

Geometrical Optics from the Ground up

Jan Vrbik

Department of Mathematics, Brock University, St. Catharines, Ontario, Canada

Email: jvrbik@brocku.ca

How to cite this paper: Vrbik, J. (2020) Geometrical Optics from the Ground up. *Applied Mathematics*, 11, 1021-1040. <https://doi.org/10.4236/am.2020.1110068>

Received: September 15, 2020

Accepted: October 23, 2020

Published: October 26, 2020

Copyright © 2020 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

We review, with proper derivation and proofs, the common undergraduate formulas for building images of objects using a system of lenses with spherical surfaces. This is done using the first-order approximation which assumes that light rays deviate from the symmetry axis by only small angles. Yet, even this most basic approximation results in surprisingly complex theory, which is then applied to explain workings of everyday optical instruments.

Keywords

Paraxial Approximation, Cardinal Points, Ray-Transfer Matrix, Keystone Distortion, Exit Pupil, Aperture and Vignetting

1. Introduction

Undergraduate Physics curriculum usually includes a brief introduction to Geometrical Optics, but students are often given a few basic formulas to learn only how to apply them, using graphs as the primary tool. The derivation of these formulas is left to more advanced books (such as [1]) whose level of presentation makes them too difficult as supplementary reading for most undergraduates (the corresponding literature is vast and easily searchable; we make no attempts at its review).

The article's goal is to bridge this gap and make the underlying theory more accessible; the only prerequisites are: a basic knowledge of vectors and two by two matrices, and some prior exposure to the topic itself, roughly at the level of [2] (a classic Physics textbook). Following individual steps of our reasoning should give students a better appreciation of the subject and demonstrate how intimately are realities of our world linked to Mathematics. Our presentation may be also of interest to people from other fields (Mathematics in particular), wishing to familiarize themselves with a new and interesting subject, without having to read an extensive monograph.

Throughout the presentation we assume that light is a collection of rays emanating, as straight lines, from a light source (all physical objects in daylight are such sources); the wavelike nature of light is totally ignored. When reaching a boundary between two optical media with different speeds of light (light travels faster in air than in water or glass), a ray instantaneously changes its direction before continuing in a straight-line path again. This is referred to as the light's refraction; it enables us to trace light rays passing through an optical system consisting of one or more lenses. Light can also be reflected by some surfaces due to space limitations, we do *not* discuss the related issue of optical mirrors and their applications. For the same reason, we do not include numerical examples and limit the number of graphs (a multitude of these is easily found on the internet; readers are also encouraged to draw their own, as an aid to proper understanding of individual formulas).

We then investigate what happens when a conical pencil of light rays emanating from a *point-like* source enters an axially symmetric system consisting of several different optical media, separated from each other by spherically shaped boundaries. We then trace a ray's path as it changes direction at each boundary until it emerges from the last surface. We thus discover that all these rays converge, to a good approximation, to a single point, which to our eyes appears to be a nearly perfect image of the original object. To end up with this, rather idealized picture, we need to assume that the rays' *directions* always deviate from the axis of symmetry by only *small* angles; we can then replace $\sin \alpha$ of any such angle by α itself (in radians). We also ignore the fact that light consists of different colours; we assume, slightly incorrectly, that colour does not affect the light's speed.

This enables us to show how properties of any such optical system can be summarized by the location of a handful of so-called cardinal points (focal and principal points in particular), derive formulas for finding the image of any specific object, and explain basic principles behind common optical instruments (loupe, camera, projector, microscope, telescope and periscope).

2. First-Order Approximation

When a ray of a *unit* direction \mathbf{v} (or, more explicitly $\langle v_x, v_y, v_z \rangle$; this is the notation we use later on) hits a smooth surface separating one optical medium from the next, the new unit direction becomes

$$\mathbf{u} = \frac{\mathbf{v} + \mathbf{m} \left(\sqrt{n^2 - 1 + (\mathbf{v} \cdot \mathbf{m})^2} - \mathbf{v} \cdot \mathbf{m} \right)}{n} \quad (1)$$

where n is the speed of light in the original medium, divided by the speed of light in the new medium, this is called the relative index of refraction, and \mathbf{m} is a unit vector, *normal* to the surface at the point of the ray's entry and oriented *into* the new medium; this implies that both $\mathbf{v} \cdot \mathbf{m}$ and $\mathbf{u} \cdot \mathbf{m}$ are positive. Note that all our directions and normal vectors are of unit length.

Proof. Basic Physics tells us that \mathbf{u} must be in the plane defined by \mathbf{v} and \mathbf{m} (having the form of $a\mathbf{v} + b\mathbf{m}$ clearly meets this condition), and it must also comply with Snell’s law, which reads

$$1 - (\mathbf{v} \cdot \mathbf{m})^2 = n^2 (1 - (\mathbf{u} \cdot \mathbf{m})^2) \tag{2}$$

where the dot products $\mathbf{v} \cdot \mathbf{m}$ and $\mathbf{u} \cdot \mathbf{m}$ provide the cosines of the corresponding angles. The usual version of the law states that the ratio of the sin’s of the two angles is equal to n ; our formulation is clearly equivalent.

To show that (1) complies with (2), take the dot product of each of its sides with \mathbf{m} (recall that $\mathbf{m} \cdot \mathbf{m} = 1$) and multiply by n , thus getting

$$n(\mathbf{u} \cdot \mathbf{m}) = \sqrt{n^2 - 1 + (\mathbf{v} \cdot \mathbf{m})^2} \tag{3}$$

Squaring each side of (3) and subtracting from n^2 yields (2).

Finally, (3) makes $\mathbf{u} \cdot \mathbf{m}$ positive, implying that \mathbf{u} has the correct orientation. Note that $n^2 - 1 + (\mathbf{v} \cdot \mathbf{m})^2$ must be positive to achieve refraction, rather than total reflection in this article, we do not need to consider the latter possibility. ■

When the x and y components of all three (\mathbf{u} , \mathbf{v} and \mathbf{m}) vectors are small, (1) can be, to a *linear* accuracy discarding second and higher powers of small quantities, simplified to

$$\langle u_x, u_y \rangle = \frac{\langle v_x, v_y \rangle + (n-1)\langle m_x, m_y \rangle}{n} \tag{4}$$

since, in this approximation, $\mathbf{v} \cdot \mathbf{m} \approx 1$. Note that the z component of *all* our unit vectors is approximately equal to 1, thus becoming a redundant piece of information (our notation simply leaves it out).

All subsequent equations are correct only to the linear accuracy; this is referred to as paraxial approximation.

