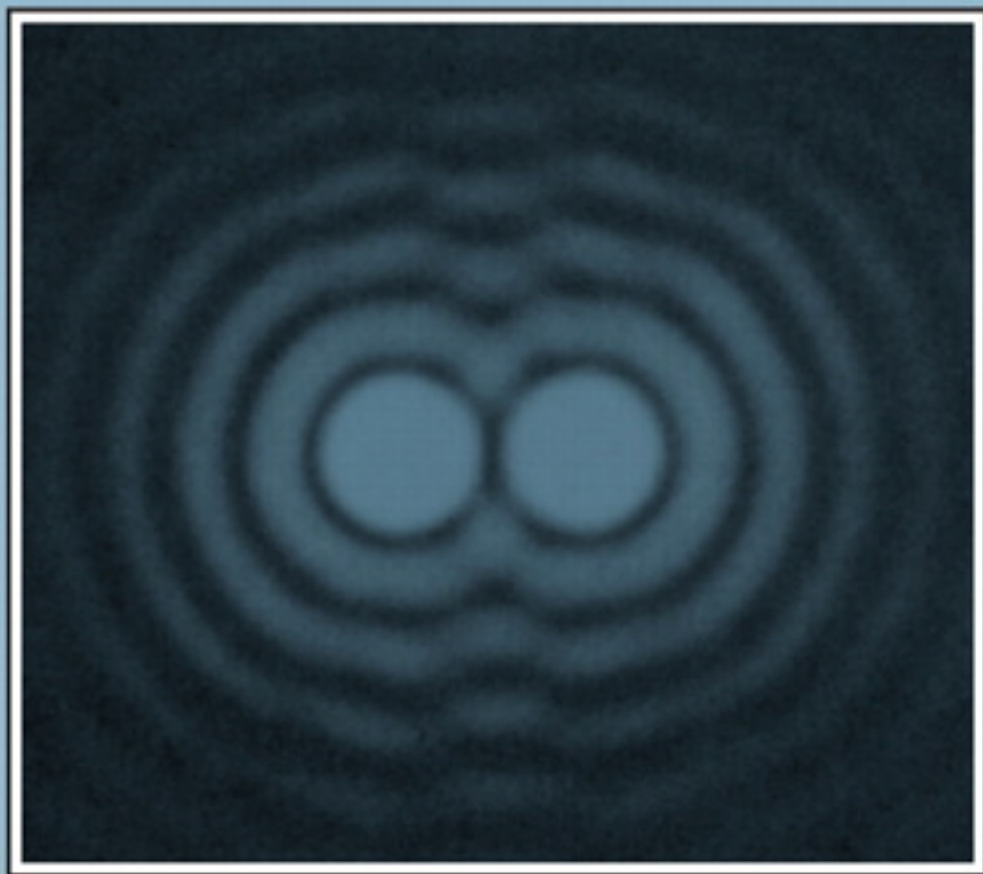


INTRODUCTION TO OPTICS

THIRD EDITION



Frank L. Pedrotti, S.J. • Leno M. Pedrotti • Leno S. Pedrotti

Third Edition

Introduction to Optics

FRANK L. PEDROTTI, S.J.

LENO M. PEDROTTI

LENO S. PEDROTTI

This page intentionally left blank

PHYSICAL CONTSTANTS

Speed of light	$c = 2.998 \times 10^8 \text{ m/s}$
Electron charge	$e = 1.602 \times 10^{-19} \text{ C}$
Electron rest mass	$m = 9.109 \times 10^{-31} \text{ kg}$
Planck constant	$h = 6.626 \times 10^{-34} \text{ Js}$
Boltzmann constant	$k = 1.3805 \times 10^{-23} \text{ J/K}$
Permittivity of vacuum	$\epsilon_0 = 8.854 \times 10^{-12} \text{ C}^2/\text{N}\cdot\text{m}^2$
Permeability of vacuum	$\mu_0 = 4\pi \times 10^{-7} \text{ T}\cdot\text{m/A}$

This page intentionally left blank

List of Tables

Table 1.1	Radiometric Terms	11
Table 2.1	Summary of Gaussian Mirror and Lens Formulas	39
Table 3.1	Fraunhofer Lines	64
Table 3.2	Standard Relative Apertures and Irradiance Available on Cameras	71
Table 6.1	Laser Parameters for Several Common Lasers	159
Table 8.1	Fabry-Perot Figures of Merit	217
Table 8.2	Fabry-Perot Parameters	217
Table 10.1	Characterization of Several Optical Fibers	248
Table 12.1	Fabry-Perot Interferometer and Diffraction Grating Figures of Merit	298
Table 13.1	Fresnel Integrals	323
Table 14.1	Summary of Jones Vectors	340
Table 14.2	Summary of Jones Matrices	346
Table 15.1	Refractive Indices for Several Materials	359
Table 15.2	Specific Rotation of Quartz	364
Table 15.3	Refractive Indices for Quartz	366
Table 18.1	Summary of Some Simple Ray-Transfer Matrices	404
Table 18.2	Cardinal Point Locations in Terms of System Matrix Elements	410
Table 18.3	Meridional Ray-Tracing Equations	415
Table 19.1	Radiometric and Photometric Terms	423

Table 19.2	Constants of a Schematic Eye	425
Table 20.1	Sample of Optical Glasses	455
Table 22.1	Refractive Indices for Several Coating Materials	485
Table 22.2	Reflectance of a High-Low Quarter-Wave Stack	488
Table 24.1	Linear and Nonlinear Processes	517
Table 24.2	Linear Electro-optic Coefficients for Representative Materials	519
Table 24.3	Kerr Constant for Selected Materials	523
Table 24.4	Verdet Constant for Selected Materials	525
Table 26.1	Laser Diode Wavelengths	578

Contents

	<i>Physical Constants</i>	<i>iii</i>
	<i>List of Tables</i>	<i>v</i>
1	<i>Nature of Light</i>	1
	Introduction	1
	1.1 A Brief History	2
	1.2 Particles and Photons	4
	1.3 The Electromagnetic Spectrum	6
	1.4 Radiometry	11
	Problems	15
2	<i>Geometrical Optics</i>	16
	Introduction	16
	2.1 Huygens' Principle	17
	2.2 Fermat's Principle	20
	2.3 Principle of Reversibility	22
	2.4 Reflection in Plane Mirrors	22
	2.5 Refraction Through Plane Surfaces	23

2.6	Imaging by an Optical System	25
2.7	Reflection at a Spherical Surface	27
2.8	Refraction at a Spherical Surface	32
2.9	Thin Lenses	35
2.10	Vergence and Refractive Power	39
2.11	Newtonian Equation for the Thin Lens	42
2.12	Cylindrical Lenses	42
	Problems	46
3	<i>Optical Instrumentation</i>	50
	Introduction	50
3.1	Stops, Pupils, and Windows	50
3.2	A Brief Look at Aberrations	58
3.3	Prisms	60
3.4	The Camera	69
3.5	Simple Magnifiers and Eyepieces	75
3.6	Microscopes	79
3.7	Telescopes	82
	Problems	89
4	<i>Wave Equations</i>	94
	Introduction	94
4.1	One-Dimensional Wave Equation	94
4.2	Harmonic Waves	96
4.3	Complex Numbers	99
4.4	Harmonic Waves as Complex Functions	100
4.5	Plane Waves	100
4.6	Spherical Waves	102
4.7	Other Harmonic Waveforms	103
4.8	Electromagnetic Waves	104
4.9	Light Polarization	108

4.10	Doppler Effect	110
	Problems	111
5	<i>Superposition of Waves</i>	113
	Introduction	113
5.1	Superposition Principle	113
5.2	Superposition of Waves of the Same Frequency	114
5.3	Random and Coherent Sources	119
5.4	Standing Waves	120
5.5	The Beat Phenomenon	123
5.6	Phase and Group Velocities	125
	Problems	129
6	<i>Properties of Lasers</i>	131
	Introduction	131
6.1	Energy Quantization in Light and Matter	132
6.2	Thermal Equilibrium and Blackbody Radiation	135
6.3	Nonlaser Sources of Electromagnetic Radiation	138
6.4	Einstein's Theory of Light-Matter Interaction	143
6.5	Essential Elements of a Laser	146
6.6	Simplified Description of Laser Operation	149
6.7	Characteristics of Laser Light	153
6.8	Laser Types and Parameters	158
	Problems	161
7	<i>Interference of Light</i>	163
	Introduction	163
7.1	Two-Beam Interference	163
7.2	Young's Double-Slit Experiment	169
7.3	Double-Slit Interference with Virtual Sources	173
7.4	Interference in Dielectric Films	175

7.5	Fringes of Equal Thickness	180
7.6	Newton's Rings	181
7.7	Film-Thickness Measurement by Interference	182
7.8	Stokes Relations	184
7.9	Multiple-Beam Interference in a Parallel Plane	185
	Problems	189
8	<i>Optical Interferometry</i>	192
	Introduction	192
8.1	The Michelson Interferometer	193
8.2	Applications of the Michelson Interferometer	196
8.3	Variations of the Michelson Interferometer	198
8.4	The Fabry-Perot Interferometer	199
8.5	Fabry-Perot Transmission: The Airy Function	201
8.6	Scanning Fabry-Perot Interferometer	206
8.7	Variable-Input-Frequency Fabry-Perot Interferometers	211
8.8	Lasers and the Fabry-Perot Cavity	213
8.9	Fabry-Perot Figures of Merit	216
8.10	Gravitational Wave Detectors	217
	Problems	220
9	<i>Coherence</i>	224
	Introduction	224
9.1	Fourier Analysis	224
9.2	Fourier Analysis of a Finite Harmonic Wave Train	228
9.3	Temporal Coherence and Line Width	230
9.4	Partial Coherence	231
9.5	Spatial Coherence	237
9.6	Spatial Coherence Width	238
	Problems	241

10	<i>Fiber Optics</i>	243
	Introduction	243
10.1	Applications	243
10.2	Communications System Overview	244
10.3	Bandwidth and Data Rate	246
10.4	Optics of Propagation	246
10.5	Allowed Modes	249
10.6	Attenuation	251
10.7	Distortion	253
10.8	High-Bit-Rate Optical-Fiber Communications	260
	Problems	264
11	<i>Fraunhofer Diffraction</i>	267
	Introduction	267
11.1	Diffraction from a Single Slit	268
11.2	Beam Spreading	273
11.3	Rectangular and Circular Apertures	274
11.4	Resolution	279
11.5	Double-Slit Diffraction	281
11.6	Diffraction from Many Slits	284
	Problems	289
12	<i>The Diffraction Grating</i>	292
	Introduction	292
12.1	The Grating Equation	292
12.2	Free Spectral Range of a Grating	293
12.3	Dispersion of a Grating	295
12.4	Resolution of a Grating	296
12.5	Types of Gratings	298
12.6	Blazed Gratings	299
12.7	Grating Replicas	301

12.8	Interference Gratings	302
12.9	Grating Instruments	303
	Problems	305
13	<i>Fresnel Diffraction</i>	308
	Introduction	308
13.1	Fresnel-Kirchhoff Diffraction Integral	308
13.2	Criterion for Fresnel Diffraction	311
13.3	The Obliquity Factor	312
13.4	Fresnel Diffraction from Circular Apertures	312
13.5	Phase Shift of the Diffracted Light	316
13.6	The Fresnel Zone Plate	316
13.7	Fresnel Diffraction from Apertures with Rectangular Symmetry	318
13.8	The Cornu Spiral	320
13.9	Applications of the Cornu Spiral	324
13.10	Babinet's Principle	330
	Problems	331
14	<i>Matrix Treatment of Polarization</i>	333
	Introduction	333
14.1	Mathematical Representation of Polarized Light: Jones Vectors	334
14.2	Mathematical Representation of Polarizers: Jones Matrices	341
	Problems	347
15	<i>Production of Polarized Light</i>	350
	Introduction	350
15.1	Dichroism: Polarization by Selective Absorption	350
15.2	Polarization by Reflection from Dielectric Surfaces	353
15.3	Polarization by Scattering	355
15.4	Birefringence: Polarization with Two Refractive Indices	357
15.5	Double Refraction	361

15.6	Optical Activity	363
15.7	Photoelasticity	367
	Problems	369
16	<i>Holography</i>	372
	Introduction	372
16.1	Conventional versus Holographic Photography	372
16.2	Hologram of a Point Source	373
16.3	Hologram of an Extended Object	375
16.4	Hologram Properties	379
16.5	White-Light (Rainbow) Properties	379
16.6	Other Applications of Holography	381
	Problems	384
17	<i>Optical Detectors and Displays</i>	386
	Introduction	386
17.1	Thermal Detectors of Radiation	386
17.2	Quantum Detectors of Radiation	387
17.3	Image Detection	389
17.4	Optical Detectors: Noise and Sensitivity	390
17.5	Optical Displays	391
	Problems	394
18	<i>Matrix Methods in Paraxial Optics</i>	396
	Introduction	396
18.1	The Thick Lens	396
18.2	The Matrix Method	399
18.3	The Translation Matrix	400
18.4	The Refraction Matrix	400
18.5	The Reflection Matrix	401
18.6	Thick-Lens and Thin-Lens Matrices	402

