Matrix Representation of Gaussian Optics*

KLAUS HALBACH
Lawrence Radiation Laboratory, University of California, Berkeley, California
(Received 5 September 1963)

The ray-tracing method is used to discuss Gaussian optics. After proving that Gaussian optics can be described by matrices, some often-used general formulas for telescopic and focusing systems are derived. This formalism is then used to solve several problems. They are selected to make the reader familiar with the application of the matrix representation of Gaussian optics and to acquaint him with some optical systems that are very useful but not well known among physicists who use optical methods only occasionally.

I. INTRODUCTION

ALTHOUGH matrices are often used to describe electron-optical systems,[1] and although there exists an excellent modern book on geometrical (light) optics[2] which uses matrices, the matrix technique of dealing with optical problems is by no means common knowledge among physicists, nor is it, as a rule, taught at colleges and universities. One probable reason is that, to the knowledge of this author, there is no paper or book that addresses itself not to the professional optical systems designer, but to the physicist who uses optical methods only occasionally in the laboratory. This article attempts to fill that gap.

Since the laboratory physicist usually has to assemble his optical system with stock items, his possibilities for correcting any but chromatic aberrations are limited. We, therefore, deal primarily with first-order (Gaussian) optics and mention aberrations only to see what the limitations of Gaussian optics are and what has to be avoided in the design of an optical system. We try to demonstrate that the description of Gaussian optics with matrices makes both the analysis and synthesis of optical systems so simple that they can be done in a systematic way even by a person inexperienced in this field. Also, we try to show that the matrix representation of Gaussian optics is well suited to be taught at our colleges and universities, since it would acquaint the student with Gaussian optics and would make him familiar with matrices in connection with a very simple subject.

We intentionally limit ourselves to light optics, since the application of this formalism to electron- and ion-optical problems is trivial to the physicists concerned. The only concession we make is that in our general formulas we do not assume the object and image space to have the same index of refraction, thus making these expressions directly applicable to electron-optical problems.

II. RAY-TRACING FORMULAE AND INTRODUCTION OF MATRICES

To derive the ray-tracing formulae, we require, at least at first, that the optical system consist of lenses with rotational symmetry with all axes coinciding, thus forming the common optical axis. Therefore, we momentarily exclude cylinder lenses. We furthermore restrict the discussion in the beginning to meridional rays, i.e., to rays that lie in a common plane with the optical axis of the system.

To describe a ray at a reference plane (RP), which is always perpendicular to the optical axis, we introduce the two quantities r and r' (Fig. 1). The distance between the optical axis and the intersection between the ray and the RP is given by r, whereas r' is an abbreviation for dr/du, describing the resulting change dr of r when the RP is displaced by du. The object of tracing a ray is to establish the relation between r, r' of some initial RP and r, r' of any other RP of interest. In

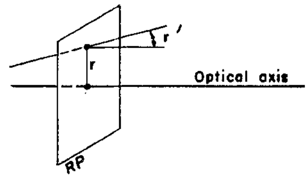

FIG. 1. Definition of r and r'.

* Work done under the auspices of the U. S. Atomic Energy Commission.

[1] M. S. Livingston and J. P. Blewett, *Particle Accelerators* (McGraw-Hill Book Company, Inc., New York, 1962).

[2] M. Herzberger, *Modern Geometrical Optics* (Interscience Publishers, Inc., New York, 1958).

deriving these relations, we assume both r and r' to be small enough so that only the lowest significant powers of r and r' have to be taken into account. This leads to equations that are linear in r and r', giving first-order (Gaussian) optics. Since we almost always use this approximation, for simplicity we usually refer to r' as the angle between the ray and the optical axis (although r' is, as introduced above, really the tangent of that angle).

Since a ray goes either through a homogeneous medium or is refracted at the interface between two media with different indices of refraction, we have to derive, at least in principle, only the two basic relations that we discuss in Secs. IIA and IIB.