2.1. Ray Tracing

Consider a ray that starts at $\langle x, y, 0 \rangle$, the so-called object (note that our “objects” are single points, but are often represented by an arrow from z axis to this point), and follows direction \mathbf{v} , at a small angle to the z axis (horizontal, in our graphs) until it reaches a spherical-surface medium of relative index n . Note that the sphere’s radius R is measured *from* the apex of this surface, whose z coordinate is g , to the sphere’s center, as in **Figure 1**; R is thus negative for *concave* surfaces and *infinite* for flat surfaces (most of our quantities may have negative values; this does not affect our formulas, they remain correct without a need for any modification).

The point of entry into the new medium is therefore $\langle x, y, 0 \rangle + g \langle v_x, v_y, 1 \rangle$, making the normal direction to the surface at this point equal to

$$\left\langle \frac{x + gv_x}{R}, \frac{y + gv_y}{R} \right\rangle \tag{5}$$

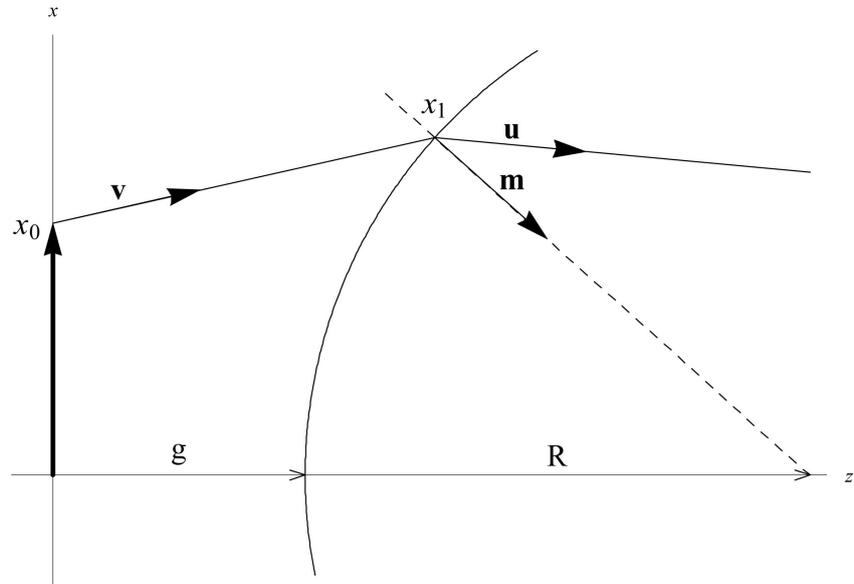


Figure 1. Ray refraction.

Combined with (4), this yields the ray’s new direction, namely

$$\left\langle \frac{v_x + \frac{(n-1)(x + gv_x)}{nR}}{n}, \frac{v_y + \frac{(n-1)(y + gv_y)}{nR}}{n} \right\rangle \tag{6}$$

Based on the last expression, we conclude that it is sufficient to follow a ray’s path in the x - z plane *only*, since its x and y components fully *decouple* (are algebraically independent of each other), implying that results obtained for the x component then have their automatic analog (just replace x by y and v_x by v_y) in the y direction. This makes it unnecessary to consider the y coordinate any further; from now on, we discard it and work strictly in the two-dimensional x - z plane. We can then correspondingly simplify our notation: v will stand for the original v_x , and our object may now be placed at $\langle x_0, 0, 0 \rangle$, instead of $\langle x, y, 0 \rangle$.

2.2. Matrix Notation

Based on these results, the x component of the ray’s new location and of its direction can then be conveniently computed using the usual transfer-matrix technique (see [3]); note that location and direction have become elements of a one-column matrix.

$$\begin{bmatrix} x_1 \\ v_1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \frac{1-n_1}{n_1 R_1} & \frac{1}{n_1} \end{bmatrix} \begin{bmatrix} 1 & g \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_0 \\ v_0 \end{bmatrix} \stackrel{\text{def}}{=} \mathbb{S}_1 \mathbb{T}_0 \begin{bmatrix} x_0 \\ v_0 \end{bmatrix} \tag{7}$$

where subscripts 0 and 1 now indicate the original and the new values respectively; similarly subscripting R and n implies that the ray has just entered the *first* such surface.

Extending this to a system of k axially symmetrical, spherical surfaces, each with its own R , n (relative to the *previous* medium) and distance q_j from the

apex of the previous, j^{th} surface (where $j = 1, 2, \dots, k - 1$), we get the x coordinate and direction of the ray at the point of leaving the *last* surface by a simple iteration (*i.e.* repeated application) of (7), namely

$$\begin{bmatrix} x_k \\ v_k \end{bmatrix} = \mathbb{S}_k \mathbb{T}_{k-1} \cdots \mathbb{S}_2 \mathbb{T}_2 \mathbb{S}_1 \mathbb{T}_1 \mathbb{S}_0 \begin{bmatrix} x_0 \\ v_0 \end{bmatrix} \tag{8}$$

where

$$\mathbb{S}_j := \begin{bmatrix} 1 & 0 \\ \frac{1-n_j}{n_j R_j} & \frac{1}{n_j} \end{bmatrix} \text{ and } \mathbb{T}_j := \begin{bmatrix} 1 & q_j \\ 0 & 1 \end{bmatrix} \tag{9}$$

It is convenient to rewrite (8) as

$$\begin{bmatrix} x_k \\ v_k \end{bmatrix} \stackrel{\text{def}}{=} \begin{bmatrix} A & C \\ B & D \end{bmatrix} \mathbb{T}_0 \begin{bmatrix} x_0 \\ v_0 \end{bmatrix} \tag{10}$$

(slightly departing from the usual notation of other authors). Note that the determinant of \mathbb{S}_j (of \mathbb{T}_j) is equal to $\frac{1}{n_j}$ (to 1); therefore, when the final

medium has the same speed of light as the original one (which is the most common situation, as the ambient medium of a system of lenses is usually the air), the product of all these determinants must equal to 1. This further implies that, in such a case, $\Delta := AD - CB = 1$. Nevertheless, most of our subsequent formulas are *not* making this assumption, to keep them fully general; note that Δ must be always positive.