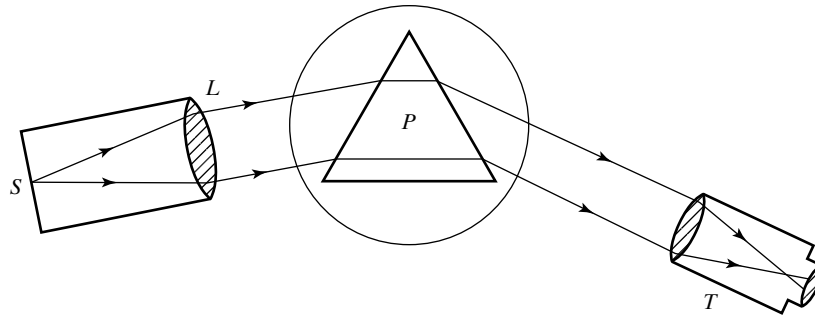
18.7	System-Ray Transfer Matrix	404
18.8	Significance of System Matrix Elements	406
18.9	Location of Cardinal Points for an Optical System	408
18.10	Examples Using the System Matrix and Cardinal Points	410
18.11	Ray Tracing	412
	Problems	416
19	<i>Optics of the Eye</i>	419
	Introduction	419
19.1	Biological Structure of the Eye	419
19.2	Photometry	421
19.3	Optical Representation of the Eye	424
19.4	Functions of the Eye	425
19.5	Vision Correction with External Lenses	428
19.6	Surgical Vision Correction	434
	Problems	436
20	<i>Aberration Theory</i>	438
	Introduction	438
20.1	Ray and Wave Aberrations	439
20.2	Third-Order Treatment of Refraction at a Spherical Interface	440
20.3	Spherical Aberrations	444
20.4	Coma	447
20.5	Astigmatism and Curvature of Field	449
20.6	Distortion	451
20.7	Chromatic Aberration	451
	Problems	456
21	<i>Fourier Optics</i>	458
	Introduction	458
21.1	Optical Data Imaging and Processing	459

21.2	Fourier-Transform Spectroscopy	471
	Problems	474
22	<i>Theory of Multilayer Films</i>	476
	Introduction	476
22.1	Transfer Matrix	477
22.2	Reflectance at Normal Incidence	481
22.3	Two-Layer Antireflecting Films	483
22.4	Three-Layer Antireflecting Films	486
22.5	High-Reflective Layers	486
	Problems	489
23	<i>Fresnel Equations</i>	491
	Introduction	491
23.1	The Fresnel Equations	491
23.2	External and Internal Reflections	497
23.3	Phase Changes on Reflection	499
23.4	Conservation of Energy	502
23.5	Evanescent Waves	504
23.6	Complex Refractive Index	506
23.7	Reflection from Metals	507
	Problems	508
24	<i>Nonlinear Optics and the Modulation of Light</i>	510
	Introduction	510
24.1	The Nonlinear Medium	511
24.2	Second Harmonic Generation and Frequency Mixing	513
24.3	Electro-Optic Effects	517
24.4	The Faraday Effect	524
24.5	The Acousto-Optic Effect	526
24.6	Optical Phase Conjugation	529

24.7	Optical Nonlinearities in Fibers	531
	Problems	533
25	<i>Optical Properties of Materials</i>	535
	Introduction	535
25.1	Polarization of a Dielectric Medium	535
25.2	Propagation of Light Waves	539
25.3	Conduction Current in a Metal	544
25.4	Propagation of Light Waves in a Metal	544
25.5	Skin Depth	545
25.6	Plasma Frequency	546
	Problems	548
26	<i>Laser Operation</i>	549
	Introduction	549
26.1	Rate Equations	549
26.2	Absorption	553
26.3	Gain Media	557
26.4	Steady-State Laser Output	561
26.5	Homogeneous Broadening	564
26.6	Inhomogeneous Broadening	567
26.7	Time-Dependent Phenomena	569
26.8	Pulsed Operation	571
26.9	Some Important Laser Systems	575
26.10	Diode Lasers	577
	Problems	579
27	<i>Characteristics of Laser Beams</i>	582
	Introduction	582
27.1	Three-Dimensional Wave Equation and Electromagnetic Waves	582
27.2	Gaussian Beams	583

27.3	Spot Size and Radius of Curvature of a Gaussian Beam	586
27.4	Characteristics of Gaussian Beams	587
27.5	Modes of Spherical Mirror Cavities	591
27.6	Laser Propagation Through Arbitrary Optical Systems	593
27.7	Higher-Order Gaussian Beams	600
	Problems	605
 <i>References</i>		 607
 <i>Answers to Selected Problems</i>		 611

This page intentionally left blank



3

Optical Instrumentation

INTRODUCTION

The principles of geometrical optics are applied in this chapter in order to discuss several practical optical instruments. The discussion begins with an introduction to the operation of *stops*, *pupils*, and *windows*, of great practical importance to *light control* in optical instrumentation. We follow this with a brief overview of *aberrations* and then examine in turn, the optics and operation of prisms, cameras, eyepieces, microscopes, and telescopes.

1 STOPS, PUPILS, AND WINDOWS

You should be familiar with ways to trace rays through an optical system using the step-by-step application of Gaussian formulas and ray tracing. However, not every light ray from an object point, directed toward or into an optical system, reaches the final image. Depending on the location of the object point and the ray angle, many of these rays are blocked by the limiting apertures of lenses and mirrors or by physical apertures intentionally inserted into the optical system. An *aperture*, in its broadest sense, is an opening defined by a geometrical boundary. In this section, we wish to concentrate on the effects of such spatial limitations of light beams in an optical system.

The apertures dealt with are often purposely inserted into an optical system to achieve various practical purposes. Apertures can be used to modify the effects of spherical aberration, astigmatism, and distortion. In other applications, apertures may be introduced to produce a sharp border to the image, like the sharp outline we see looking into the eyepiece of an optical instrument. Apertures may also be used to shield the image from

undesirable light scattered from optical components. In any case, apertures are inevitably present because every lens or mirror has a finite diameter that effectively introduces an aperture into the system.

The presence of apertures in an optical system influences its image-forming properties in two important ways—by limiting the *field of view* and by controlling the *image brightness*. The first limitation determines how much of the surface of a broad object can be seen by looking back through the optical system; the second determines how bright the image can be, that is, how much irradiance (W/m^2) reaches the image. Both of these limitations depend directly on the optical behavior of bundles of rays that leave points on an object and thread their way through an optical system—and its aperture—to a conjugate image point.

There exists in the optical industry today many different practical optical systems, each with many choices available for the placement of apertures. In this introduction we limit ourselves to a basic description of how apertures can affect both image brightness and field of view. We examine first the role of *aperture stops* and their related *pupils* in fixing image brightness, then the role of *field stops* and their related *windows* in determining a field of view.

Although the actions of stops, pupils, and windows depend quite clearly on the principles of geometrical optics, the details can at times be confusing. In the descriptive text and correlated figures that follow, it will benefit the reader to coordinate the reading of the text and examination of associated figures closely. Thus, for example, in the text and correlated Figures 1a, b, and c, we define carefully the meaning of aperture stops (AS), entrance pupils (E_nP), and exit pupils (E_xP) and illustrate their effects in several simple optical systems.

Image Brightness: Aperture Stops and Pupils

Aperture Stop (AS) The *aperture stop* of an optical system is the actual physical component that limits the size of the *maximum cone* of rays—from an axial object point to a conjugate image point—that can be processed by the entire system. Thus it controls the brightness of the image. The diaphragm of a camera and the iris of the human eye are examples of aperture stops. Another example is found in the telescope, in which the first, or objective, lens determines how much light is accepted by the telescope to form a final image on the retina of the eye. In the telescope, then, the objective lens is the aperture stop of the optical system. However, the aperture stop is not always identical with the first component of an optical system. For example, refer to Figure 1a. As shown, the aperture stop (AS) in front of the lens determines the extreme (or *marginal*) rays that can be accepted by the lens. But if the object OO' is moved to a position nearer to AS, at some point the lens rim becomes the limiting aperture. This position is just the point of intersection between the optical axis and a line drawn from the lens rim through the edge of the aperture stop AS. In this case, the angle ($L'OL$) subtended by the lens rim at O becomes smaller than the angle ($M'OM$) subtended by the edge of the aperture, and so we designate the lens as the aperture stop.

Entrance Pupil (E_nP) The *entrance pupil* is the *limiting aperture* (opening) that the light rays “see,” looking into the optical system from any object point. In Figure 1a, this is simply the aperture stop itself, so in this case, AS and E_nP are identical. To see that this is not always the case, look at Figure 1b, where the aperture stop is behind the lens (a *rear stop*), as in most cameras. Which component now limits the cone of light rays? You can see that it is that component whose aperture edges limit rays from O to their smallest angle relative to the axis. Looking into the optical system from object space, one

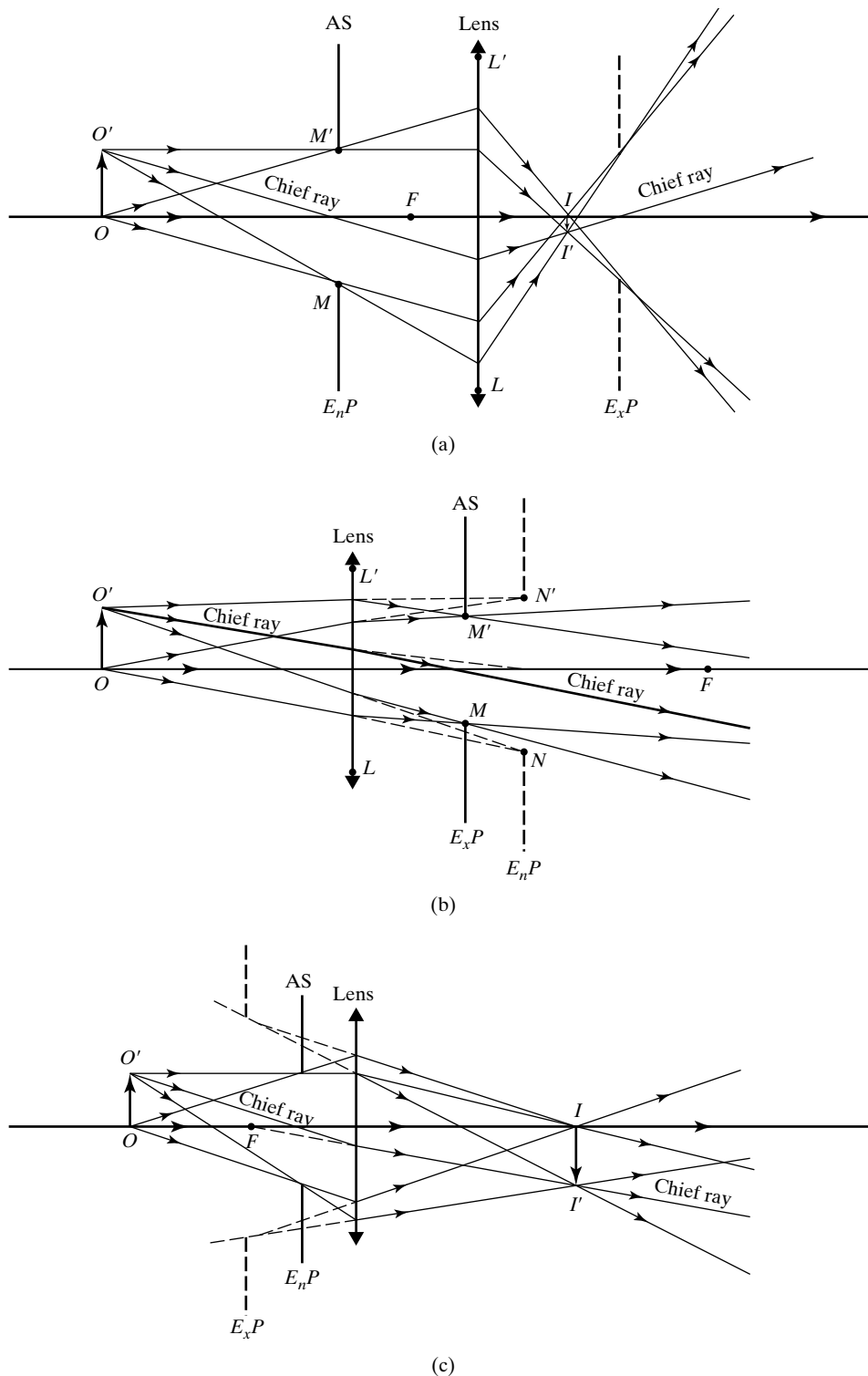


Figure 1 Limitation of light rays by various combinations of a positive lens and aperture.

sees the lens directly but sees the AS through the lens. In other words, the *effective aperture* due to the AS is its *image* formed by the lens, the dashed line marked E_nP . Since rays from O , directed toward this virtual aperture, make a smaller angle ($N'ON$) than rays directed toward the lens edge ($L'OL$), this virtual aperture serves as the effective entrance pupil for the system. Notice that rays from O , directed toward the edges of E_nP , are in fact first refracted by the lens so as to pass through the edges of the real aperture stop. This must be the case, since AS and E_nP are, by definition, conjugate

planes: The edges of E_nP are the images of the edges of AS. This example illustrates the general rule: *The entrance pupil E_nP is the image of the controlling aperture stop formed by the imaging elements preceding it.*¹ When the controlling aperture stop is the first such element (a *front stop*), it serves itself as the entrance pupil.

Another example, in which an aperture placed in front of the lens functions as the AS for the system, is shown in Figure 1c. It is different from Figure 1a in that the aperture is placed *inside* the focal length of the lens. Nevertheless, the aperture is the AS for the system because it, not the lens, limits the system rays to their smallest angle with the axis. Furthermore, it is the E_nP of the system because it is the first element encountered by the light from the object.