A. Transit Through a Homogeneous Medium

In Fig. 2, and all other figures of this paper, the location of the RP's are marked by numbered points on the optical axis. If RP_1 and RP_2 are separated by the distance D_{12}, we directly obtain from Fig. 2

$$r_2 = r_1 + D_{12} \cdot r_1' \quad \text{and} \quad r_2' = r_1'. \tag{1a}$$

B. Refraction at Interface Between Different Media

In our approximation we can represent all interfaces by spherical surfaces. In accordance with Fig. 3, we obtain for $r_1' \ll 1$, $r_2' \ll 1$, $r_1 \ll R$ from Snell's law,

$$n_1 \cdot (r_1' + r_1/R) = n_2 \cdot (r_2' + r_1/R).$$

After rearrangement, and with $1/R = K$, we obtain for the relations between r_1, r_1' and r_2, r_2':

$$\begin{aligned} r_2 &= r_1 \\ r_2' &= K[(n_1 - n_2)/n_2]r_1 + (n_1/n_2)r_1'. \end{aligned} \tag{2a}$$

Since we carry only first-order terms, we can assume that the refraction takes place at the plane that is tangential to the interface at the intersection between the axis and the interface (vertex), thus establishing the same set of RP's for all refracted rays; this does not change Eq. (2a).

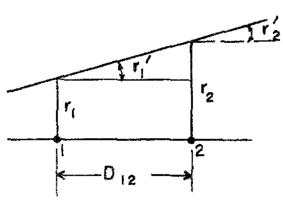

FIG. 2. Ray transit through a homogeneous medium.

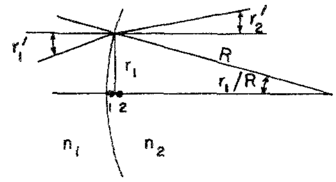

FIG. 3. Refraction of a ray at an interface between different media.

C. Consequences and Applications of Eqs. (1a) and (2a)

In our approximation, the relations between r_1, r_1' and r_2, r_2' are linear for both the transit through a homogeneous medium and for refraction at any interface. Since tracing through a whole optical system consists of a sequence of these steps, the relation between the r, r' of any two RP's must be a linear one, too. It is, therefore, appropriate and practical to describe these relations by matrices. If we introduce the abbreviation

$$\mathbf{r} = \begin{pmatrix} r \\ r' \end{pmatrix} \tag{3}$$

for the column vector

$$\begin{pmatrix} r \\ r' \end{pmatrix}$$

we can rewrite Eqs. (1a) and (2a) as

$$\mathbf{r}_2 = \begin{pmatrix} 1 & D_{12} \\ 0 & 1 \end{pmatrix} \mathbf{r}_1 \quad \text{(homogeneous medium)}, \tag{1b}$$

and

$$\mathbf{r}_2 = \begin{pmatrix} 1 & 0 \\ K[(n_1 - n_2)/n_2] & (n_1/n_2) \end{pmatrix} \mathbf{r}_1 \quad \text{(refraction)}. \tag{2b}$$

We notice that the determinant of the matrix in Eq. (2b) (which describes the transition from a medium with a refractive index n_1 to a medium with refractive index n_2) is n_1/n_2, whereas the determinant of the matrix in Eq. (1b) (no

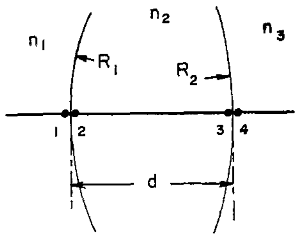

FIG. 4. Thick lens.

change of n) is one. Although we use a considerably simpler method later, the matrix that describes the relation between any two RP's can be obtained by multiplication of matrices of the type appearing in Eqs. (1b) and (2b), as we see in the following example. Since the determinant of the product of matrices equals the product of the determinants of the matrices, the determinant of the matrix describing the relations of the $\mathbf{r}$ in the two RP's must be equal to the ratio of the indices of refraction surrounding these RP's. We, therefore, have the general result for any two RP's

$$\mathbf{r}_2 = A\cdot\mathbf{r}_1; \quad A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}; \quad \|A\| = n_1/n_2. \tag{4}$$

To demonstrate how A could be obtained in principle, we calculate the matrix connecting the two vertex planes of a thick biconvex lens. According to Fig. 4 we have

$$\begin{aligned} A_{12}\mathbf{r}_1 &= \mathbf{r}_2; \\ A_{23}\mathbf{r}_2 &= \mathbf{r}_3; \\ A_{34}\mathbf{r}_3 &= \mathbf{r}_4. \end{aligned} \tag{5a}$$