2.3. Image Creation

When a ray originating at an object leaves the last surface of such a system and continues for a horizontal distance h , its location and direction are then given by

$$\begin{bmatrix} 1 & h \\ 0 & 1 \end{bmatrix} \begin{bmatrix} A & C \\ B & D \end{bmatrix} \begin{bmatrix} 1 & g \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_0 \\ v_0 \end{bmatrix} = \begin{bmatrix} (A + Bh)x_0 + (C + Dh + Ag + Bhg)v_0 \\ Bx_0 + (D + Bg)v_0 \end{bmatrix} \tag{11}$$

Now, consider a pencil (implying: many different v_0 values) of such rays originating at $\langle x_0, 0, 0 \rangle$ and converging to a *single* point again, thus creating an image of the original object where all these rays *intersect*. This clearly happens at the distance h_i , obtained by making the coefficient of v_0 in the expression for the final x location, the first component of (11), equal to 0, *i.e.* by solving

$$C + gA + Dh + gBh = 0 \tag{12}$$

for h . This gives

$$h_i = -\frac{gA + C}{gB + D} \tag{13}$$

yielding the following x coordinate of the resulting image

$$(A + Bh_i)x_0 = \frac{AD - CB}{gB + D}x_0 = \frac{\Delta}{gB + D}x_0 \stackrel{\text{def}}{=} Mx_0 \tag{14}$$

where M is the corresponding magnification (often negative, indicating that

the image is upside down or inverted); when $|M| < 1$, the image is *smaller* than the object, making “magnification” an euphemism. Note that h_i may also turn out to be negative; this then implies that the image is only virtual, as the outgoing rays actually *diverge* (only by extending them *backwards* makes them converge to an image; our eyes can see it there, but it cannot be captured on film). A real (as opposed to virtual) image *can* be imprinted onto a light-sensitive layer of “receptors” (to use a generic term), placed at the distance h_i from the last surface.

In this context, it is important to mention that rays preserve their colour and intensity (at least proportionately: a brighter object creates a brighter image) of the light of their source. Similarly, one has to realize that an optical system hardly ever deals with a single object, it is always a continuous multitude of them (we call it a “scenery”), transforming them into the corresponding multitude of images. Our formulas imply that images of objects with the same value of g (defining the so-called object plane) form a re-scaled (by factor M), but otherwise undistorted and sharp image of the original scenery at distance h_i from the last surface (in a z -perpendicular image plane), often captured on a planar layer of film. Point-like objects at distance $g' \neq g$ will appear on this film as small, uniformly illuminated *disks*, whose radius is proportional to $|g' - g|$; we call such images “out of focus”.

2.4. Focal Points

From (13), it is easy to see that, when the distance from the object to the first surface is

$$g = -\frac{D}{B} \stackrel{\text{def}}{=} f_1 \tag{15}$$

the image (*i.e.* the value of h_i) is at infinity, implying that the outgoing rays are parallel to each other; f_1 thus establishes the z location of the *first* focal plane, perpendicular to the z axis; this plane intersects the z axis at the first focal point (labelled f_1 in **Figure 2**; but otherwise, use definition (15) in all subsequent formulas).

Proof. The fact that the outgoing rays must be parallel follows from

$$\left(\begin{bmatrix} A & C \\ B & D \end{bmatrix} - \frac{D}{B} \begin{bmatrix} 0 & A \\ 0 & B \end{bmatrix} \right) \begin{bmatrix} x_0 \\ v_0 \end{bmatrix} = \begin{bmatrix} Ax_0 + \frac{\Delta}{B}v_0 \\ Bx_0 \end{bmatrix} \tag{16}$$

since the final *direction* is the same for all values of v_0 . ■

This implies that the final direction of *any* incoming ray *passing* through the first focal point (*i.e.* when $x_0 = 0$) is equal to 0, the ray must come out of the system parallel to z ; note that when f_1 is negative, the incoming ray needs to be extended to have it cross the focal point.

Similarly, as g approaches infinity, implying that the *incoming* rays from any such distant object become parallel to each other, the corresponding h_i tends to

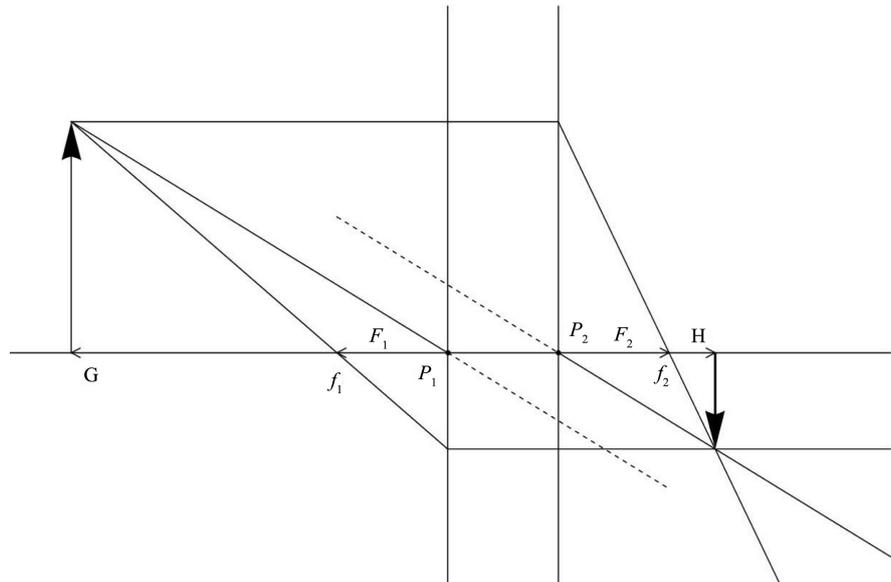


Figure 2. Image construction.

$$f_2 := -\frac{A}{B} \tag{17}$$

which yields the z distance *from* the last surface *to* the *second* focal plane, where the resulting image is located. In general: incoming *parallel* rays always converge to a single point in the second focal *plane*.

Proof. The x coordinate of a ray (which starts with arbitrary values of x_0 and v_0) at distance f_2 from the last surface is given by

$$\begin{bmatrix} 1 & -\frac{A}{B} \end{bmatrix} \left(\begin{bmatrix} A & C \\ B & D \end{bmatrix} + g \begin{bmatrix} 0 & A \\ 0 & B \end{bmatrix} \right) \begin{bmatrix} x_0 \\ v_0 \end{bmatrix} = -\frac{\Delta}{B} v_0 \tag{18}$$

Rays with the same value of v_0 (regardless of x_0) must therefore intersect at this same point. ■

This further implies that a ray originally parallel to z (*i.e.* having $v_0 = 0$) must pass, upon leaving the system, through the second focal *point*; when f_2 is negative, this statement becomes “virtual”, *i.e.* the outgoing straight line, when extended *backwards*, passes through this point.