Exit Pupil (E_xP) We have described the E_nP of an optical system as the image of the AS one sees by looking into the optical system from the object. If one looks into the optical system from the image, another image of the AS can be seen that appears to limit the output beam size. This image is called the *exit pupil* of the optical system. Thus, *the exit pupil is the image of the controlling aperture stop formed by the imaging elements following it* (or to the right of it in our figures). The rear stop in Figure 1b is automatically the E_xP for the system because it is the last optical component. According to our definition of the E_xP , the exit pupil is the optical conjugate of the AS; the E_xP and AS are conjugate planes. It follows that the E_xP is also conjugate with the E_nP . In Figure 1a, the E_xP is the real image of the E_nP ; in Figure 1c, it is the virtual image. Notice that in each case, rays intersecting the edges of the entrance pupil also (actually, or when extended) intersect the edges of the exit pupil.

In a system like that of Figure 1a, a screen held at the position of the exit pupil receives a sharp image of the circular opening of the aperture stop. If the system represents the eyepiece of some optical instrument, the exit pupil is matched in position and diameter to the pupil of the eye. Notice further that if the screen is moved closer to the lens, it intercepts a sharp image II' of the object OO' . The exit pupil is seen to limit the solid angle of rays forming each point of the image and therefore determines the image brightness, point by point.

Chief Ray The *chief*, or *principal*, ray is a ray from an object point that passes through the *axial point*, in the *plane of the entrance pupil*. Given the conjugate nature of the entrance pupil with both the aperture stop and the exit pupil, this ray must also pass (actually or when extended) through their axial points. The chief ray in the cone of rays leaving object point O' is shown in all three systems of Figure 1. The chief ray in the cone of rays leaving the *axial point* O always coincides with the optical axis.

Before adding to our collection of new concepts that arise from a consideration of apertures in optical systems, we consider a system slightly more complex than those of Figure 1. In Figure 2, we specify a particular optical system consisting of two lenses, $L1$ and $L2$, with an aperture A placed between them, as shown. The first question to be answered is: Which element serves as the effective AS for the whole system? The answer to this question is not always obvious. It can always be answered, however, by determining which of the actual elements in the given system—in this case, A , $L1$, or $L2$ —has an *entrance pupil* that confines rays to their *smallest angle with the axis*, as seen from the object point. To decide which candidate presents the limiting aperture, it is necessary to find the entrance pupil for each by imaging each one through that part of the optical system *lying to its left*:

¹“Preceding” is used in the sense that light must pass through those imaging elements first. If we always use light rays directed from left to right, we can simply say, “by all imaging elements to its left.”

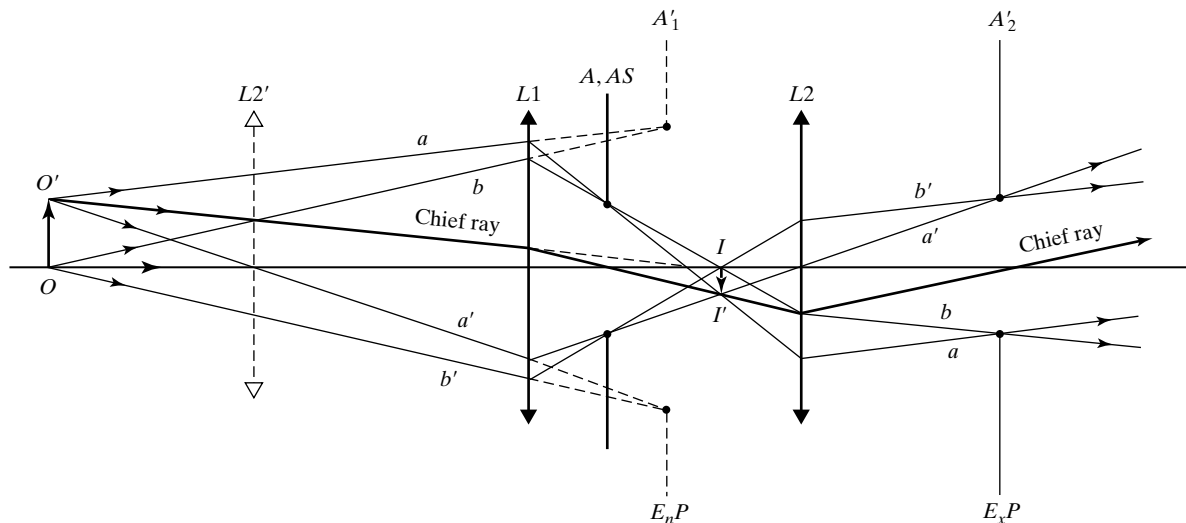


Figure 2 Limitation of light rays in an optical system consisting of two positive lenses and an aperture A . The labels a, a', b and b' assist in tracking the various rays.

$L2$: By ray diagram or by calculation, the image of lens $L2$, formed by $L1$ (as if light went from right to left), is $L2'$, and is real. Both its location and size (magnification) are shown.

A : The image of aperture A backward through $L1$ is virtual and is shown as A_1' .

$L1$: Since lens $L1$ is the first element, it acts as its own entrance pupil.

The three candidates for entrance pupils, $L2'$, $L1$, and A_1' , are next viewed from the axial point O . Since the rim of A_1' subtends the smallest angle at O , as is clear from the figure, we conclude that its conjugate aperture A is the aperture stop (AS) of the system.

Once the AS is identified, it is imaged through the optical elements to its right to find the exit pupil. In this case, aperture A is imaged through $L2$ to form A_2' . The chief ray, together with its two marginal rays, a and a' is drawn from the tip O' of the object. Notice that the chief ray passes (actually or when extended) through the centers of AS and its conjugate planes, E_nP and E_xP . The chief ray from O' intersects the optical axis at A , at A_1' (which is virtual), and at A_2' . The two cones—defined by the pairs of rays a, a' and b, b' —emerging from the points O and O' are limited by the size of the entrance pupil A_1' and just make it through the exit pupil A_2' . The image of OO' formed by $L1$ is shown as II' ; the final image (not shown) is virtual, since the rays from either O or O' diverge on leaving $L2$.

Field of View: Field Stops and Windows

In describing the limitations of a cone of rays from an axial object point, we have seen that entrance and exit pupils are related to the aperture stop and so govern the brightness of the image. As mentioned earlier, apertures also determine the field of view handled by the system. The controlling element in this connection is called the field stop, and it is related to an entrance window and an exit window in the same way that the aperture stop is related to entrance and exit pupils.

A simple experience of a limitation in the field of view is that of looking through an ordinary window. The edges of the window determine how much of the outdoors we can see. This field of view can be described in terms of the lateral dimensions of the object viewed, or in terms of the angular extent of the window, relative to the line of sight. One can talk about the field of view in terms of the object being viewed or in terms of the image formed at the viewer (on the retina, in this example).

To see how an aperture restricts the field of view—using diagrams that could well be applied to the case of window and eye lens, just discussed—look at Figure 3. In part (a), the optical system is a single aperture A placed in front of a single lens. Object and image planes are also shown. Rays from an axial point O are limited in angle by the aperture and focused by the lens at point O' . The same is true for the off-axis point T and its image, T' . In both cases, the lens is large enough to intercept the entire cone of rays admitted by A . If the object plane is uniformly bright and the aperture is a circular hole, then a circle of radius $O'T'$ is uniformly illuminated in the image plane.

However, if one considers object points below T , the upper rays from such points, passing through the aperture, miss the lens. Such a point, U , is shown in part (b) of the figure, the same optical system as (a), but redrawn for clarity. It is chosen such that the chief or central ray of the bundle from U just misses the top of the lens. About half the beam is lost so that image point U' receives only about half as much light as points O' and T' . Thus, the circular image begins to dim as its radius increases. This partial shielding of the outer portion of the image by the aperture for off-axis object points is called *vignetting*. Excessive vignetting may make the image of a point appear astigmatic. Finally, object point V is chosen such that all its rays through the aperture miss the lens entirely. The lateral field of view processed by this optical system is at most a circle of radius OV . It is often defined as the smaller circle of radius OU if one considers the usable field of view as that which consists of all object points that produce image points having at least half the maximum

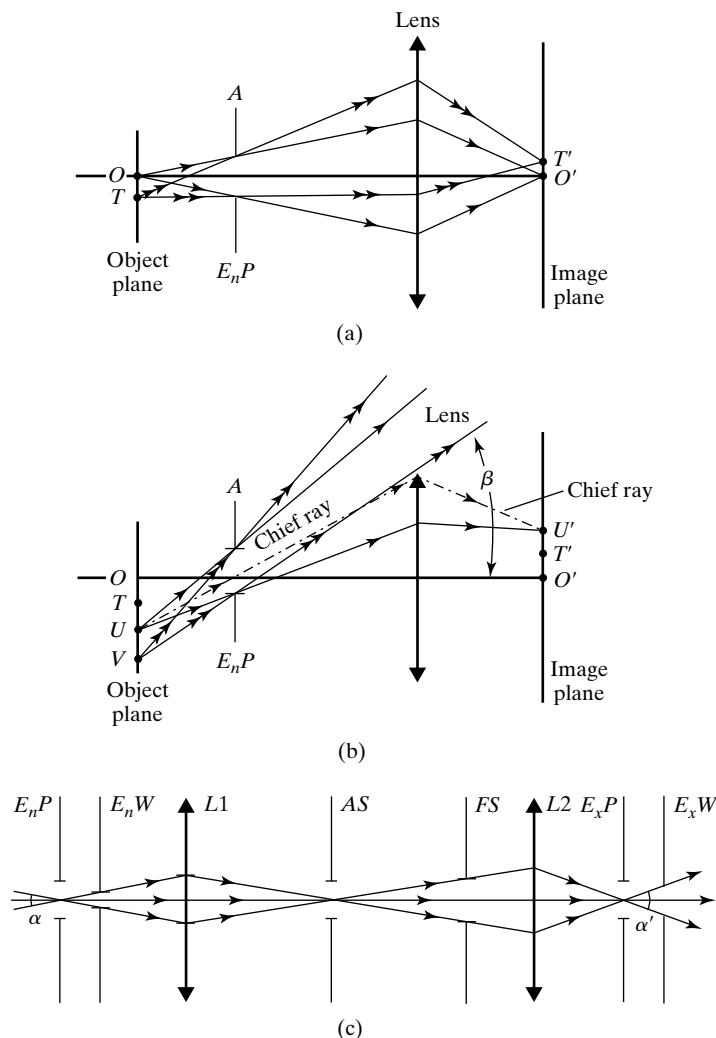


Figure 3 Referring to the same optical system, diagrams (a) and (b) illustrate both the way in which an aperture limits the field of view and the process by *vignetting*. Diagram (c) is an example of a more complex optical system, showing the angular field of view in object and image space, α and α' , respectively.

irradiance near the center. One can then also define the *angular field of view* as twice the angle β made by the chief ray with the axis at the center of the opening represented by the entrance pupil. The lens itself, in this case, acts both as the field stop and the entrance window. In other, more complex, systems, the relevant quantities may be described as follows.

Field Stop (FS) The field stop is the aperture that controls the field of view by limiting the solid angle formed by chief rays. *As seen from the center of the entrance pupil, the field stop (or its image) subtends the smallest angle.* When the edge of the field of view is to be sharply delineated, the field stop should be placed in an image plane so that it is sharply focused along with the final image. A simple example of such a field stop is the opening directly in front of the film that outlines the final image in a camera. Intentional limitation of the field of view using an aperture is desirable when either far-off axis imaging is of unacceptable quality due to aberrations or when vignetting severely reduces the illumination in the outer portions of the image.

Entrance Window (E_nW) *The entrance window is the image of the field stop formed by all optical elements preceding it.* The entrance window delineates the lateral dimensions of the object to be viewed, as in the viewfinder of a camera, and its angular diameter determines the angular field of view. When the field stop is located in an image plane, the entrance window lies in the conjugate object plane, where it outlines directly the lateral dimensions of the object field imaged by the optical system.

Exit Window (E_xW) *The exit window is the image of the field stop formed by all optical elements following it.* To an observer in image space, the exit window seems to limit the area of the image in the same way as an outdoor scene appears limited by the window of a room.

In Figure 3c, the locations of the stops, pupils, and windows are shown in a more complex optical system consisting of two lenses and two apertures. The first aperture is the AS of the system and, as we have seen, is related to an entrance pupil, its image in $L1$, and an exit pupil, its image in $L2$. The second aperture is the field stop, FS, with its corresponding images through the lenses: the entrance window to the left and the exit window to the right. The field of view in object space can then be described by α , the angle subtended by the entrance window at the center of the entrance pupil. Similarly, the field of view in image space can be described by α' , the angle subtended by the exit window at the center of the exit pupil. We see that the size of the field imaged by the optical system is effectively determined by the entrance window and, actually, by the size of the field stop. Notice that since E_nW and E_xW are both images of the FS, they are conjugate planes. Thus, the same bundle of rays that fills the entrance window also fills the field stop and the exit window.

The Summary of Terms that follows is provided as a convenient reference for a subject that requires patience, and experience with many examples, to master.

SUMMARY OF TERMS

Brightness

Aperture stop AS:	The <i>real</i> element in an optical system that limits the size of the cone of rays accepted by the system from an axial object point.
Entrance pupil E_nP :	The image of the aperture stop formed by the optical elements (if any) that precede it.
Exit pupil E_xP :	The image of the aperture stop formed by the optical elements (if any) that follow it.