Here, A_{23} is equal to the matrix in Eq. (1b), with D_{12} being replaced by d; the matrices A_{12} and A_{34} are equal to the matrix in Eq. (2b) with, respectively, K being replaced by $K_1 = 1/R_1$ for A_{12}, and K, n_1, and n_2 being replaced by $-K_2 = -1/R_2$, n_2 and n_3 for A_{34}. (For A_{34}, K has to be replaced by $-K_2$ because of opposite curvature.) Using Eqs. (5a), (1b), and (2b), we obtain

$$\mathbf{r}_4 = A_{34}\mathbf{r}_3 = A_{34}A_{23}\mathbf{r}_2 = A_{34}A_{23}A_{12}\mathbf{r}_1 = A_{14}\mathbf{r}_1,$$

$$\begin{aligned} A_{14} &= A_{34}\begin{pmatrix} 1 & d \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ K_1[(n_1-n_2)/n_2] & (n_1/n_2) \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ K_2[(n_3-n_2)/n_3] & (n_2/n_1) \end{pmatrix}\begin{pmatrix} 1+dK_1[(n_1-n_2)/n_2] & d(n_1/n_2) \\ K_1[(n_1-n_2)/n_2] & (n_1/n_2) \end{pmatrix}, \end{aligned}$$

and

$$\mathbf{r}_4 = \begin{bmatrix} 1+dK_1[(n_1-n_2)/n_2] & d(n_1/n_2) \\ K_1[(n_1-n_2)/n_3]+K_2[(n_3-n_2)/n_3] & \\ \quad +dK_1K_2[(n_1-n_2)(n_3-n_2)/n_2n_3] & (n_1/n_3)\{1+dK_2[(n_3-n_2)/n_2]\} \end{bmatrix}\mathbf{r}_1. \tag{5b}$$

Although Eq. (5b) contains all the necessary information for ray-tracing purposes, the individual matrix elements do not yet give us direct information about the focal length, position of the focal points, etc., of the thick lens. Before we discuss this subject in a general way (Sec. IV), we make some generalizations in the next section. With the exception of the discussion of skew rays, Sec. III can be omitted at the first reading without impairing one's understanding of the rest of this paper.

III. GENERALIZATIONS

A. Skew Rays

To discuss skew rays (i.e., rays that are not in a plane with the optical axis), we use Cartesian coordinates x–y in all RP's, with all the x axes parallel to each other (and therefore the y axes are similarly parallel). If we generalize the column vector $\mathbf{r}$ introduced in Eq. (3) to have the components x, x', y, and y' in this order, we want again to establish the relation between the $\mathbf{r}$'s in different RP's. As in Sec. II, by dropping all terms of higher than first order, which particularly excludes products between any of the quantities x, x', y, and y', the relation between the $\mathbf{r}$'s in different RP's can again be represented by a matrix, which is this time a 4×4 matrix. By representing this 4×4 matrix with the help of four 2×2 submatrices $b_{\nu\mu}$ in the following way,

$$B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}, \tag{6}$$

it is easy to see that $b_{12}=b_{21}\equiv 0$. Since a ray starting in a meridional plane always remains in that same meridional plane, the special meridional ray defined by x and $x'\neq 0$ and $y=y'=0$ in one RP must also satisfy $y=y'=0$ in any other RP. This is evidently only possible if $b_{21}\equiv 0$; by analogy one obtains $b_{12}\equiv 0$. Because of the rotational symmetry around the optical axis, the matrices describing meridional rays must be identical for all meridional planes, which necessitates $b_{11}\equiv b_{22}$. This means that the 2×2 matrices introduced in Sec. II are sufficient to describe skew rays too. Therefore, the r, r' introduced in Sec. II can be interpreted as the values describing the projection of a skew ray onto any meridional plane of interest, and a change in the notation introduced in Sec. II is not necessary.