2.5. Principal Points

The value of g which results in unit magnification (*i.e.* $M = 1$) is, based on (14), equal to

$$p_1 := \frac{\Delta - D}{B} \tag{19}$$

This defines the location (in terms of its z distance *to* the apex of the first surface) of the *first* principal point and of the corresponding, perpendicular to z , principal *plane*, as most of our distances, p_1 may turn out to be negative. The image of a (potentially virtual) object whose $g = p_1$ is then located in the *second* principal plane, which crosses the z axis at the second principal point and whose z

distance *from* the *last* surface (again, possibly negative) is, based on (13),

$$p_2 := -\frac{p_1 A + C}{p_1 B + D} = -\frac{A \frac{\Delta}{B} - \frac{\Delta}{B}}{\Delta} = \frac{1 - A}{B} \tag{20}$$

The significance of the two principal planes is this: they allow us to simplify the path of a ray entering a system of several surfaces by letting the incoming straight line reach the first principal plane, then continue by a *horizontal* line to the second principal plane, from where the new path follows the correct final direction, given by the second component of (11). The result is clearly *not* a true representation of the ray's *full* path, but it is a proper way of connecting its incoming and outgoing parts and thus sufficient for finding the image of an object, as done in the next section.

Proof. Assuming that the ray starts at $\langle x_0, 0, 0 \rangle$, all we have to do is to show that the new, simplified path, being *given* the correct outgoing direction, has also the correct x value, *i.e.* (14), at distance (13) from the last surface. This is confirmed by

$$\begin{aligned} & x_0 + (g - p_1)v_0 + (h_i - p_2)v_k \\ &= x_0 + \left(g - \frac{\Delta - D}{B}\right)v_0 + \left(-\frac{gA + C}{gB + D} - \frac{1 - A}{B}\right)(Bx_0 + v_0(Bg + D)) \\ &= \left(g - \frac{\Delta - D}{B} - gA - C - g(1 - A) - \frac{(1 - A)D}{B}\right)v_0 \\ & \quad + \left(1 - (1 - A) - \frac{(gA + C)B}{gB + D}\right)x_0 \\ &= \left(\frac{gBA + DA - gAB - CB}{gB + D}\right)x_0 = Mx_0 \end{aligned} \tag{21}$$

Note that it is possible to reverse the logic of the last step of this construction: with the help of principal planes, we can construct the outgoing ray when, instead of knowing its final direction, we know the exact location of the image; after reaching the second principal plane, we run a straight line through the image. ■

2.6. Main Formulas

We now derive a simple formula relating the following two focal lengths, namely

$$\begin{aligned} F_1 &:= f_1 - p_1 = -\frac{\Delta}{B} \\ F_2 &:= f_2 - p_2 = -\frac{1}{B} \end{aligned} \tag{22}$$

to distances of the object (image) to (from) the first (second) principal plane, namely

$$\begin{aligned} G &:= g - p_1 = g - \frac{\Delta - D}{B} \\ H &:= h_i - p_2 = -\frac{gA + C}{gB + D} + \frac{A}{B} - \frac{1}{B} = \frac{\Delta}{B(gB + D)} - \frac{1}{B} \end{aligned} \tag{23}$$

Having defined these, it is easy to show that

$$\frac{F_1}{G} + \frac{F_2}{H} = 1 \quad (24)$$

Proof.

$$\frac{\Delta}{\Delta - D - gB} + \frac{1}{1 - \frac{\Delta}{gB + D}} = \frac{\Delta - gB - D}{\Delta - D - gB} = 1 \quad (25)$$

■

when $\Delta = 1$ (the ambient medium is air), $F_1 = F_2$, with the corresponding simplification of (24). Let us also note that object-related distances such as G and F_1 have *opposite* orientation to H and F_2 , as seen in **Figure 2**.

The RHS of (14) then provides a quick and simple way of finding the image's x location, where magnification is now computed by

$$M = \frac{F_1}{F_1 - G} \quad (26)$$

Proof.

$$\frac{F_1}{F_1 - G} = \frac{\Delta}{gB - (\Delta - D) + \Delta} = \frac{\Delta}{gB + D} \quad (27)$$

■

Similarly, solving (24) for H yields the image's z location, namely

$$H = \frac{F_2}{1 - \frac{F_1}{G}} = \frac{GF_2}{G - F_1} \quad (28)$$

and provides an alternate formula for

$$M = \frac{F_2 - H}{F_2} \quad (29)$$

In addition to this algebraic way of finding the image, the simplified-path construction of the previous section provides an equivalent *geometrical* technique of achieving the same goal by:

- starting at the object, run a straight line parallel to z axis till reaching the *second* principal plane, then continue by a straight line, running it through the *second* focal point,
- run a straight line from the object through the *first* focal point till reaching the *first* principal plane, then continue parallel to z .

The image is located where these two lines intersect. A simple graph of **Figure 2** confirms that this is equivalent to (26) and (28).

This means that, in the current approximation, optical properties of any system of lenses are fully specified by the location of the two focal and the two principal points.

2.7. Nodal Points

There is one more pair of the so called cardinal points, defined in the following

manner.

For each object, it is possible to find a ray whose incoming and outgoing segments are parallel to each other; this happens when the outgoing direction, the second component of (11), is equal to v_0 , implying

$$v_0 = \frac{Bx_0}{1-D-gB} \stackrel{\text{def}}{=} v_* \tag{30}$$

The incoming ray then intersects the z axis at

$$-\frac{x_0}{v_*} = \frac{D-1}{B} + g \tag{31}$$

thus defining the first nodal point, whose distance *to* the first surface is then *minus* (31), namely

$$d_1 := \frac{1-D}{B} \tag{32}$$

Similarly, the emerging ray intersects the z axis at the second nodal point, whose distance *from* the last surface is

$$\begin{aligned} d_2 &:= h_i - \frac{Mx_0}{v_*} = -\frac{gA+C}{gB+D} - \frac{\Delta(1-gB-D)}{B(gB+D)} \\ &= -\frac{gAB+AD-\Delta(gB+D)}{B(gB+D)} = -\frac{A-\Delta}{B} \end{aligned} \tag{33}$$

The two nodal points enable us to add yet another path to our geometric construction of the image, namely: a straight line from the object to the first nodal point, horizontally connected to the second nodal point, and then continued in the original direction, as shown in **Figure 2**; note that the graph assumes that $\Delta = 1$, resulting in its simplification, as described shortly.