Field of view

Field stop FS:	The <i>real</i> element that limits the angular field of view formed by an optical system.
----------------	--

- Entrance window E_nW : The image of the field stop formed by the optical elements (if any) that precede it.
- Exit window E_xW : The image of the field stop formed by the optical elements (if any) that follow it.

The following example will provide practice with the thin-lens formula $1/s + 1/s' = 1/f$ and procedures for determining stops and pupils for a specific optical system. As you follow through the steps in the solution, be sure to verify the correct use of the sign convention related to the object distance s , image distance s' , focal length f , and image magnification $m = -s'/s$ for each calculation.

Example 1

An optical system (see sketch in Figure 4 below) is made up of a positive thin lens L_1 of diameter 6 cm and focal length $f_1 = 6$ cm, a negative thin lens L_2 of diameter 6 cm and focal length $f_2 = -10$ cm, and an aperture A of diameter 3 cm. The aperture A is located 3 cm in front of lens L_1 , which is located 4 cm in front of lens L_2 . An object OP , 3 cm high is located 18 cm to the left of L_1 .

Problem

- Determine which element (A , L_1 , or L_2) serves as the aperture stop AS.
- Determine size and location of the entrance and exit pupils.
- Determine the location and size of the intermediate image of OP formed by L_1 and the final image formed by the system.
- Using a scale of $1 \text{ cm} = \frac{1}{4} \text{ in.}$, draw a diagram of the optical system and locate to scale, on the drawing, the two pupils, intermediate image $O'P'$ and final image $O''P''$.
- Draw the chief ray from object point P to its conjugate in the final image, P'' .

Solution

- Determine first which element (A , L_1 , or the image of L_2 in L_1) subtends the smallest half-angle from rim to point O .

Elements A and L_1 have no “optics” to their left, so each subtends a half-angle directly:

$$\text{For element } A: \theta_A \approx 1.5/15 = 0.1 \text{ rad}$$

$$\text{For element } L_1: \theta_{L_1} \approx 3/18 = 0.17 \text{ rad}$$

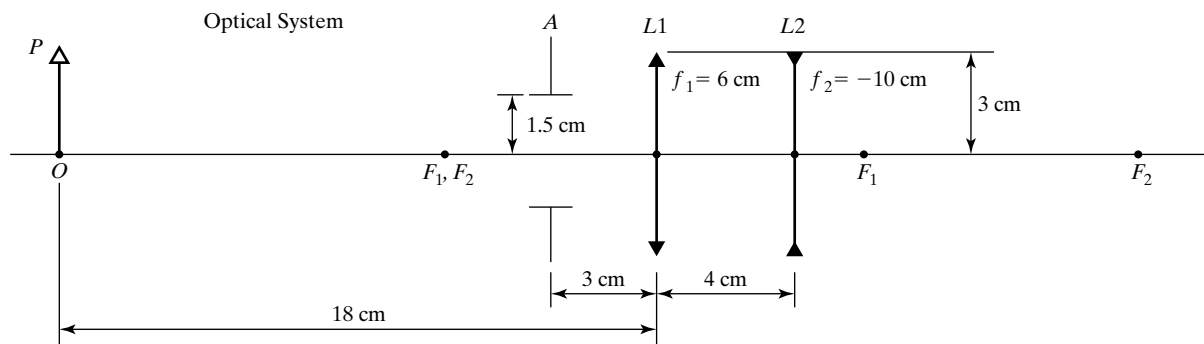


Figure 4 Sketch of optical system of Example 1.

Find the image of in with light traveling right to left so that $s = 4$ cm, $f = 6$ cm. Thus, $\frac{1}{4} + \frac{1}{s'} = \frac{1}{6}$ gives $s' = -12$ cm. So the image is virtual, 12 cm to the right of L_1 , of half-size given by $(3 \text{ cm}) \times \left(\frac{-s'}{s}\right) = (3 \text{ cm})\left(-\frac{-12}{4}\right) = 9$ cm. (This virtual image is shown as L'_2 in the final drawing.) So, for image L'_2 :

$$\theta_{L'_2} \simeq 9/30 = 0.3 \text{ rad}$$

Comparing half-angles, element A subtends the smallest half-angle and so serves as the aperture stop AS.

- b. Location and size of entrance and exit pupils.

Entrance Pupil E_nP : There are no optics to the left of the aperture stop A , so it serves also as the entrance pupil E_nP .

Exit Pupil E_xP : L_1 and L_2 are to the right of AS, so we must image AS through both to locate the position and size of the exit pupil E_xP .

Through lens L_1 : $s_1 = 3$ cm, $f_1 = 6$ cm so $\frac{1}{3} + \frac{1}{s'_1} = \frac{1}{6}$ gives $s'_1 = -6$ cm. This is a virtual image that serves as the object for lens L_2 , with $s_2 = 6 \text{ cm} + 4 \text{ cm} = 10$ cm and $f_2 = -10$ cm. Thus, $\frac{1}{10} + \frac{1}{s'_2} = \frac{1}{-10}$ yields $s'_2 = -5$ cm. So the exit pupil E_xP is located 5 cm to the left of L_2 or 1 cm to the left of L_1 . Its size is $(3 \text{ cm})\left(-\frac{-6}{3}\right)\left(-\frac{-5}{10}\right) = 3$ cm, as shown in the final drawing.

- c. Locating the intermediate and final image.

OP imaged through L_1 : $s_1 = 18$ cm, $f_1 = 6$ cm; $\frac{1}{18} + \frac{1}{s'_1} = \frac{1}{6}$ gives $s'_1 = 9$ cm right of L_1 or 5 cm to the right of L_2 . The size of $O'P'$, the intermediate image, is thus $(3 \text{ cm})\left(-\frac{9}{18}\right) = -1.5$ cm. So, $O'P'$ is inverted, 1.5 cm long, and 5 cm to the right of L_2 . (But it never forms there due to the presence of L_2 .)

$O'P'$ imaged through L_2 : s_2 , a virtual object for L_2 , equals -5 cm. So, $\frac{1}{-5} + \frac{1}{s'_2} = \frac{1}{-10}$ gives $s'_2 = 10$ cm. The final image $O''P''$ is then 10 cm to the right of L_2 . Its size, based on $O'P'$, is $(-1.5 \text{ cm})\left(-\frac{10}{-5}\right) = -3$ cm. So, $O''P''$, as shown in the final drawing, is real, inverted, 3 cm long (same as the object), and 10 cm from L_2 .

- d. The final drawing, based on an original scale of $1 \text{ cm} = \frac{1}{4} \text{ in.}$, with all items of interest, is shown in Figure 5.
- e. The chief ray, from point P to conjugate point P'' , is shown in the final drawing. Note that it leaves P , passes through M , the center of AS and E_nP , undergoes refraction at L_1 , heads for $O'P'$, refracts again at L_2 before reaching $O'P'$, and heads for $O''P''$, the final image. Note also that the segment of the chief ray from L_2 to P'' , if traced backward, will appear to be coming from point N , the center of the exit pupil E_xP . Thus, the chief ray involves the centers of AS, E_nP , and E_xP , as defined.

2 A BRIEF LOOK AT ABERRATIONS

All aberrations lead to a blurring of an image formed by an optical system, thereby frustrating the optical designer intent on producing an ideal image—that is, a faithful point-by-point re-creation of corresponding object points. *Chromatic* aberration and the five *monochromatic* aberrations (spherical aberration, coma, astigmatism, curvature of field, and distortion) occur largely as a result of the form and shape of lenses and mirrors and as a result of

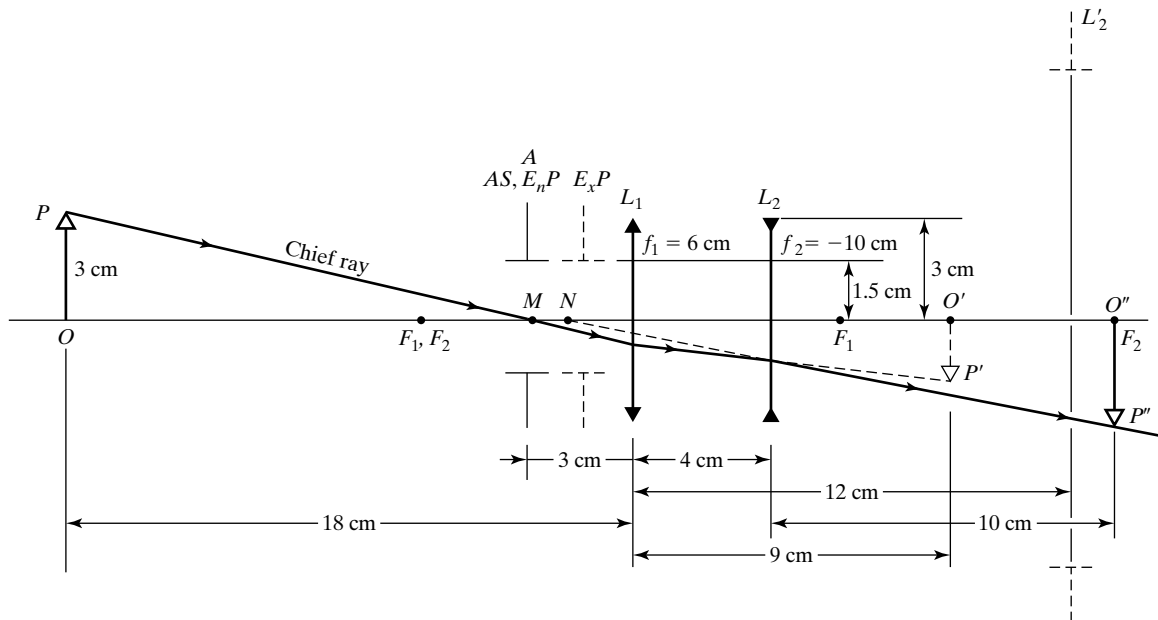


Figure 5 Solution to Example 1 (d).

rays from object points striking spherical surfaces at angles that exceed those set by the paraxial approximation.

A brief description of spherical and chromatic aberration will serve us here as a useful background for the treatment of optical elements and instruments that conclude this chapter.

Spherical Aberration

In the paraxial ray approximation, all rays emanating from an object point, after reflecting from a spherical mirror or passing through a lens with spherical surfaces, either intersect at, or to the viewer appear to intersect at, a common image point. In fact, rays emanating from an object point that are incident on an optical element (spherical mirror or lens) at different distances from the optical axis, after reflection or refraction, either intersect, or to the viewer appear to intersect, at different positions. The result is that point objects are not imaged as points but rather as small blurred lines. The effect of spherical aberration for a concave spherical mirror that images an axial object point is shown in Figure 6a. Note that rays incident on the mirror at symmetrically placed points Q converge at axial point M , whereas rays incident on the mirror at points P converge at axial point N . Consequently, light rays from a single point O form a blurred image along the line segment containing M and N . Figure 6b illustrates the effect of spherical aberration in a thin-lens system that forms a blurred line image of an axial point object. In the case shown, rays from axial point object O that encounter the lens at symmetrically placed points P converge at axial point N , while rays from O that encounter the lens at points Q converge at axial point M . As in the spherical mirror case, the result is a blurred line image along the line segment containing M and N . The constructions shown in Figure 6a and 6b suggest that “stopping down” the optical system, so that only nearly paraxial rays get through the entrance pupil, will limit the effect of spherical aberration. This remedy for spherical aberration is, of course, accompanied by a reduction in image brightness. Another remedy for spherical aberration is achieved by combining positive and negative lenses in an arrangement such that the spherical aberration from one tends to cancel that from the other.

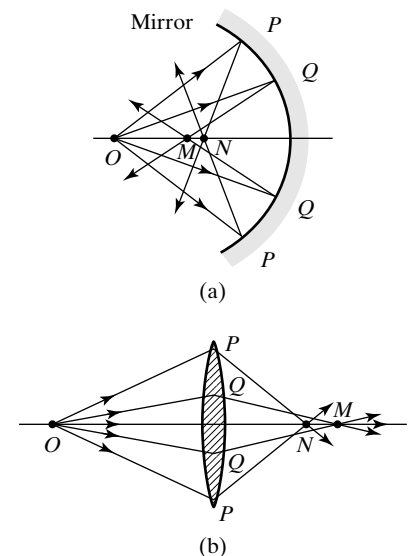


Figure 6 Spherical aberration (exaggerated) in (a) a spherical mirror and (b) a thin lens. For both arrangements, the rays from object point O fail to converge at a single image point.

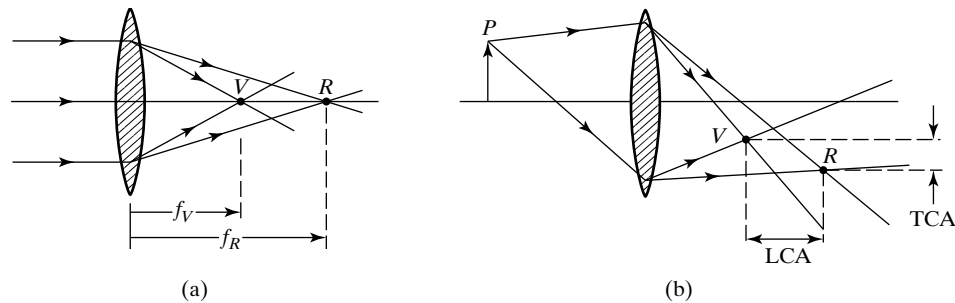


Figure 7 Chromatic aberration (exaggerated) for (a) parallel rays of white light incident on a thin lens and (b) white light incident on a thin lens from an off-axis point P .