B. Cylinder Lenses

If an optical system contains cylinder lenses also, we redefine the optical axis as the axis that perpendicularly intersects all interfaces between different media, and we require that such an optical axis exist. Using the same notation as in Sec. IIIA, and arguing the same way, one again arrives at the conclusion that the relation between the $\mathbf{r}$ in different RP's can be expressed by a 4×4 matrix, which we split up into four 2×2 matrices as in Sec. IIIA. To find out under what circumstances $b_{12}=b_{21}=0$ [see Eq. (6)], let us imagine that we calculate the matrix describing the refraction at the interface between two different media. It can be shown that for any second-order interface (which are the only ones of interest in our approximation), there exists one Cartesian coordinate system x, y such that if $y=y'=0$ before refraction, $y=y'=0$ holds also after refraction, and the same is true for x and x'. By using these coordinates, the refraction at the interface is described by a 4×4 matrix with $b_{12}=b_{21}=0$ and, since we are dealing with cylinder lenses, $b_{11}\neq b_{22}$. Furthermore, since the transition through a homogeneous medium is described by a matrix with $b_{12}=b_{21}=0$ and $b_{11}=b_{22}=\begin{pmatrix}1 & D\\ 0 & 1\end{pmatrix}$, b_{12} and b_{21} vanish for a system containing cylinder lenses when all the coordinate systems mentioned above have the same orientation. Consequently, in this case it is possible to describe the relation between x, x' at two different RP's by a 2×2 matrix (b_{11}), and to independently describe the relation of the y, y''s by another different 2×2 matrix (b_{22}). This means that the methods outlined in this paper can be directly applied to systems containing cylinder lenses, provided their principal axes are aligned as specified above. If this is not the case, b_{12} and b_{21} in general do not vanish and one has to use 4×4 matrices. Because of its relative unimportance, we do not discuss this case further.

C. Limitations of Gaussian Optics[3]

The obvious procedure for treating aberrations (i.e., deviations from Gaussian optics) is to carry not only first-order terms but also higher powers of r and r' in the calculation of the relation between r, r' in different RP's. Since a reversal of the signs of r, r' in one RP must necessarily lead to a reversal of the signs of r, r' in all other RP's, the expressions describing the relation between two RP's can depend only on odd powers of r, r'. Dropping higher than third-order terms, r in some RP would then depend on r_1, r_1' in some other RP in the following way

$$r=a_{11}r_1+a_{12}r_1'+a_{13}r_1^3+a_{14}r_1^2r_1'+a_{15}r_1r_1'^2+a_{16}r_1'^3, \quad (7)$$

and the expression for r' would have the same structure. The expression describing the transition through a homogeneous medium [Eq. (1a)] would remain unchanged, and corrective terms would appear only in the formulae describing the refraction at interfaces.

If Eq. (7) describes such a refraction, a_{13}, for example, has to be proportional to the inverse second power of the radius of curvature R of that interface, since $a_{13}r_1^3$ has to have the dimension of a length and R is the only length besides r_1 entering that problem. Requiring that the third-order terms be negligible compared to the first-order terms, one has, therefore, to fulfill, at least order of magnitude-wise, the conditions

$$(r_1/R)^2\ll 1 \quad \text{and} \quad r_1'^2\ll 1.$$

We see that these conditions are weaker than the ones originally set forth in Sec. IIB. Because we

[3] For simplicity, we talk here only about meridional rays; all the arguments hold equally for skew rays, where we would use the four variables x, x', y, and y'.

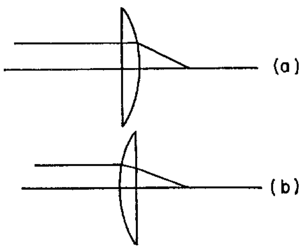

FIG. 5. Image formation with planoconvex lens.

find later that for thin lenses $1/f=(n-1)\times(1/R_1+1/R_2)$, we can replace $(r_1/R)^2 \ll 1$ by $(r_1/f)^2 \ll 1$ for a thin lens, again describing this condition only to an order of magnitude. Simply stated, this means that since the laboratory physicist usually does not have the means to obtain lens systems that are corrected for his particular application, he should try to use his lenses so that $r'^2 \ll 1$ and $(r_1/f)^2 \ll 1$, the latter condition requiring that the square of the effective f-number of each lens be large compared to one.