Defining

$$\begin{aligned} F'_1 &:= f_1 - d_1 = -\frac{1}{B} \\ F'_2 &:= f_2 - d_2 = -\frac{\Delta}{B} \end{aligned} \tag{34}$$

and

$$\begin{aligned} G' &:= g - d_1 = g - \frac{1-D}{B} \\ H &:= h_i - d_2 = -\frac{gA+C}{gB+D} + \frac{A}{B} - \frac{\Delta}{B} = \frac{\Delta}{B(gB+D)} - \frac{\Delta}{B} \end{aligned} \tag{35}$$

enables us to derive a twin equation to (24), namely

$$\frac{F'_1}{G'} + \frac{F'_2}{H'} = 1 \tag{36}$$

Proof.

$$\frac{1}{1-D-gB} + \frac{1}{1-\frac{1}{gB+D}} = \frac{1-gB-D}{1-D-gB} = 1 \tag{37}$$



when $\Delta = 1$, there is no difference between principal and nodal points, *i.e.* $d_1 = p_1$ and $d_2 = p_2$; furthermore $F_1 = F_2 = F'_1 = F'_2$, resulting in a welcomed simplification of all previous formulas. This is what is assumed in our next section.

2.8. Simple Lens

For a lens with two spherical surfaces R_1 and R_2 of apex-to-apex distance q , made of a medium (usually glass) with an index of refraction n , we get, based on (10)

$$\begin{aligned} \begin{bmatrix} A & C \\ B & D \end{bmatrix} &= \left(\begin{bmatrix} 1 & 0 \\ \frac{n-1}{R_2} & n \end{bmatrix} + q \begin{bmatrix} 0 & 1 \\ 0 & \frac{n-1}{R_2} \end{bmatrix} \right) \begin{bmatrix} 1 & 0 \\ \frac{1-n}{nR_1} & \frac{1}{n} \end{bmatrix} \\ &= \begin{bmatrix} 1 - \frac{(n-1)q}{nR_1} & \frac{q}{n} \\ (n-1)\frac{n(R_1 - R_2 - q) + q}{nR_1R_2} & 1 + \frac{(n-1)q}{nR_2} \end{bmatrix} \end{aligned} \tag{38}$$

Formula (22) then yields the lens' focal length

$$F = \frac{n}{n-1} \cdot \frac{R_1R_2}{n(R_2 - R_1 + q) - q} \tag{39}$$

while, from (32), we get the distance of the first nodal point *to* the first spherical surface

$$p_1 = \frac{R_1q}{n(R_2 - R_1 + q) - q} \tag{40}$$

and, from (33), the distance of the second nodal point *from* the second spherical surface

$$p_2 = \frac{-R_2q}{n(R_2 - R_1 + q) - q} \tag{41}$$

The distance from p_1 to p_2 is thus equal to

$$\Gamma := \frac{(R_1 - R_2)q}{n(R_2 - R_1 + q) - q} + q = \frac{q(n-1)(R_2 - R_1 + q)}{n(R_2 - R_1 + q) - q} \tag{42}$$

Formula (39) is often presented in the following form

$$\frac{1}{F} = (n-1) \left(\frac{1}{R_1} - \frac{1}{R_2} + \frac{(n-1)q}{nR_1R_2} \right) \tag{43}$$

For a so-called thin lens, q can be approximated by zero, yielding, for the focal length's *reciprocal*, the following simple expression

$$(n-1) \left(\frac{1}{R_1} - \frac{1}{R_2} \right) \tag{44}$$

while both p_1 and p_2 are then equal to zero.

2.9. Keystone Distortion

Another interesting consequence of our main formulas (even though of lesser practical importance) is that objects located in the same object plane (this time *not* necessarily perpendicular to z) form their images in the corresponding image plane (see the subsequent proof); when the planes *are* tilted with respect to z , the collection of such images builds a sharp but *distorted* rendition of the original scenery, since magnification varies with the image's z coordinate, this creates the so-called keystone effect (a square tilted on one of its sides becomes a trapezoid, the shape of a keystone; a different picture emerges when a square array of dots is tilted on one of its vertices, as **Figure 3** demonstrates).

Proof. Going back to the full, three-dimensional description, we consider objects at $\langle x, y, ax + by + c \rangle$ where a, b and c are fixed parameters, while x and y have variety of values (thus defining an arbitrary plane); note that the first refracting surface is now placed at $z = 0$. This implies that the corresponding $G_{x,y} = -ax - by - c - p_1$ and the resulting image is thus at

$$\mathbf{I} := \left\langle \frac{F_1 x}{F_1 - G_{x,y}}, \frac{F_1 y}{F_1 - G_{x,y}}, p_2 + \frac{G_{x,y} F_2}{G_{x,y} - F_1} \right\rangle \tag{45}$$

which is a parametric (x and y are now the variable “parameters”, just a semantic switch) representation of the image plane. The same plane is more conventionally defined by one of its points, namely

$$\mathbf{I}_0 := \left\langle 0, 0, p_2 + \frac{(c + p_1) F_2}{c + p_1 + F_1} \right\rangle \tag{46}$$

found by evaluating \mathbf{I} at $x = y = 0$, and the following *normal* (but not necessarily unit) vector

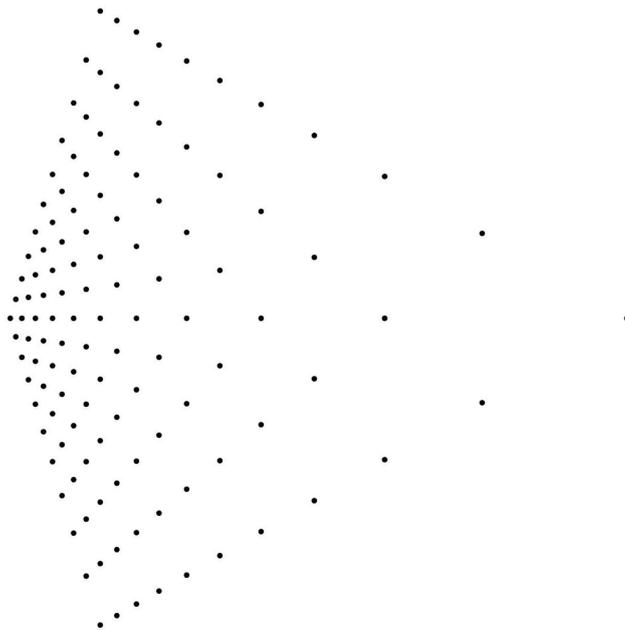


Figure 3. Keystone distortion.

$$\mathbf{N} := \langle aF_2, bF_2, -(c + p_1 + F_1) \rangle \tag{47}$$

This is verified by showing that the dot product of $\mathbf{I} - \mathbf{I}_0$ and \mathbf{N} equals to zero for any x and y ; to simplify this task, we first multiply the former vector by $F_1 - G_{x,y}$. The corresponding dot product is then

$$\begin{aligned} & axF_1F_2 + byF_1F_2 - (c + p_1 + F_1)G_{x,y}F_2 - (c + p_1)F_2(F_1 - G_{x,y}) \\ &= (ax + by)F_1F_2 + F_1G_{x,y}F_2 + (c + p_1)F_2F_1 \\ &= -(c + p_1)F_1F_2 + (c + p_1)F_2F_1 = 0 \end{aligned} \tag{48}$$

■

Similarly one can show that an image of a straight line is also a straight line.

