehrke and R. Tiedeken, "Lens system," US Patent 2187780 (Washington, DC: 1940).
3. A. Cox, *A System of Optical Design*, Focal Press (London: 1964).
4. J. DeGroote Nelson, R. N. Youngworth, and D. M. Aikens, "The cost of tolerancing," *Proc. SPIE* **7433**, 74330E (2009) [doi: 10.1117/12.828433].
5. Optimax Systems, "Manufacturing tolerance chart," accessed Spetember 21, 2021, http://www.optimaxsi.com/innovation/optical-manufacturing-tolerance-chart/.
6. ISO, *Preparation of Drawings for Optical Elements and Systems—Part 1: General*, ISO 10110-1:2019 (Geneva: 2019).
7. ISO, *Preparation of Drawings for Optical Elements and Systems—Part 18: Stress Birefringence, Bubbles and Inclusions, Homogeneity, and Striae*, ISO 10110-18:2018 (Geneva: 2018).
8. ISO, *Preparation of Drawings for Optical Elements and Systems—Part 9: Surface Treatment and Coating*, ISO 10110-9:2016 (Geneva: 2016).
9. ISO, *Optical Coatings—Part 2: Optical Properties*, ISO 9211-2:2010 (Geneva: 2010).
10. ISO, *Optical Coatings—Part 3: Environmental Durability*, ISO 9211-3:2008 (Geneva: 2008).
11. ISO, *Preparation of Drawings for Optical Elements and Systems—Part 5: Surface Form Tolerances*, ISO 10110-5:2015 (Geneva: 2015).

12. ISO, *Preparation of Drawings for Optical Elements and Systems—Part 7: Surface Imperfections*, 03/03/2017, ISO 10110-7:2017 (Geneva: 2017).
13. ISO, *Preparation of Drawings for Optical Elements and Systems—Part 8: Surface Texture*, ISO 10110-8:2019 (Geneva: 2019).
14. ISO, *Preparation of Drawings for Optical Elements and Systems—Part 6: Centring Tolerances*, ISO 10110-6:2015 (Geneva: 2015).
15. ISO, *Preparation of Drawings for Optical Elements and Systems—Part 12: Aspheric Surfaces*, ISO 10110-12:2019 (Geneva: 2019).
16. ISO, *Preparation of Drawings for Optical Elements and Systems—Part 14: Wavefront Deformation Tolerance*, 2018 (Geneva: 2018).

Epilogue
The Path Forward

Hopefully, we have made a compelling case for the use of international standards in general and of the ISO 10110 standard in particular. We are providing this book as a guide for those who have elected to embrace the future and begin using this drawing notation system. However, standards are living documents, and by the time you read this book, some of the standards referenced in this book may have changed.

It is always best to use the latest revisions of standards when possible. It is appropriate, therefore, for us to provide one last bit of guidance to the reader, where the authors have some insight into what may be happening to the standards in the years to come. And while such speculative sections rarely age well, we will take out our crystal ball and do some gazing, even if it is at our own peril.

The safest thing to postulate is backward compatibility. ISO/TC 172/SC 1 has worked very hard in every revision to make the drawing notations backward compatible with each new part or revision to the standard. There are a few examples where this has not been possible, but by-and-large we should expect that drawings made today using the current revisions of ISO 10110 will still be completely valid in the years to come.

As of this writing, there are several projects in progress that will result in changes to the standards, and we can be fairly certain that they will be completed in the years ahead. First, an entirely new part of ISO 10110 is expected in 2022 or 2023: ISO 10110 Part 16, Diffractive surfaces. The subject of this document includes the presentation, description, and dimensioning of diffractive surfaces such as linear or circular gratings or computer-generated holograms.

Two other projects will revise the centering standard ISO 10110-6 and the surface form standard ISO 10110-5. While at this time the scope of these projects is modest—correcting errors and making some of the notations more clear—it is possible that more form or centering notations might be added or some of the existing notations might be augmented in some way. There has also been discussion on harmonizing different standards that should complement each other, such as ISO 10110-5, ISO 10110-14, and ISO 14999-4. Due to different cycles in revision, sometimes such synchronization is challenging without a specific focus to achieve that goal.

One other project is underway that will impact the environmental standards: revision of ISO 10109, guidance for the selection of environmental tests. This important standard currently provides six “standard environments” that are loosely based on IEC environmental standards. The revision will almost double this number of standard environments and provide much better guidance on how to specify environmental testing. This revision and those of Parts 5 and 6 are expected in 2022 or perhaps 2023.

That is as far into the future as our current information on the work program of TC 172/SC 1 allows us to see. But there are other things expected to change in the next five years that are worth including. Two other parts of ISO 10110 are in need of re-work and will likely get attention soon. First is the laser damage notation in Part 17 and the associated metrology standards ISO 21254, Parts 1 through 4. It is difficult to say today what form those revisions will take, but as likely as not, the entire method of specifying and conducting laser damage testing will change to reflect current practice. The other area under discussion is the standards supporting new manufacturing methods for aspheric and freeform optical surfaces. Molded plastic and glass elements have additional requirements for notations unique to the molding process. Diamond turned optics and deterministically finished components require a more elaborate method for specifying surface texture than is currently provided in Part 8, and the surface notation given in Part 19 was really just a starting place for CAD/CAM data transfer for optics.

But the greatest change that we expect to see is that of adoption. The revisions to the surface imperfection and materials imperfections standards removes the last barriers to broad adoption of ISO 10110 in the US and other countries outside of Europe. Now that we can specify surface imperfections using the classic MIL scratch and dig visibility method and the blank material rather than the finished part for birefringence, bubbles, homogeneity, and striae, it will be an easy transition for most manufacturers. From the perspective of 2021, we see a tipping point, where the rate of adoption will swing from 15−20% now to 70−80% by the end of the decade. At long last, the international optics drawings standard is truly international, and we can all work from a global system of notations in the 21st century.

Index

Eric Herman is an optical design engineer with Zygo Corporation. He has over a decade of optical design experience. He received his M.S. in Optical Sciences from the University of Arizona's James C. Wyant College of Optical Sciences. He is a Senior Member of SPIE and has served in a variety of volunteer experiences for both SPIE and the Optics and Electro-Optics Standards Council (OEOSC). He is currently the Chair-Elect on the Board of Directors for OEOSC.

Dave Aikens is an optical design engineer at Savvy Optics Corp, a company he founded in 2007. He has been working on the standardization of optics drawings since his first job working with Ron Kimmel at Hughes Aircraft Company in 1983. He is a graduate of the University of Rochester (B.S. 1983, M.S. 1984), a Fellow of SPIE, and the 2021 recipient of the A. E. Conrady award. He currently serves as the Chair of the Technical Advisory Group to ISO Technical Committee TC172, responsible for standards in optics and photonics.

Richard N. Youngworth is a Fellow of SPIE and Optica. He is the founder and Chief Engineer of Riyo LLC, an optical design and engineering firm. He obtained his Ph.D. in optics at the University of Rochester. He has authored and delivered numerous short courses, papers, talks, and lectures on optical design and engineering. He is active in optical standards development and in professional society work.