The correct use of a given lens can help to minimize aberrations: If a given plano-convex lens has to be used to form an image of an object that is far away (Fig. 5), the use of the lens as shown in Fig. 5(b) leads to smaller aberrations than if used as in Fig. 5(a). Although in Fig. 5(b) the ray is refracted twice, the aberrations are smaller than in Fig. 5(a), since at each interface in Fig. 5(b) the aberrations are of the order of $\frac{1}{4}$ of the aberration introduced at the second refracting interface in Fig. 5(a) due to the smaller change of r' at each refracting interface.

Because chromatic aberrations are caused by the wavelength dependence of the refractive index of materials, their effects cannot always be minimized by reducing the effective aperture of lenses, and do not belong in the category of limitations of Gaussian optics. With the exception of one application (Sec. VH), we do not discuss chromatic aberrations because a great variety of achromats are available at reasonable prices.

IV. DISCUSSION OF EQ. (4)

A. Significance of the Disappearance of Matrix Elements

Referring to Eq. (4), we notice first that because of $\|A\|=n_1/n_2>0$, at most only two of the matrix elements of A can be zero. If two elements actually are zero, they must be either the diagonal elements or the off-diagonal elements. To get a better understanding of the meaning the different matrix elements can have (besides their obvious role for ray tracing), we let the matrix elements of A individually vanish:

(a) We set $a_{12}=0$. The equation for r_2 then reads $r_2=a_{11}\cdot r_1$. This means that the two RP's have an object–image relation, with the lateral magnification $m=r_2/r_1=a_{11}$.

(b) We set $a_{21}=0$. The equation for r_2' then reads $r_2'=a_{22}\cdot r_1'$. This means that a parallel beam of light entering the optical system also leaves the system as a parallel beam. Since this is the description of a telescope focused at infinity, we call systems with $a_{21}=0$ telescopic systems and introduce the angular magnification or power of the telescope, $p=r_2'/r_1'=a_{22}$. A further discussion of telescopic systems follows in Sec. IVB. In Sec. IVC the definition of focusing systems is given.

(c) We set $a_{11}=0$. This means that for $r_1'=0$, then $r_2=0$ independently of the value of r_1; i.e., a beam of light parallel to the axis focuses in the second RP to $r_2=0$. Since this is the definition of a focal point, $a_{11}=0$ indicates that the second RP is a focal plane (FP).

(d) We set $a_{22}=0$. A consideration equivalent to the one above leads to the conclusion that for $a_{22}=0$, the first RP is a FP. A further discussion of focusing systems follows in Sec. IVC.

B. Telescopic Systems

As an introductory remark to the discussion of telescopic systems it might be worthwhile to mention that we discuss them not only because of the importance of the telescope as an instrument for observational purposes, such as in astronomy, but particularly because they can be used advantageously to set up optical systems in the laboratory, as we see in Sec. VC. If we limit ourselves to telescopic systems with $n_1=n_2$, which is fulfilled for an air–air system, and if we introduce the power p of the telescope as in Sec. IVA(b), the matrix describing the relation of the $\mathbf{r}$ in the two RP's becomes

$$A_{12}=\begin{pmatrix} 1/p & a_{12} \\ 0 & p \end{pmatrix}, \tag{8}$$

where the value of the upper left element results from the condition $\|A_{12}\|=1$.

To find out whether there are pairs of RP's that have an object–image relation,[4] we introduce RP_0 at the distance D_1 to the left of RP_1, and RP_3 at the distance D_2 to the right of RP_2 (Fig. 6). Analogously to the derivation of Eq. (5b), by connecting $\mathbf{r}_0$ to $\mathbf{r}_3$ through $\mathbf{r}_3=A_{03}\mathbf{r}_0$ we obtain with Eqs. (1b) and (8) for the matrix A_{03}:

$$A_{03}=\begin{pmatrix}1 & D_2\\ 0 & 1\end{pmatrix}\begin{pmatrix}1/p & a_{12}\\ 0 & p\end{pmatrix}\begin{pmatrix}1 & D_1\\ 0 & 1\end{pmatrix}$$

and, therefore,

$$A_{03}=\begin{pmatrix}1/p & a_{12}+D_2p+D_1/p\\ 0 & p\end{pmatrix}. \tag{9}$$

The position of the RP's that have an object–image relation is obtained by setting the upper right element of this matrix equal to zero, giving

$$D_2=-D_1/p^2-a_{12}/p,$$

and (10)

$$\Delta D_2=-\Delta D_1/p^2.$$

From Eqs. (9) and (10) we draw the following important conclusions:

(a) The lateral magnification m, introduced before, is the same for all pairs of RP's that have an object–image relation, and is the reciprocal of the power of the telescopic system

$$m=1/p\,(=\text{const}). \tag{11}$$

(b) If the object is moved by the distance ΔD_1, the image moves the distance $\Delta D_1/p^2$ in the same direction. This is equivalent to a constant axial magnification m_{ax}:

$$m_{ax}=1/p^2=m^2\,(=\text{const}). \tag{12}$$

Equation (11) allows a very simple and well-known determination of the power of a telescope: The ratio of the sizes of the objective aperture (entrance pupil) and its image (exit pupil) gives directly the power of a telescope. In Sec. VC, we show that Eq. (12) is very important for the design of optical systems that use both telescopic and focusing systems.

C. Focusing Systems

Referring to Eq. (4), we define systems with $a_{21}\neq 0$ as focusing systems because, as we show now, they have one FP (real or virtual—see Ref. 4) on either side of the optical system. To show this, we introduce, as in the previous section, two new RP's (Fig. 6) and find the FP's by setting the diagonal elements of A_{03} equal to zero; we have

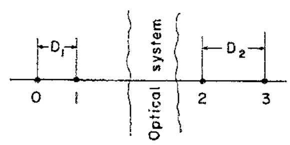

FIG. 6. Relative position of RP's for the derivation of Eq. (9) (D_1 and D_2 are positive in this figure).

$$A_{03}=\begin{pmatrix}1 & D_2\\ 0 & 1\end{pmatrix}\begin{pmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{pmatrix}\begin{pmatrix}1 & D_1\\ 0 & 1\end{pmatrix}.$$

After the multiplications are carried out, we have

$$A_{03}=\begin{pmatrix}a_{11}+D_2a_{21} & a_{12}+a_{11}D_1+D_2(a_{22}+a_{21}D_1)\\ a_{21} & a_{22}+D_1a_{21}\end{pmatrix}. \tag{13}$$

Setting the diagonal elements of this matrix equal to zero, we find that there is one and only one FP associated with each side of the optical system.

To relate the matrix elements of Eqs. (4) or (13) to the focal lengths (FL's) of the optical system, we have to generalize the customary definition of the FL for a thin lens, which is usually defined as the distance between the lens and the FP (Fig. 7). If one imagines that the lens is made inaccessible by placing the lens between two very thin glass plates, as indicated by the dashed lines in Fig. 7, one is led to the following generalized definition of the FL:

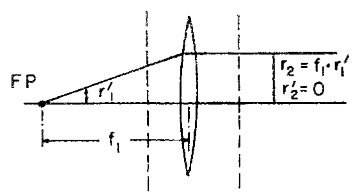

FIG. 7. Definition of the FL for a thin lens.

[4] In general, it is of course impossible to distinguish between real and virtual objects and images, the difference being only their physical accessibility. In the design of specific systems, the distance between physical barriers (lens surfaces) and RP's are known and have to be taken into account if accessibility is of importance.

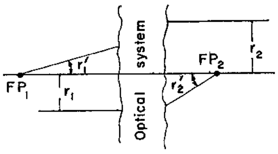

FIG. 8. Generalized definition of the FL's.

Referring to the ray originating from focal point 1 in Fig. 7 or 8, we have

$$f_1 = r_2/r_1'. \tag{14a}$$

Referring to the ray going through focal point 2 in Fig. 8 and taking into account that for a positive FL r_1 and r_2' have opposite signs, we obtain

$$f_2 = -r_1/r_2'. \tag{14b}$$

We introduce two different FL's since we do not know yet under what circumstances they are equal.

With Eqs. (14) we can now easily construct the matrix connecting the two FP's of an optical system:

For a ray originating from focal point one ($r_1=0$), we get in FP_2 $r_2 = f_1 \cdot r_1'$ and $r_2'=0$. Using the notation of Eq. (4) we get $a_{12}=f_1$ and $a_{22}=0$. A ray parallel to the optical axis ($r_1'=0$), intersecting FP_1 at r_1, gives in FP_2 $r_2=0$ and r_2'

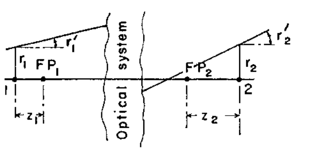

FIG. 9. Definition of z_1 and z_2 for Eq. (15) (z_1 and z_2 are positive in this figure).