Proof. Now, the object is at $\langle x, ax + c, bx + d \rangle$ where x is arbitrary, implying that $G_x = -bx - d - p_1$, and yielding the following parametric representation of the image line and one of its points

$$\mathbf{I} := \left\langle \frac{F_1x}{F_1 - G_x}, \frac{F_1(ax + c)}{F_1 - G_x}, y, p_2 + \frac{G_xF_2}{G_x - F_1} \right\rangle \tag{49}$$

$$\mathbf{I}_0 := \left\langle 0, \frac{F_1c}{d + p_1 + F_1}, p_2 + \frac{(d + p_1)F_2}{d + p_1 + F_1} \right\rangle \tag{50}$$

respectively. Since $\mathbf{I} - \mathbf{I}_0$ is proportional to the *constant* vector

$$\langle F_1 + d + p_1, (F_1 + d + p_1)a - bc, F_2b \rangle \tag{51}$$

for any value of x (just divide the former by F_1x and multiply by the common denominator), the line is straight. ■

The keystone distortion thus presents the original scenery as if seen from a different perspective.

3. Optical Instruments

We now turn our attention to more practical issues (see [4]), discussing basic principles behind common optical instruments, including some technical challenges which arise in this context. Since the ambition of this article does not go beyond the first-order approximation, we avoid the topic of optimizing the *quality* of resulting images by using *groups* of lenses (more on this in our Conclusion); we aim at a rudimental understanding of workings of these instruments only.

3.1. Loupe

When a lens is used as a magnifying glass (“loupe” is an alternate name, usually applied to the watchmaker’s version), the object is placed slightly to the right (in terms of our z axis) of the first focal plane (F must be positive), so as to create its *virtual* image at $H \approx -25$ cm (since 25 cm is the closest distance a normal eye can focus at). From (24) we find that, to achieve this, we need

$$G = \frac{F}{1 - \frac{F}{H}} = \frac{F}{1 + \frac{F}{25}} \tag{52}$$

where F is also in cm. The image thus appears on the object's side of the lens; it is then viewed (from the opposite side) with the observer's eye close to the lens, implying that the corresponding magnification (26) is

$$\frac{F}{F-G} = \frac{\left(1 + \frac{F}{25}\right)F}{\left(1 + \frac{F}{25} - 1\right)F} = \frac{25}{F} + 1 \quad (53)$$

The image is thus bigger (by the extra $\frac{25}{F}$ term) than the actual object as it would appear when viewed from the same 25 cm distance *without* a magnifying glass.

A few things to note:

- To achieve an actual, *i.e.* bigger than 1, magnification, F must be positive (as mentioned already).
- The image is upright (or erect) rather than inverted, since M is *positive*.
- When the object is placed directly *in* the first focal plane, (26) becomes infinite, but consequently not very meaningful; what needs to be compared now is the image's *angular* size (regular size divided by the image's distance, which is given by $\frac{Mx_0}{-H}$ and tends to $\frac{x_0}{F}$ as H approaches infinity) with the object's angular size when viewed from a distance of 25 cm (*i.e.* $\frac{x_0}{25}$), resulting in $\frac{25}{F}$. The advantage of this arrangement is that the observer's eyes do not need to strain; the image appears at infinity, the most comfortable focus for a normal eye, but the resulting magnification is then smaller than (53).

In this context we should mention that the eye itself is a rather amazing optical instrument, consisting of a lens of adjustable focal length and aperture; with the help of the brain, it is capable of building a *three-dimensional* (due to having two eyes, which see each object from slightly different angles) and sharp (due to its ability to change its focus and direction so quickly that it can scan the whole scenery in an instant) representation of the outside world.

Some eyes have difficulty focusing over the usual 25 cm to infinity range; this can be corrected by placing a pair of simple lenses with the appropriate focal length in front of them. We have thus described the most common optical instrument: the eyeglasses.

3.2. Object (Image) at Infinity

When g is orders of magnitude larger than the size of the optical system itself (e.g. a camera in a landscape mode, an eye watching a distant object, etc.), rather than converting the existing formulas to their $g \rightarrow \infty$ limit (which would be rather difficult), it is more expedient to

- have a pencil of rays starting at the object reach the first surface as *parallel* lines,

- identify the object's location by the value of v_0 (instead of x_0 , which would be infinite as well),
- use $g = 0$; for each ray of the pencil, x_0 is then the x coordinate of where the ray enters the first surface.

The image is found at the point of convergence of the outgoing rays; this, as we already know, happens at the *second* focal plane. The corresponding x value (the image's size) is, based on (11) with $g = 0$ and $h = f_2$, equal to

$$\left(C - \frac{A}{B}D\right)v_0 = -\frac{\Delta}{B}v_0 = F_1v_0 \quad (54)$$

Since the image needs to be *real* (as opposed to *virtual*), F_1 needs to be positive. As v_0 is the incoming angle, rays from objects *above* the z axis arrive with *negative* values of v_0 (and vice versa); the camera thus builds an *inverted* picture of the scenery.

Conversely, when placing an object of vertical size x_0 in the *first* focal plane, we know that the outgoing rays of the corresponding pencil are parallel to each other, leaving the system at an angle given by (11) with $g = f_1$, namely

$$Bx_0 = -\frac{x_0}{F_2} \quad (55)$$

This describes, to a good approximation, objects (originally captured on a planar film) being projected onto a large *distant* screen. The image is again inverted; this is easily fixed by inverting the film instead. F_2 must be positive since the film needs to be located *before* the first surface and be properly illuminated. To get the image's size, it is sufficient to multiply (55) by the screen's distance.

The illumination is facilitated by a small source of intense light directed at the film. This necessitates placing an extra collector lens between the source and the film, to make the light converge at the projecting lens; also: having a spherical mirror behind the source to reflect its rays in the film's direction.

3.3. Microscope and Telescope

The same approach can be also used to explain workings of a microscope: a tiny object of size x_0 is similarly "projected" to create a real image at a relatively large distance t (called tube length); the resulting magnification is then given by the size of this image (see the previous section) divided by x_0 , *i.e.*

$$-\frac{x_0}{F_2} \times t \div x_0 = -\frac{t}{F_2} \quad (56)$$

The lens, called an objective, must therefore have a small focal length to achieve high magnification. The image thus created is then observed through (becomes the object of) another, ocular lens, also called an *eyepiece*, which functions in a manner of a loupe; typically, an ocular is a system of two or more lenses, to improve its optical properties. The overall magnification of the microscope is then the product of (56) and (53).