Chromatic Aberration

Chromatic aberration results because the index of refraction of a material differs for different wavelengths. Since the focal length of a lens is dependent on the index of refraction of the lens material, focal lengths and image positions differ for different wavelength components of the light used in the optical system. Thus, polychromatic light from a point object images not as a point but as a series of points, one for each distinct wavelength.

Figure 7 illustrates two related aspects of chromatic aberration. In Figure 7a, parallel rays of light focus nearer the lens, at point V , for violet light and further from the lens, at point R , for red light. Thus, *white* light coming from a single distant object point fails to image as a single point. Rather, the different wavelength components refract to form image points between the focal lengths f_V and f_R , as indicated in Figure 7a. Figure 7b shows a slightly different view of chromatic aberration evident in the behavior of white light incident on a lens from an off-axis point P . The violet and red components of the white light leaving point P refract differently at the lens and so converge at different image points, labeled V and R , respectively. The amount of chromatic aberration, in such a case, can be described by two distances, shown in Figure 7b, one called the *longitudinal* chromatic aberration (LCA) and the other the *lateral* or *transverse* chromatic aberration (TCA).

Chromatic aberration in lenses can be effectively reduced by using multiple refractive elements of opposite powers. Of course, images formed in mirrors do not suffer from chromatic aberration since the focal length of a mirror is independent of wavelength.

3 PRISMS

Angular Deviation of a Prism

The top half of a double-convex, spherical lens can form an image of an axial object point within the paraxial approximation, as shown in Figure 8. If the lens surfaces are flat, a prism is formed, and paraxial rays can no longer produce a unique image point. It is nevertheless helpful in some cases to think of a prism as functioning approximately like one-half of a convex lens.

In the following we derive the relationships that describe exactly the progress of a single ray of light through a prism. The bending that occurs at each face is determined by Snell's law. The degree of bending is a function of the refractive index of the prism and is, therefore, a function of the wavelength of the incident light. The variation of refractive index and light speed with wavelength is called *dispersion* and is discussed later. For the present, we assume monochromatic light, which has its own characteristic refractive index in the prism medium. The relevant angles describing the progress of the ray through the prism are defined in Figure 9. Angles of incidence and refraction at each prism face are shown relative to the normals constructed at the point of intersection with the light ray. The total angular deviation δ of the ray

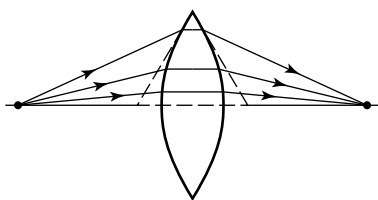


Figure 8 Focusing due to half of a convex lens approximates the action of a prism.

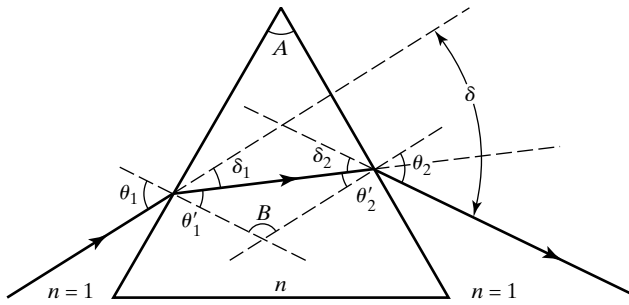


Figure 9 Progress of an arbitrary ray through a prism.

due to the action of the prism as a whole is the sum of the angular deviations δ_1 and δ_2 at the first and second faces, respectively. Snell's law at each prism face requires that

$$\sin \theta_1 = n \sin \theta'_1 \quad (1)$$

$$n \sin \theta'_2 = \sin \theta_2 \quad (2)$$

Inspection will show that the following geometrical relations must hold between the angles:

$$\delta_1 = \theta_1 - \theta'_1 \quad (3)$$

$$\delta_2 = \theta_2 - \theta'_2 \quad (4)$$

$$B = 180 - \theta'_1 - \theta'_2 = 180 - A \quad (5)$$

$$A = \theta'_1 + \theta'_2 \quad (6)$$

The two members of Eq. (5) follow because the sum of the angles of a triangle is 180° and because the sum of the angles of a quadrilateral must be 360° . Notice that the angles A and B and the two right angles formed by the normals with the prism sides constitute such a quadrilateral.

Using Eqs. (1) through (6), a programmable calculator or computer may easily be programmed to perform the sequential operations that finally determine the *angle of deviation*, δ . Assuming that the *prism angle* A and refractive index n are given, then the stepwise calculation for a ray incident at an angle θ_1 is as follows:

$$\theta'_1 = \sin^{-1}\left(\frac{\sin \theta_1}{n}\right) \quad (7)$$

$$\delta_1 = \theta_1 - \theta'_1 \quad (8)$$

$$\theta'_2 = A - \theta'_1 \quad (9)$$

$$\theta_2 = \sin^{-1}(n \sin \theta'_2) \quad (10)$$

$$\delta = \theta_1 + \theta_2 - \theta'_1 - \theta'_2 \quad (11)$$

The variation of deviation with angle of incidence for $A = 30^\circ$ and $n = 1.50$ is shown in Figure 10. Notice that a minimum deviation occurs for $\theta_1 = 23^\circ$. Refraction by a prism under the condition of minimum deviation is most often utilized in practice. We may argue rather neatly that when minimum deviation occurs, the ray of light passes symmetrically through the prism, making it unnecessary to subscript angles, as shown in Figure 11. Suppose this were not the case, and minimum deviation occurred for a nonsymmetrical case, as in Figure 9. Then if the ray were reversed, following the same path backward, it would have the same total deviation as the forward ray, which we

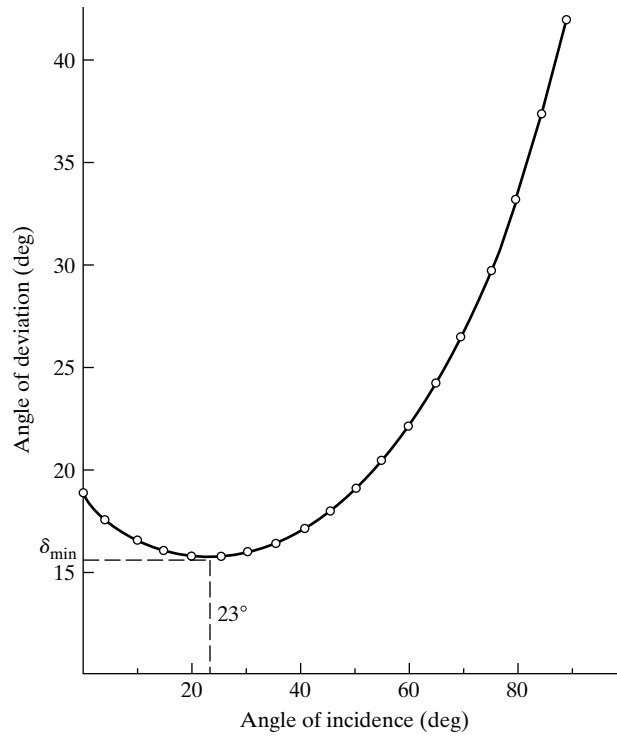


Figure 10 Graph of total deviation versus angle of incidence for a light ray through a prism with $A = 30^\circ$ and $n = 1.50$. Minimum deviation occurs for an angle of 23° .

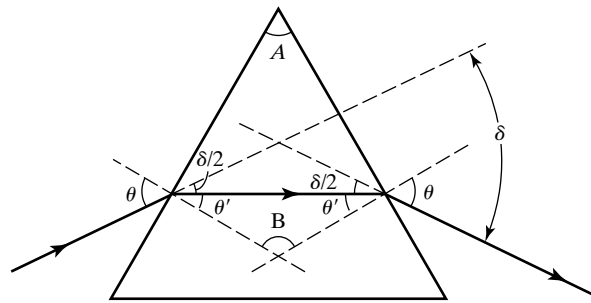


Figure 11 Progress of a ray through a prism under the condition of minimum deviation.

have supposed to be a minimum. Hence there would be two angles of incidence, θ_1 and θ_2 , producing minimum deviation, contrary to experience. The geometric relations simplify in this case. From Eq. (11),

$$\delta = 2\theta - 2\theta' \quad (12)$$

and from Eq. (6),

$$A = 2\theta' \quad (13)$$

Together these allow us to write

$$\theta' = \frac{A}{2} \quad \text{and} \quad \theta = \frac{\delta + A}{2} \quad (14)$$

so that Eq. (1) becomes

$$\sin\left(\frac{A + \delta}{2}\right) = n \sin\left(\frac{A}{2}\right)$$

or

$$n = \frac{\sin[(A + \delta)/2]}{\sin(A/2)} \quad (15)$$

Eq. (15) provides a method of determining the refractive index of a material that can be produced in the form of a prism. Measurement of both prism angle and minimum deviation of the sample determines n . An approximate form of Eq. (15) follows for the case of small prism angles and, consequently, small deviations. Approximating the sine of the angles by the angles in radians, we may write

$$n \cong \frac{(A + \delta)/2}{A/2}$$

or

$$\delta \cong A(n - 1), \text{ minimum deviation, small } A, \text{ prism in air} \quad (16)$$

For $A = 15^\circ$, the deviation given by Eq. (16) is correct to within about 1%. For $A = 30^\circ$, the error is about 5%.

Dispersion

The minimum deviation of a monochromatic beam through a prism is given implicitly by Eq. (15) in terms of the refractive index. The refractive index, however, depends on the wavelength, so that it would be better to write n_λ for this quantity. As a result, the total deviation is itself wavelength dependent, which means that various wavelength components of the incident light are separated on refraction from the prism. A typical *normal dispersion* curve and the nature of the resulting color separation are shown in Figure 12. Notice that shorter wavelengths have larger refractive indices and, therefore, smaller speeds in the prism. Consequently, violet light is deviated most in refraction through the prism. The dispersion indicated in the graph n versus λ of Figure 12 is called “normal” dispersion. When the refracting medium has characteristic excitations that absorb light of wavelengths within the range of the dispersion curve, the curve is monotonically decreasing, as shown, but has a positive slope in the wavelength region of the absorption. When this occurs, the term *anomalous dispersion* is used, although there is nothing anomalous about it. The normal dispersion curve shown is typical but varies somewhat for different materials. An empirical relation that approximates the curve, introduced by Augustin Cauchy, is

$$n_\lambda = A + \frac{B}{\lambda^2} + \frac{C}{\lambda^4} + \dots \quad (17)$$

where $A, B, C, \dots$ are empirical constants to be fitted to the dispersion data of a particular material. Often the first two terms are sufficient to provide a reasonable fit, in which case experimental knowledge of n at two distinct wavelengths is sufficient to determine values of A and B that represent the dispersion. The *dispersion*, defined as $dn/d\lambda$, is then approximately, using Cauchy's formula, $dn/d\lambda = -2B/\lambda^3$.

It is important to distinguish dispersion from deviation. Although prism materials of large n produce a large deviation at a given wavelength, the dispersion or separation of neighboring wavelengths need not be correspondingly large. Figure 13 depicts extreme cases illustrating the distinction. Historically, dispersion has been characterized by using three wavelengths of light near the middle and ends of the visible spectrum. They are called *Fraunhofer lines*. These lines were among those that appeared in the solar spectrum studied by J. von Fraunhofer. Their wavelengths, together with refractive indices, are given in Table 1. The *F* and *C* dark lines are due to absorption by hydrogen atoms, and the *D* dark line is due to absorption by the

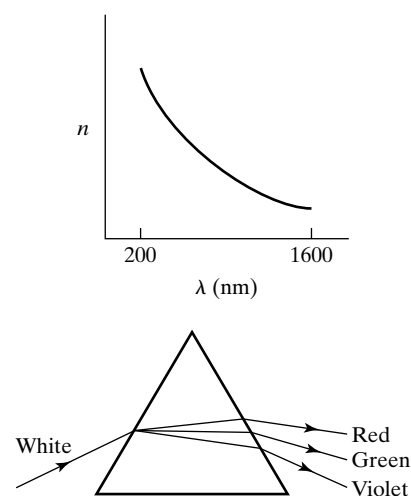


Figure 12 Typical normal dispersion curve and consequent color separation for white light refracted through a prism.

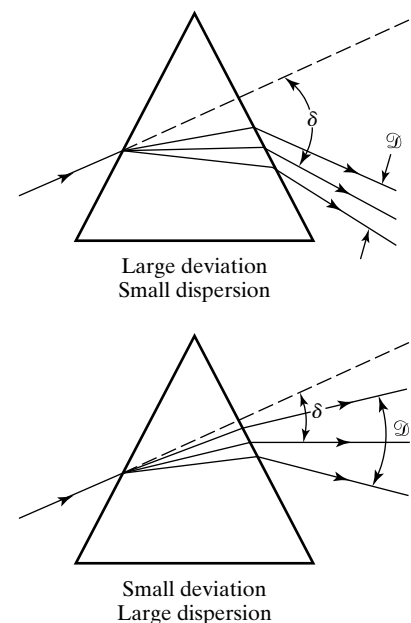


Figure 13 Extreme cases showing the dispersion Δ for three wavelengths and the deviation δ for the intermediate wavelength.