Finally, a telescope (a pair of these makes so called *binoculars*) is an arrange-

ment of two lenses such that the *first* focal point of the *second* (ocular) lens coincides with the *second* focal point of the *first* (objective) lens. The object is effectively at infinity, which means that its location must be specified by the angle, say v_0 , at which parallel rays from the object enter the first lens. Based on (54), this creates an image of size $\dot{F}_1 v_0$ (a single dot implying the *first* lens) in the objective's second focal plane. The ocular (again, essentially a loupe) then converts this image into a pencil of parallel rays leaving the telescope at an angle given by (55), namely

$$-\frac{\dot{F}_1}{\ddot{F}_2} v_0 \tag{57}$$

(two dots implying the *second* lens). The coefficient of v_0 is the corresponding angular magnification (to make it large, \dot{F}_1 must be close to the tube length, whereas \ddot{F}_2 is kept relatively small, several ocular lenses may be interchangeably used with the same telescope, to allow for different overall magnifications).

When both focal lengths are positive, the image is inverted, when \dot{F}_1 is positive (as it must be, to create a real image) and \ddot{F}_2 is negative, the image is upright; the latter design is used only for relatively small magnifications (opera glasses, etc.). Using the former design for land-based binoculars (where an erect image is essential inversion makes no difference to astronomers) necessitates inserting, between the objective and the ocular, either four reflecting mirrors (this is what makes the usual binoculars so bulky), or an extra erector lens. This lens takes the image created by the objective, making it its object, and creates a new image in the ocular's first focal plane, using 1 magnification (achieved by $G = 2F$ and $H = 2F$, where F is the *erector's* focal length). This increases the length of the instrument by $4F + \Gamma$, where Γ was defined in (42); such a design is used mainly in rifles' telescopic sights, which need to be of a small diameter at the expense of a longer tube.

We mention in passing that wave properties of light put a limit on magnification achievable in this manner.

3.4. Combining Two Systems

Assuming that the ambient medium is air and, consequently, $\Delta = 1$, we now explore what happens when two such optical systems (e.g. two surfaces of a simple lens; a tandem of two lenses; objective and ocular of an instrument; etc.) are combined by arranging them, one after the other, along the z axis. Assuming that the principal points and the focal length of each of the two systems are known (we place single and double dot over quantities of the first and second system, respectively) and taking the distance from \dot{p}_2 to \ddot{p}_1 to be s (potentially negative), we want to find p_1, p_2 and F of the new, *combined* system, as these fully determine its overall optical properties; the two individual sub-systems no longer matter. We claim that

$$p_1 = \dot{p}_1 - \frac{\dot{F}s}{\dot{F} + \ddot{F} - s} \tag{58}$$

$$p_2 = \ddot{p}_2 + \frac{\ddot{F}s}{\dot{F} + \ddot{F} - s} \tag{59}$$

$$F = \frac{\dot{F}\ddot{F}}{\dot{F} + \ddot{F} - s} \tag{60}$$

where p_1 (p_2) is the distance of the first (second) principal point of the combined system to (from) the first (last) surface of the first (second) sub-system.

Proof. Placing on object at p_1 which, from the perspective of the first sub-system, see (23), corresponds to

$$\dot{G} = p_1 - \dot{p}_1 = -\frac{\dot{F}s}{\dot{F} + \ddot{F} - s} \tag{61}$$

we get, based on (28), for the location of the resulting image (specifically, its distance from \dot{p}_2)

$$\dot{H} = \frac{\dot{G}\dot{F}}{\dot{G} - \dot{F}} = \frac{-\dot{F}^2s}{-\dot{F}s - \dot{F}^2 - \dot{F}\ddot{F} + \dot{F}s} = \frac{\dot{F}s}{\dot{F} + \ddot{F}} \tag{62}$$

implying that

$$\ddot{G} = s - \dot{H} = \frac{\ddot{F}s}{\dot{F} + \ddot{F}} \tag{63}$$

The combined system's magnification is clearly just a product of the two individual magnifications, namely

$$\frac{\dot{F}}{\dot{F} - \dot{G}} \cdot \frac{\ddot{F}}{\ddot{F} - \ddot{G}} = \frac{\dot{F}^2 + \dot{F}\ddot{F} - \dot{F}s}{\dot{F}^2 + \dot{F}\ddot{F}} \cdot \frac{\dot{F}\ddot{F} + \dot{F}^2}{\dot{F}\ddot{F} + \dot{F}^2 - \dot{F}s} = 1 \tag{64}$$

resulting in overall *unit* magnification. This proves that p_1 is the correct first principal point of the combined system.

From the perspective of the second sub-system, an object at \ddot{G} , given by (63), has its image at the following distance from \ddot{p}_2

$$\ddot{H} = \frac{\ddot{G}\ddot{F}}{\ddot{G} - \ddot{F}} = \frac{\ddot{F}^2s}{\ddot{F}s - \ddot{F}\ddot{F} - \ddot{F}^2} = \frac{\ddot{F}s}{s - \ddot{F} - \ddot{F}} \tag{65}$$

which verifies (59).

Finally, starting with an object at $\dot{G} = \infty$ yields $\dot{H} = \dot{F}$, $\ddot{G} = s - \dot{F}$ and

$$\ddot{H} = \frac{\ddot{G}\ddot{F}}{\ddot{G} - \ddot{F}} = \frac{\ddot{F}s - \dot{F}\ddot{F}}{s - \dot{F} - \ddot{F}} \tag{66}$$

Converting \ddot{H} to the distance from p_2 (instead of \ddot{p}_2) yields

$$F = \frac{\ddot{F}s - \dot{F}\ddot{F}}{s - \dot{F} - \ddot{F}} + \ddot{p}_2 - p_2 = \frac{\ddot{F}s - \dot{F}\ddot{F}}{s - \dot{F} - \ddot{F}} + \frac{\ddot{F}s}{\dot{F} + \ddot{F} - s} = \frac{\dot{F}\ddot{F}}{\dot{F} + \ddot{F} - s} \tag{67}$$

■

In the same manner, it is possible to combine three or more such optical sub-systems into a single one, whose principal points and focal length (and again: to the current level of approximation, nothing else is needed to know the combined system's optical properties) are then established by a repeated application of (58) to (60).

Formula (60) explains workings of the so-called zoom lens: by changing the value of s , we can give the combined system an almost arbitrary F , *i.e.* the corresponding angle-to-size magnification given by (54). Note that the smallest possible value of s is $-(\dot{p}_2 + \dot{p}_1)$ when the *physical* distance between the two original sub-systems becomes zero, they come into direct contact with each other.