TABLE 1 FRAUNHOFER LINES

λ (nm)	Characterization	n	
		Crown glass	Flint glass
486.1	F , blue	1.5286	1.7328
589.2	D , yellow	1.5230	1.7205
656.3	C , red	1.5205	1.7076

sodium atoms in the sun's outer atmosphere.² Using the thin prism at minimum deviation for the D line, for example, the ratio of angular spread of the F and C wavelengths to the deviation of the D wavelength, as suggested in Figure 13, is

$$\frac{\mathfrak{D}}{\delta} = \frac{n_F - n_C}{n_D - 1}$$

This measure of the ratio of dispersion $\mathfrak{D}$ to deviation δ is defined as the *dispersive power* Δ , so that

$$\Delta = \frac{n_F - n_C}{n_D - 1} \quad (18)$$

Using Table 1, we may calculate the dispersive power of crown glass to be $1/65$, while that of flint glass is $1/29$, more than twice as great. The reciprocal of the dispersive power is known as the *Abbe number*.

Prism Spectrometers

An analytical instrument employing a prism as a dispersive element, together with the means of measuring the prism angle and the angles of deviation of various wavelength components in the incident light, is called a *prism spectrometer*. Its essential components are shown in Figure 14. Light to be analyzed is focused onto a narrow slit S and then collimated by lens L and refracted by the prism P , which typically rests on a rotatable indexed platform. Rays of light corresponding to each wavelength component emerge mutually parallel after refraction by the prism and are viewed by a telescope focused for infinity. As the telescope is rotated around the indexed prism table, a focused image of the slit is seen for each wavelength component at its corresponding angular deviation. The deviation δ is measured relative to the telescope position when viewing the slit without the

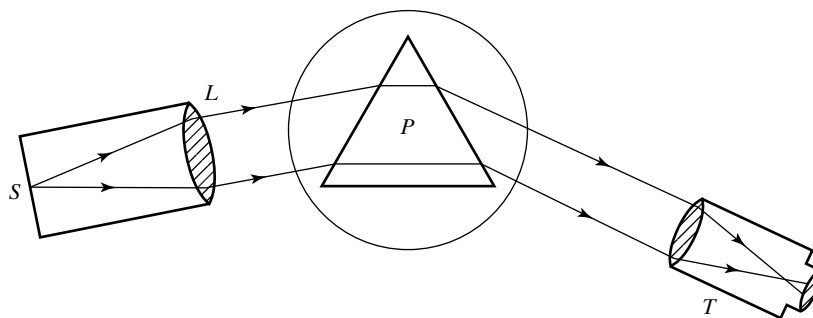


Figure 14 Essentials of a spectrometer.

²Because the yellow sodium D line is a doublet (589.0 and 589.6 nm), the more monochromatic d line of helium at 587.56 nm is often preferred to characterize the center of the visible spectrum. The green line of mercury at 546.07, which lies nearer to the center of the luminosity curve, is also used.

prism in place. When the instrument is used for visual observations without the capability of measuring the angular displacement of the spectral “lines,” it is called a *spectroscope*. If means are provided for recording the spectrum, for example, with a photographic film in the focal plane of the telescope objective, the instrument is called a *spectrograph*. When the prism is made of some type of glass, its wavelength range is limited by the absorption of glass outside the visible region. To extend the usefulness of the spectrograph farther into the ultraviolet, for example, prisms made from quartz (SiO_2) and fluorite (CaF_2) have been used. Wavelengths extending further into the infrared can be handled by prisms made of salt (NaCl , KCl) and sapphire (Al_2O_3).

Chromatic Resolving Power

If the wavelength difference between two components of the light incident on a prism is allowed to decrease, the ability of the prism to resolve them will ultimately fail. The resolving power of a prism spectrograph thus represents an important performance parameter, which we shall evaluate in this section. Imagine two spectral lines formed on a photographic film in a prism spectrograph. The lines are images of the slit, so that for precise wavelength measurements the entrance slit should be kept as narrow as possible consistent with the requirement of adequate illumination of the film. Even with the narrowest of slit widths, however, the spectral line image is found to possess a width, directly traceable to the limitation that the edges of the collimating lens or prism face impose on the light beam. The phenomenon is therefore due to the *diffraction* of light, treated later. Since the line images have an irreducible width due to diffraction, as $\Delta\lambda$ decreases and the lines approach one another, a point is reached where the two lines appear as one, and the *limit of resolution* of the instrument is realized. No amount of magnification of the images can produce a higher resolution or enhancement of the ability to discriminate between two such closely spaced spectral lines.

Consider Figure 15a, in which a monochromatic parallel beam of light is incident on a prism, such that it fills the prism face. Employing Fermat's principle, the ray FTW is isochronous with ray GTX , since they begin and end on the same plane wavefronts, GF and XW , respectively. Their optical paths can be equated to give

$$FT + TW = nb$$

where b is the base of the prism and n is the refractive index of the prism, corresponding to the wavelength λ . If a second neighboring wavelength component λ' is now also present in the incident beam, such that $\lambda' - \lambda = \Delta\lambda$, the component λ' will be associated with a different refractive index, $n' = n - \Delta n$. For normal dispersion, Δn will be a small, positive quantity. The

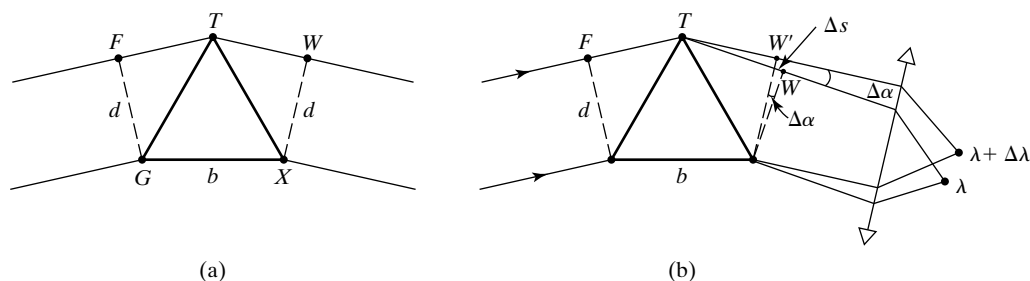


Figure 15 Constructions used to determine chromatic resolving power of a prism. (a) Refraction of monochromatic light. (b) Refraction of two wavelength components separated by $\Delta\lambda$.

emerging wavefronts for the two components, shown in Figure 15b, are thus separated by a small angular difference, $\Delta\alpha$, and are accordingly focused at different points in the focal plane of the telescope objective. Fermat's principle, applied to the second component λ' , gives

$$FT + TW - \Delta s = (n - \Delta n)b$$

Subtracting the last two equations, we conclude

$$\Delta s = b \Delta n \quad (19)$$

or, introducing the dispersion,

$$\Delta s = b \left(\frac{dn}{d\lambda} \right) \Delta \lambda \quad (20)$$

Equation (20) now relates the path difference Δs to the wavelength difference $\Delta \lambda$. The angular difference $\Delta \alpha$ can also be introduced, using

$$\Delta \alpha = \frac{\Delta s}{d} = \left(\frac{b}{d} \right) \left(\frac{dn}{d\lambda} \right) \Delta \lambda \quad (21)$$

where d is the beam width. We appeal now to *Rayleigh's criterion*, which determines the limit of resolution of the diffraction-limited line images. This criterion is explained and used in the later treatment of diffraction, where it is shown that the minimum separation $\Delta \alpha$ of the two wavefronts, such that the images formed are just barely resolvable, is given by

$$\Delta \alpha = \frac{\lambda}{d} \quad (22)$$

Combining Eqs. (21) and (22), therefore,

$$\frac{\lambda}{d} = \left(\frac{b}{d} \right) \left(\frac{dn}{d\lambda} \right) \Delta \lambda$$

or the minimum wavelength separation permissible for resolvable images is

$$(\Delta \lambda)_{\min} = \frac{\lambda}{b(dn/d\lambda)} \quad (23)$$

The *resolving power* provides an alternate way of describing the resolution limit of the instrument. By definition, the resolving power $\mathfrak{R}$

$$\mathfrak{R} = \frac{\lambda}{(\Delta \lambda)_{\min}} = b \frac{dn}{d\lambda} \quad (24)$$

where we have incorporated Eq. (23). Since dispersion is limited by the glass, prism resolving power might be improved by increasing the base b . However, this technique soon requires impractically large and heavy prisms. The dispersion $dn/d\lambda$ may be calculated, for example, from the Cauchy formula for the prism material, using Eq. (17).

Example 2

Determine the resolving power and minimum resolvable wavelength difference for a prism made from flint glass with a base of 5 cm.

Solution

With the help of Table 1 we can calculate an approximate average value of the dispersion $\frac{\Delta n}{\Delta \lambda}$ for $\lambda = 550$ nm as

$$\frac{\Delta n}{\Delta \lambda} = \frac{n_F - n_D}{\lambda_F - \lambda_D} = \frac{1.7328 - 1.7205}{486 - 589} = -1.19 \times 10^{-4} \text{ nm}^{-1}$$

Thus, the resolving power is

$$\mathfrak{R} = b \left(\frac{dn}{d\lambda} \right) = (0.05 \times 10^9 \text{ nm})(1.19 \times 10^{-4} \text{ nm}^{-1}) = 5971$$

The minimum resolvable wavelength difference in the region around 550 nm is, then,

$$(\Delta \lambda)_{\min} = \frac{\lambda}{\mathfrak{R}} = \frac{5550 \text{ Å}}{5971} \cong 1 \text{ Å}$$

Although grating spectrographs achieve higher resolving powers, they are generally more wasteful of light. Furthermore, they produce higher-order images of the same wavelength component, which can be confusing when interpreting spectral records. These instruments are discussed later.

Prisms with Special Applications

Prisms may be combined to produce *achromatic* overall behavior, that is, the net dispersion for two given wavelengths may be made zero, even though the deviation is not zero. On the other hand, a *direct vision prism*, Figure 16b, accomplishes zero deviation for a particular wavelength while at the same time providing dispersion. Schematics involving combinations of these two prism types are shown in Figure 16. The arrangement of prisms in Figure 16a, combined so that one prism cancels the dispersion of the other, can also be reversed so that the dispersion is additive, providing double dispersion.

A prism design useful in spectrometers is one that produces a constant deviation for all wavelengths as they are observed or detected. One example is the *Pellin-Broca prism*, illustrated in Figure 17. A collimated beam of light

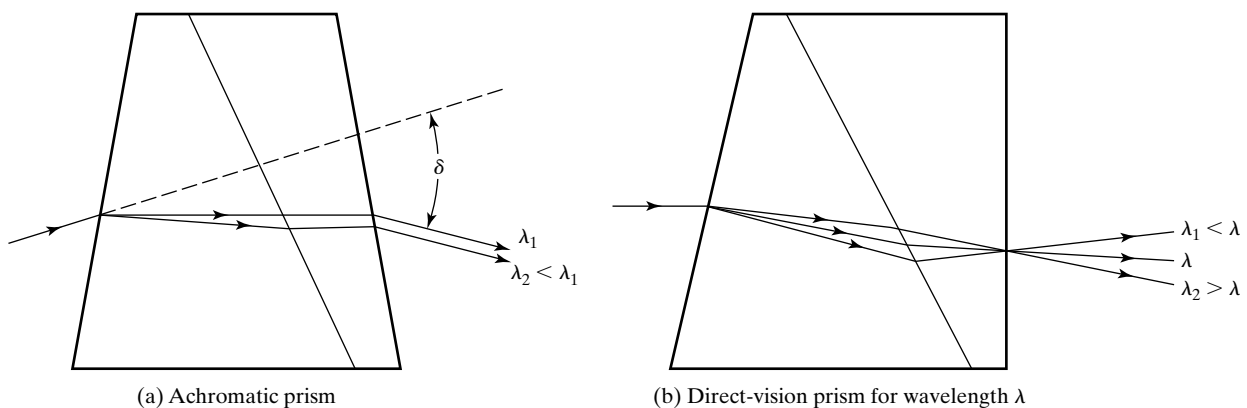


Figure 16 Nondispersive and nondeviating prisms.

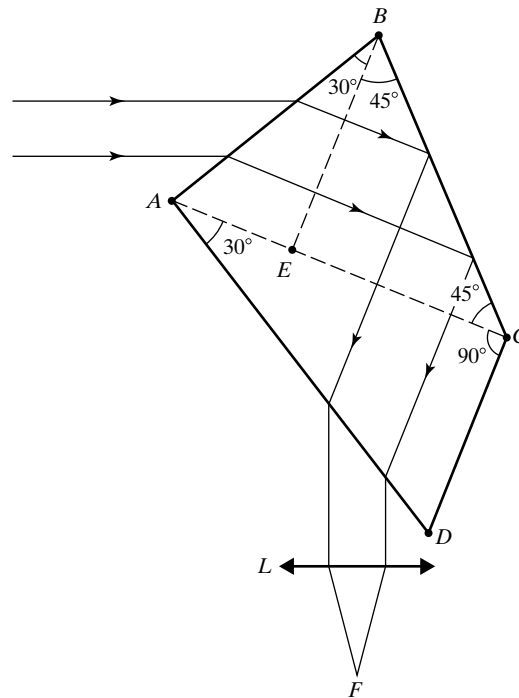


Figure 17 Pellin-Broca prism of constant deviation.

enters the prism at face AB and departs at face AD , making an angle of 90° with the incident direction. The dashed lines are merely added to assist in analyzing the operation of the prism, a single structure. Of the incident wavelengths, only one will refract at the precise angle that conforms to the case of minimum deviation, as shown, with the light rays parallel to the prism base AE . At face BC , total internal reflection occurs to direct the light beam into the prism section ACD , where it again traverses under the condition of minimum deviation. Since the prism section BEC serves only as a mirror, the beam passes effectively with minimum deviation through sections ABE and ACD , which together constitute a prism of 60° apex angle. In use, the spectral line is observed or recorded at F , the focal point of lens L . Thus, an observing telescope may be rigidly mounted. Instead, the prism is rotated on its prism table (or about an axis normal to the page), and as it rotates, various wavelengths in the incident beam successively meet the condition of incidence angle for minimum deviation, producing the path indicated, with focus at F . The prism rotation may be calibrated in terms of angle, or better, in terms of wavelength.