We need to mention that the actual design of a camera's objective (think Nikon, Zeiss, Canon, etc.) involves many additional considerations and requires more sophisticated mathematical formulation to achieve a sharp and undistorted image for a continuous variety of magnifications.

3.5. Aperture and Vignetting

Our discussion would not be complete without considering the following issue: in addition to the geometric properties discussed up to this point, it is also important for an optical system to let through the maximum amount of light, to make the corresponding image as bright as possible. To achieve this goal, we must first have some way of establishing *which* rays of a pencil starting at an object actually make it to their target, *i.e.* to the corresponding image.

To determine that, we realize that all spherical surfaces of an optical system have a finite *radial* extent; we can then visualize them as a slightly curved *disks* of different diameters. Note that the curvature is now irrelevant; it is their circular contour which matters.

By tracing individual rays of the original pencil, we discover that not all of them can reach the image, as some simply miss one of these disks and get absorbed by the instrument's enclosure. To find exactly which rays do succeed, we create images of all these disks as seen from the *object* side, *i.e.* by reversing the rays' direction (the first disk is an image of itself, the second one has its image created by the first surface, etc.); the smallest one of the resulting disk *images* (called the entrance pupil) corresponds to the so-called aperture, *i.e.* the *actual*, physical surface behind this image. A ray makes it through the whole system if and only if it can pass through the aperture; equivalently, an *incoming* ray (extended, if necessary) must pass through the entrance pupil, and a (potentially extended) *outgoing* ray must pass through the corresponding exit pupil (the aperture's image when viewed from the *image* side of the system). Looking through binoculars *from a distance* against a white background, the disk of light we see is the exit pupil; by slightly moving the binoculars we may notice that this pupil is located *in front* of the ocular; the corresponding distance is called eye relief. Ideally, the location of the exit pupil should be where the eye is normally placed, and its diameter should match the eye's pupil (about 7 mm). Similarly, looking through binoculars from a distance, but this time from the *wrong* side, the white disk we see is the *entrance* pupil, *i.e.* another image of the same aperture.

Since the entrance *images* of the previous paragraph are all disks centered on the z axis, they appear *concentric* to an object near the z axis; we already know

that only by passing through the smallest one of these (the entrance pupil) will allow a ray pass through the whole system. When the object is *off* the z axis, the disks' centers are no longer concentric but appear aligned along a straight line; they move further apart as the object's distance from the z axis increases. For a ray to pass through the whole system now requires it to pass through *all* these disks, *i.e.* through what appears (from the object's point of view) their *overlap*. The area of this overlap decreases as the object moves further away from the z axis (this effect is called vignetting), until the *edge* of one of the bigger disks (an image of a physical surface called the field stop) reaches the *center* of the entrance pupil; the angular location of such an object defines the system's field of view. Move the object a bit further and it becomes practically invisible to the observer.

An optical system (such as a camera) may have a deliberate *additional* light barrier (called a diaphragm) with adjustable circular opening to function as the system's aperture; this allows us to control the film's exposure. Similarly, the field of view is usually restricted by the physical extent of the film itself, which then defines the system's field stop.

3.6. Periscope

When one part of a system (usually, the objective) creates an intermediate image which is then taken to be the object of the systems's next part (the ocular), as done in telescopes, microscopes, etc., it is possible to improve the system's light-transmitting properties by inserting, at the intermediate image/object location, an extra (so-called field) lens with positive F . Its function is to move the image/object from its first to its second *principal* plane, leaving it otherwise intact (*i.e.* using $M=1$), while bending the passing rays so as to prevent them from hitting the instrument's tube.

This idea is essential for building systems which are limited in diameter but whose tube needs to be fairly long (such as periscopes). It should be intuitively obvious that even after precisely aligning the objective and ocular parts of such a long and narrow system (by itself a daunting task, especially when the tube may need to bend, as in endoscopes), the amount of light reaching the ocular would be extremely small (if any). To fix this problem, we need to go back to the idea of an *erector* lens (which takes an image and makes its inverted carbon copy at the distance $4F + \Gamma$), and insert as many of these as needed to link the objective with the ocular; this alleviates the problem of alignment. At the same time, half-way between these relay (as they are now known) lenses, it is necessary to insert the same number of *field* lenses described in the previous paragraph, to ensure that enough light gets through; each of these lenses adds only Γ to the instrument's length.

In endoscopes, it is necessary that the relay/field lenses have even smaller axial diameter, but are rather long (surface to surface); they are then called rod lenses. Arranging them next to each other (in spite of their different functions, they can be otherwise identical) then results in leaving only small air gaps between them,

thus allowing the tube to be both stiff (in terms of its length) and flexible (in terms of its bending). To ensure proper illumination, sufficient light needs to be transmitted through an endoscope, starting from its image end (where the eye is); after brightening the object, the reflected light then travels back to the observer.

4. Conclusions

In this article we have described only the most basic properties of a system of lenses; the attentive reader may have realized that, if everything was as perfect as presented here, there would be no need for using more than one lens for any given purpose (only its cardinal points and focal length would matter). This is definitely *not* the case in real world, and it is now time to admit that the actual situation is somehow more complicated. This is due to so-called aberrations (lens properties causing image distortion and smearing, problems with proper focusing, etc.), which arise when third and higher-order terms are included in the expansion of the sin function. The second important source of these is the fact that most optical materials (glass in particular) have an index of refraction which slightly varies with the light's colour.

To minimize these imperfections (in order of their seriousness and this would vary from one type of optical instrument to another), instead of single lens, one must use a complex system of several lenses with different indices of refraction, designed in a way to make their individual aberrations *cancel* each other. This is most important in the design of camera objectives, which must achieve such cancellation for objects at many different distances (not to mention a further complication of zooming, *i.e.* changing the resulting magnification). No wonder that these often require a collection of twelve or more lenses, not all of them with spherical surfaces (this provides the designer with extra flexibility). Such aberrations will be discussed in proper detail in our subsequent submission.

Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

References

- [1] Herzberger, M. (1980) Modern Geometrical Optics. Krieger Pub. Co., Malabar.
- [2] Halliday, D., Resnick, R. and Walker, J. (2013) Chapter 34, Fundamentals of Physics. Wiley, Hoboken.
- [3] Lin, P.D. (2016) Advanced Geometrical Optics. Springer, Berlin.
<https://doi.org/10.1007/978-981-10-2299-9>
- [4] Smith, W.J. (2008) Modern Optical Engineering. 4th Edition, McGraw-Hill, location 2 Pennsylvania Plaza New York City.