Reflecting Prisms

Total internally-reflecting prisms are frequently used in optical systems, both to alter the direction of the optical axis and to change the orientation of an image. Of course, prisms alone cannot produce images. When used in conjunction with image-forming elements, the light incident on the prism is first collimated and rendered normal to the prism face to avoid prismatic aberrations in the image. Plane mirrors may substitute for the reflecting prisms, but the prism's reflecting faces are easier to keep free of contamination, and the process of total internal reflection is capable of higher reflectivity. The stability in the angular relationship of prism faces may also be an important advantage in some applications. Some examples of reflecting prisms in use are illustrated in Figure 18. The *Porro prism*, Figure 18d, consists of two right-angle prisms, oriented in such a way that the face of one prism is partially revealed to receive the incident light and the face of the second prism is partially revealed to output the refracted light. The prism halves are separated in the figure to clarify its action. Images are inverted in both vertical and horizontal directions by the pair, so that the Porro prism is commonly used in binoculars to produce erect images.

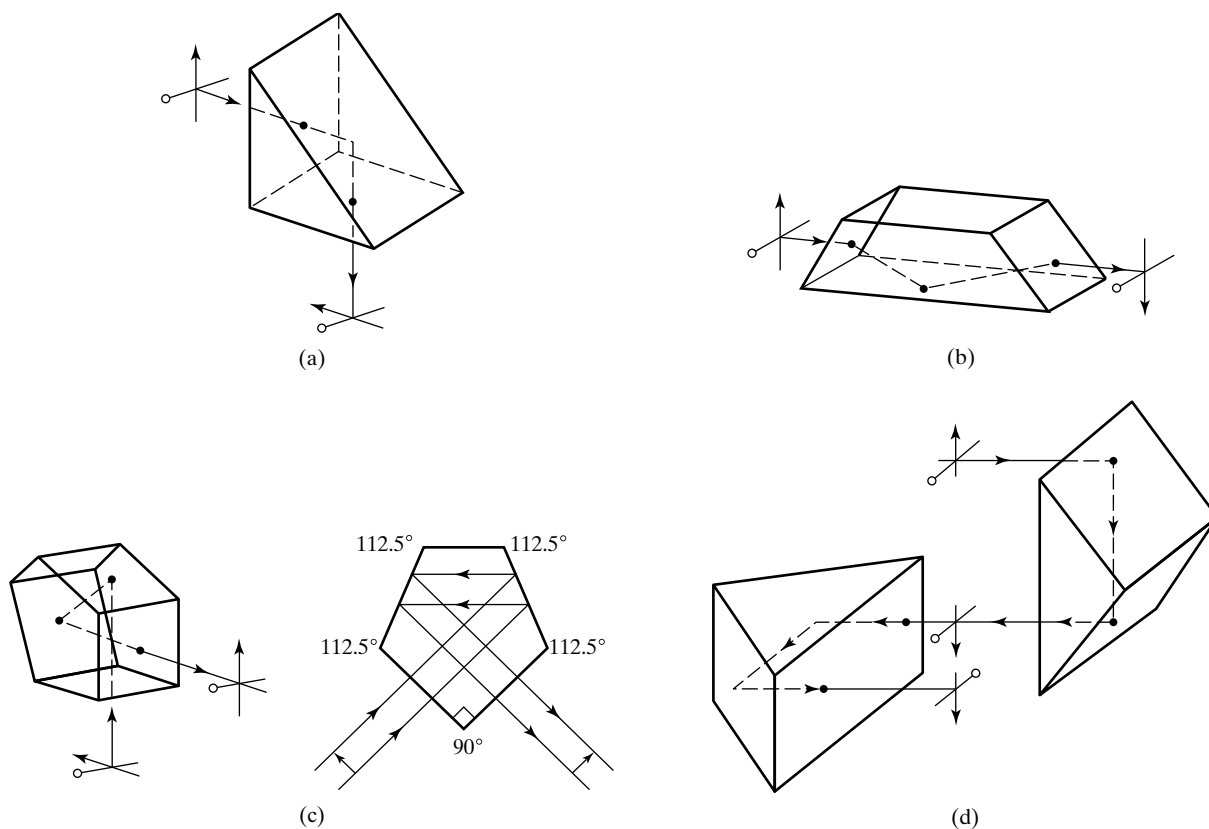


Figure 18 Image manipulation by reflecting prisms. (a) Right-angle prism. (b) Dove prism. (c) Penta prism; pentagonal cross section. (d) Porro prism.

4 THE CAMERA

The simplest type of camera is the pinhole camera, illustrated in Figure 19a. Light rays from an object are admitted into a light-tight box and onto a photographic film through a tiny pinhole, which may be provided with any simple means of shuttering, such as a piece of black tape. An image of the object is projected on the back wall of the box, which is lined with a piece of film.

As stated earlier, an image point is determined ideally when every ray from a given object point, each processed by the optical system, intersects at the corresponding image point. A pinhole does no focusing and actually blocks out most of the rays from each object point. Because of the smallness of the pinhole, however, every point in the image is reached only by rays that originate at *approximately* the same point of the object, as in Figure 19b. Alternatively, every object point sends a bundle of rays to the screen, which are limited by the small pinhole and so form a small circle of light on the screen, as in Figure 19a. The overlapping of these circles of light due to each object point maps out an image whose sharpness depends on the diameter of the individual circles. If they are too large, the image is blurred. Thus, as the pinhole is reduced in size, the image improves in clarity, until a certain pinhole size is reached. As the pinhole is reduced further, the images of each object point actually grow larger again due to diffraction, with consequent degradation of the image. Experimentally, one finds that the optimum pinhole size is around 0.5 mm when the pinhole-to-film distance is around 25 cm. The pinhole itself must be accurately formed in as thin an aperture as possible. A pinhole in aluminum foil, supported by a larger aperture, works well. The primary advantage of a pinhole camera (other than its elegant simplicity!) is that, since there is no focusing involved, all objects near and far are in focus

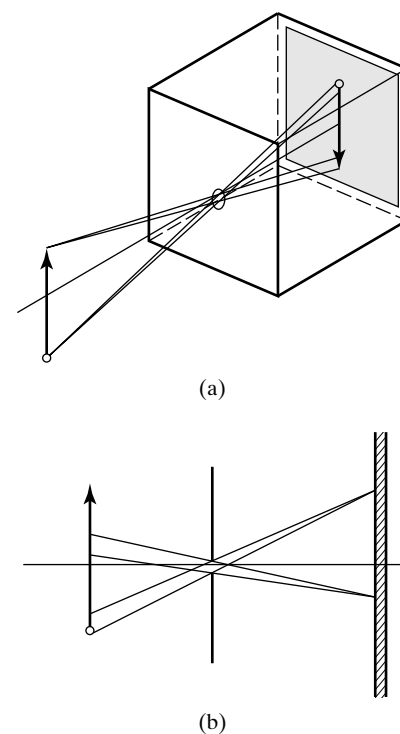


Figure 19 Imaging by a pinhole camera.

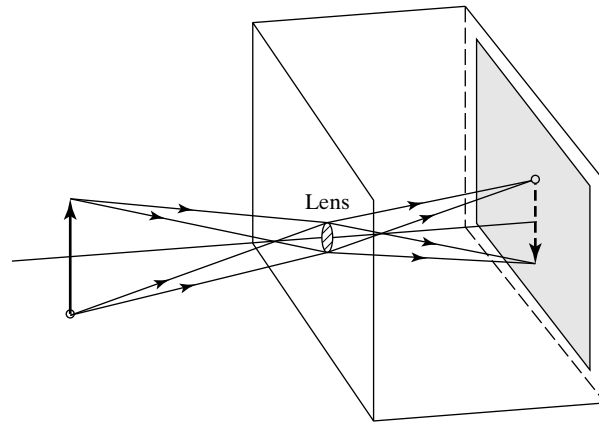


Figure 20 Simple camera.

on the screen. In other words, the *depth of field* of the camera is unlimited. The primary disadvantage is that, since the pinhole admits so little of the available light, exposure times must be long. The pinhole camera is not useful in freezing the action of moving objects. The pinhole-to-film distance, while not critical, does affect the sharpness of the image and the field of view. As this distance is reduced, the angular aperture seen by the film is larger, so that more of the scene is recorded, with corresponding decrease in size of any feature of the scene. Also, the image circles decrease in size, producing a clearer image.

If the pinhole aperture is opened sufficiently to accommodate a converging lens, we have the basic elements of the ordinary camera (Figure 20). The most immediate benefits of this modification are (1) an increase in the brightness of the image due to the focusing of *all* the rays of light from each object point onto its conjugate image point and (2) an increase in sharpness of the image, also due to the focusing power of the lens. The lens-to-film distance is now critical and depends on the object distance and lens focal length. For distant objects, the film must be situated in the focal plane of the lens. For closer objects, the focus falls beyond the film. Since the film plane is fixed, a focused image is procured by allowing the lens to be moved farther from the film, that is, by “focusing” the camera. The extreme possible position of the lens determines the nearest distance of objects that can be handled by the camera. “Close-ups” can be managed by changing to a lens with shorter focal length. Thus, the focal length of the lens determines the subject area received by the film and the corresponding image size. In general, image size is proportional to focal length. A *wide-angle lens* is a short focal-length lens with a large field of view. A *telephoto lens* is a long focal-length lens, providing magnification at the expense of subject area. The telephoto lens avoids a correspondingly “long” camera by using a positive lens, separated from a second negative lens of shorter focal length, such that the combination remains positive.

Also important to the operation of the camera is the size of its aperture, which admits light to the film. In most cameras, the aperture is variable and is coordinated with the exposure time (shutter speed) to determine the total exposure of the film to light from the scene. The light power incident at the image plane (irradiance E_e in watts per square meter) depends directly on (1) the area of the aperture and inversely on (2) the size of the image. If, as in Figure 21, the aperture is circular with diameter D and the energy of the light is assumed to be distributed uniformly over a corresponding image circle of diameter d , then

$$E_e \propto \frac{\text{area of aperture}}{\text{area of image}} = \frac{D^2}{d^2} \quad (25)$$

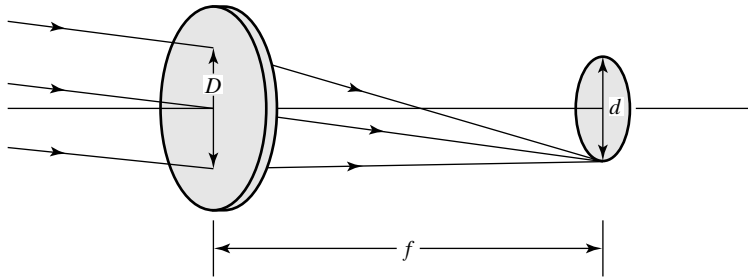


Figure 21 Illumination of image. The aperture (not shown) determines the useful diameter D of the lens.

The image size, as in Figure 21, is proportional to the focal length of the lens, so we can write

$$E_e \propto \left(\frac{D}{f}\right)^2 \quad (26)$$

The quantity f/D is the *relative aperture* of the lens (also called *f-number* or *f/stop*), which we symbolize by the letter A ,

$$A \equiv \frac{f}{D} \quad (27)$$

but is, unfortunately, usually identified by the symbol f/A . For example, a lens of 4-cm focal length that is stopped down to an aperture of 0.5 cm has a relative aperture of $A = 4/0.5 = 8$. This aperture is usually referred to by photographers as $f/8$. The irradiance is now

$$E_e \propto \frac{1}{A^2} \quad (28)$$

Most cameras provide selectable apertures that sequentially change the irradiance at each step by a factor of 2. The corresponding f -numbers, then, form a geometric series with ratio $\sqrt{2}$, as in Table 2. Larger aperture numbers correspond to smaller exposures. Since the total *exposure* (J/m^2) of the film depends on the product of irradiance ($\frac{\text{J}}{\text{m}^2 \cdot \text{s}}$) and time (s), a desirable total exposure may be met in a variety of ways. Accordingly, if a particular film (whose *speed* is described by an ISO number) is perfectly exposed by light from a particular scene with a shutter speed of $\frac{1}{50}$ s and a relative aperture of $f/8$, it will also be perfectly exposed by any other combination that gives the same total exposure, for example, by choosing a shutter speed of $\frac{1}{100}$ s and an aperture of $f/5.6$. The change in shutter speed cuts the total exposure in half, but opening the aperture to the next f /stop doubles the exposure, leaving no change in net exposure.

TABLE 2 STANDARD RELATIVE APERTURES AND IRRADIANCE AVAILABLE ON CAMERAS

$A = f\text{-number}$	$(A = f\text{-number})^2$	E_e
1	1	E_0
1.4	2	$E_0/2$
2	4	$E_0/4$
2.8	8	$E_0/8$
4	16	$E_0/16$
5.6	32	$E_0/32$
8	64	$E_0/64$
11	128	$E_0/128$
16	256	$E_0/256$
22	512	$E_0/512$

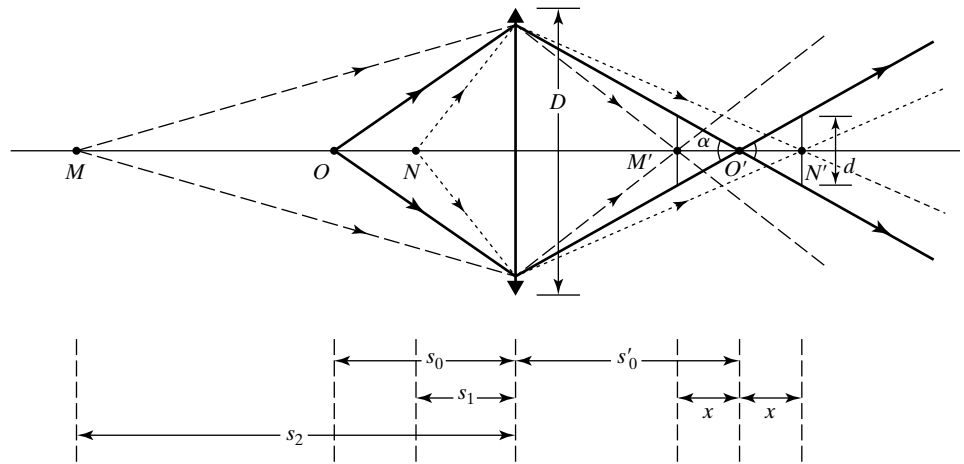


Figure 22 Construction illustrating depth of field MN . Object and image spaces are not shown to the same scale.

The particular combination of shutter speed and relative aperture chosen for an optimum total exposure is not always arbitrary. The shutter speed must be fast, of course, to capture an action shot without blurring the image. The choice of relative aperture also affects another property of the image, the *depth of field*. To define this quantity precisely, we utilize Figure 22, which shows an axial object point O at distance s_0 from a lens being imaged at O' , a distance s'_0 on the other side. All objects in the object plane are precisely focused in the image plane, disregarding the usual lens aberrations. Objects closer to (s_1), and farther from (s_2), the lens, however, send bundles of rays that focus farther ($s'_0 + x$) from and closer ($s'_0 - x$) to the image plane, respectively. Thus, a flat film, situated at distance s'_0 from the lens, intercepts *circles of confusion* corresponding to such object points. If the diameters of these circles are small enough, the resultant image is still acceptable. Suppose the largest acceptable diameter is d , as shown, such that all images within a distance x of the precise image are suitably “in focus.” The *depth of field* is then said to be the interval MN in object space conjugate to the interval $M'N'$, as shown. Notice that although the interval $M'N'$ is symmetric about s'_0 in image space, the depth of field interval (MN) is not symmetric about s_0 in object space.

The near-point and far-point distances, s_1 and s_2 , of the depth of field (MN) can be determined once the allowable blurring parameter d is chosen and the lens is specified by focal length and relative aperture. The angle α in Figure 22 may be specified in two ways,

$$\tan \alpha \cong \frac{D}{s'_0} \quad \text{and} \quad \tan \alpha \cong \frac{d}{x}$$

so that

$$x \cong \frac{ds'_0}{D} \quad (29)$$

It is then required to find, from the lens equation, the object distance s_1 corresponding to image distance $s'_0 + x$ and the object distance s_2 corresponding to image distance $s'_0 - x$. After a moderate amount of algebra, one finds

$$s_1 = \frac{s_0 f (f + Ad)}{f^2 + Ads_0} \quad (30)$$

$$s_2 = \frac{s_0 f (f - Ad)}{f^2 - Ads_0} \quad (31)$$

where the aperture is $A = f/D$. The depth of field, $MN = s_2 - s_1$, can be expressed as

$$\text{depth of field} = \frac{2Ads_0(s_0 - f)f^2}{f^4 - A^2d^2s_0^2} \quad (32)$$

Acceptable values of the circle diameter d depend on the quality of the photograph desired. A slide that will be projected or a negative that will be enlarged requires better original detail and hence a smaller value for d . For most photographic work, d is of the order of thousandths of an inch.

Example 3

A 5 cm focal length lens with an $f/16$ aperture is used to image an object 9 ft away. The blurring diameter in the image is chosen to be $d = 0.004$ mm.

Problem

Determine the location of the near point (s_1), far point (s_2), and the depth of field.

Solution

Based on Figure 22 and the given data, we have:

$$\begin{aligned} s_0 &= 9 \text{ ft} \simeq 275 \text{ cm} & d &= 0.004 \text{ cm} \\ f &= 5 \text{ cm} & A &= 16 \end{aligned}$$

For the near point, using Eq. (30),

$$s_1 = \frac{s_0f(f + Ad)}{f^2 + Ads_0} = \frac{(275)(5)[5 + 16(0.004)]}{25 + 16(0.004)(275)} \text{ cm} = 163.5 \text{ cm} \simeq 5.4 \text{ ft}$$

For the far point, using Eq. (31),

$$s_2 = \frac{s_0f(f - Ad)}{f^2 - Ads_0} = \frac{(275)(5)[5 - 16(0.004)]}{25 - 16(0.004)(275)} \text{ cm} = 1103 \text{ cm} \simeq 30 \text{ ft}$$

Thus, the depth of field, MN in Figure 22, is about 25 ft for a 5-cm focal length lens imaging an object 9 ft away. In effect this lens will image all objects from 5 ft to 30 ft with an acceptable sharpness.

Most cameras are equipped with a depth-of-field scale from which values of s_1 and s_2 can be read, once the object distance and aperture are selected. According to Eq. (32), depth of field is greater for smaller apertures (larger f -numbers), shorter focal lengths, and longer shooting distances.

The camera lens is called upon to perform a prodigious task. It must provide a large field of view, in the range of 35° to 65° for normal lenses and as large as 120° or more for wide-angle lenses. The image must be in focus and reasonably free from aberrations over the entire area of the film in the focal plane. The aberrations that must be reduced to an acceptable degree are, in addition to chromatic aberration, the five monochromatic aberrations: spherical aberration, coma, astigmatism, curvature of field, and distortion. Since a corrective measure for one type of aberration often causes greater degradation in the image due to another type of aberration, the optical solution represents one of many possible compromise lens designs. The labor involved in the design of a suitable lens that meets particular specifications within acceptable limits has been reduced considerably with the help of computer programming. Human ingenuity is nevertheless an essential component in the design task, since there

is more than one optical solution to a given set of specifications. The demands made upon a photographic lens cannot all be met using a single element. Various stages in solving the lens design problem are illustrated in Figure 23a, from the single-element meniscus lens, which may still be found in simple cameras, to the four-element Tessar lens. The use of a symmetrical placement of lenses, or groups of lenses, with respect to the aperture is often a distinctive feature of such lens designs. In such placements, one group may reverse the aberrations introduced by the other, reducing overall image degradation due to factors

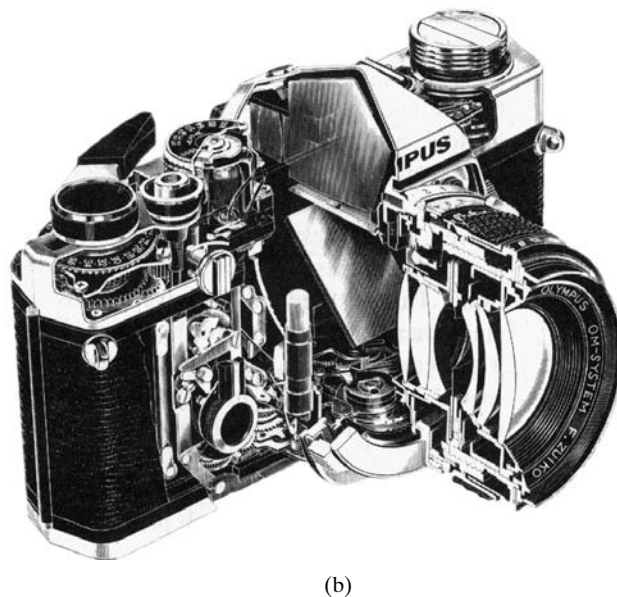
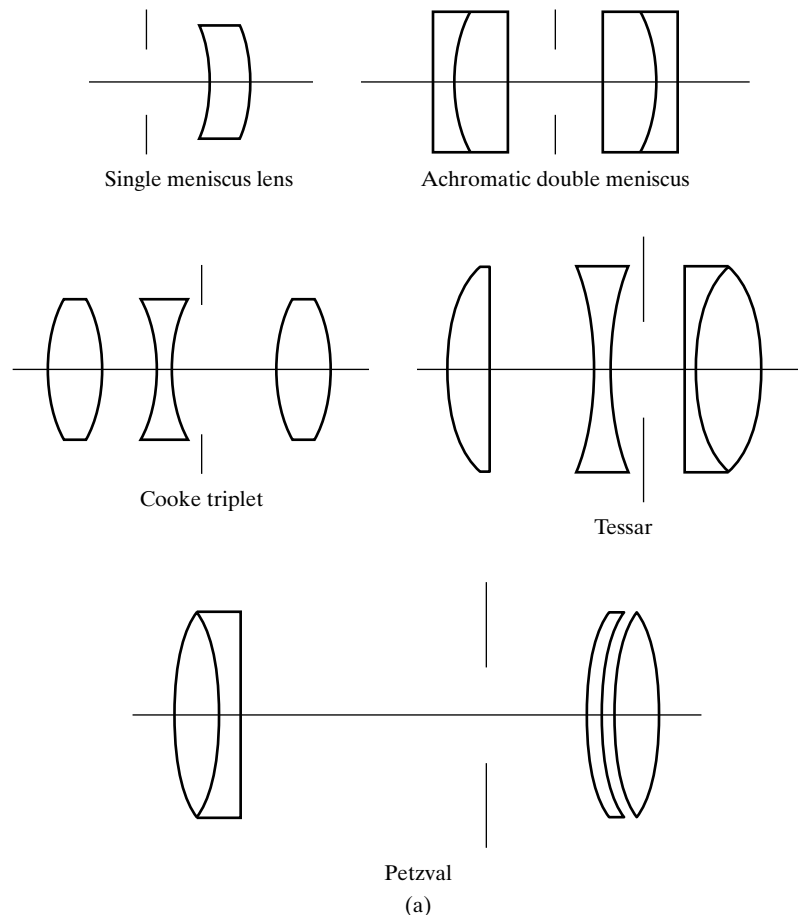


Figure 23 (a) Camera lens design. (b) Cut-away view of a 35-mm camera, revealing the multiple element lens. (Courtesy Olympus Corp., Woodbury, N.Y.)

such as coma, distortion, and lateral chromatic aberration. The multiple-element lens in a 35-mm camera is shown in the cutaway photo (Figure 23b).

5 SIMPLE MAGNIFIERS AND EYEPIECES

The simple magnifier is essentially a positive lens used to read small print, in which case it is often called a *reading glass*, or to assist the eye in examining small detail in a real object. It is often a simple convex lens but may be a doublet or a triplet, thereby providing for higher-quality images.

Figure 24 illustrates the working principle of the *simple magnifier*. A small object of dimension h , when examined by the *unaided* eye, is assumed to be held at the *near point* of the normal eye—nearest position of distinct vision—at position (a), 25 cm from the eye. At this position the object subtends an angle α_0 at the eye. To project a larger image on the retina, the simple magnifier is inserted and the object is moved physically closer to position (b), where it is at or just inside the focal point of the lens. In this position, the lens forms a virtual image subtending a larger angle α_M at the eye. The *angular magnification*³ of the simple magnifier is defined to be the ratio α_M/α_0 . In the paraxial approximation, the angles may be represented by their tangents, giving

$$\frac{\alpha_M}{\alpha_0} = \frac{h/s}{h/25} = \frac{25}{s}$$

If the image is viewed at infinity, $s = f$ and

$$M = \frac{25}{f} \quad \text{image at infinity} \quad (33)$$

At the other extreme, if the virtual image is viewed at the nearpoint of the eye, then $s' = -25$ cm, and from the thin-lens equation,

$$s = \frac{25f}{25 + f}$$

giving a magnification of

$$M = \frac{25}{f} + 1 \quad \text{image at normal near point} \quad (34)$$

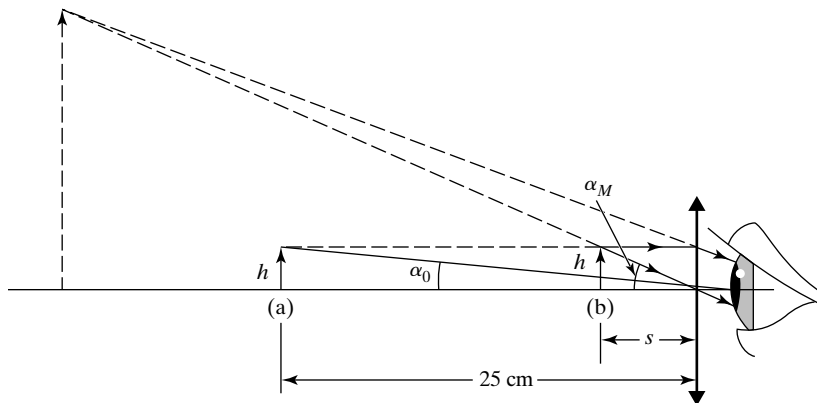


Figure 24 Operation of a simple magnifier.

³When viewing virtual images with optical instruments, the images may be at great distances, even “at infinity,” when rays entering the eye are parallel. In such cases, lateral magnifications also approach infinity and are not very useful. The more convenient *angular magnification* is clearly a measure of the image size formed on the retina and is used to describe magnification when eyepieces are involved, as in microscopes and telescopes.