$=-r_1/f_2$, resulting in $a_{11}=0$ and $a_{21}=-1/f_2$. The matrix connecting the two FP's, therefore, has the form

$$A_{\mathrm{FP}_1-\mathrm{FP}_2} = \begin{pmatrix} 0 & f_1 \\ -1/f_2 & 0 \end{pmatrix}.$$

The matrix establishing the connection between RP_1, located at the distance z_1 to the left of FP_1, and RP_2, located at the distance z_2 to the right of FP_2, is obtained by multiplying the matrix $A_{\mathrm{FP}_1-\mathrm{FP}_2}$ from the left by $\begin{pmatrix} 1 & z_2 \\ 0 & 1 \end{pmatrix}$ and from the right by $\begin{pmatrix} 1 & z_1 \\ 0 & 1 \end{pmatrix}$ (Fig. 9). Since we have made these multiplications already, we obtain the result from Eq. (13) by setting $a_{11}=a_{22}=0$, $a_{12}=f_1$, $a_{21}=-1/f_2$, $D_1=z_1$, and $D_2=z_2$; we then have

$$\left.\begin{aligned} &\mathbf{r}_2 = A_{12}\mathbf{r}_1, \\ &A_{12} = \frac{1}{f_2}\begin{pmatrix} z_2 & z_1z_2 - f_1f_2 \\ 1 & z_1 \end{pmatrix}, \\ &\text{and} \\ &\|A_{12}\| = f_1/f_2 = n_1/n_2. \end{aligned}\right\} \tag{15}$$

From Eq. (15) we learn that, for any air–air system, $f_1=f_2$. Setting the upper right element of A_{12} in Eqs. (15) equal to zero, we get for the positions of planes that have object-image relation

$$z_1 z_2 = f_1 f_2. \tag{16}$$

If Eq. (16) is fulfilled, for the lateral magnification $m=r_2/r_1$ we obtain from Eqs. 15 and 16

$$m = -z_2/f_2 = -f_1/z_1. \tag{17}$$

Contrary to the magnification of a telescopic system [Eq. (11)], the magnification obtained with a focusing system depends on the position of the object on the optical axis. A simple application of Eq. (17) is discussed in Sec. VA.

It is evident from Eq. (16) that the axial magnification of a focusing system depends also on the position of the object. It is, therefore, only possible to define an infinitesimal axial magnification, i.e., an axial magnification that is constant only over an infinitesimal region along the axis. From Eq. (16) we obtain

$$m_{\mathrm{ax}} = \frac{dz_2}{dz_1} = \frac{f_1 f_2}{z_1^2} = \frac{f_2}{f_1}\cdot m^2 = \frac{n_2}{n_1}\cdot m^2. \tag{18}$$

Equation (18) indicates that the image always moves in the same direction as the object, except when the object goes through the FP; in that case, the image moves from $+\infty$ to $-\infty$, or vice versa, depending in what direction the object moves.

To get more familiar with the very important and often used Eq. (15), we apply Eq. (15) to Eq. (5b), the latter describing the relation between the vertex planes of a thick lens. If we simplify Eq. (5b) by assuming $n_3=n_1$ (as it is in any air–air lens), we know from Eq. (15) that $f_2=f_1$, thus making the indices 1 and 2 unnecessary. If we use Eq. (15) to describe the

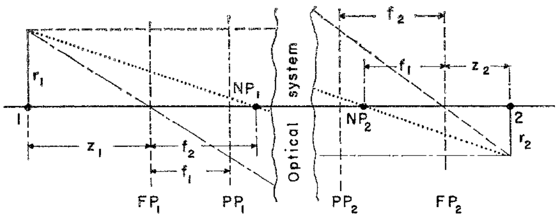

FIG. 10. Geometrical image construction and relative positions of FP's, PP's, and NP's of an optical system with positive FL's.

relation between the vertex planes of a thick lens, Eqs. (5b) and (15) must be equivalent since both equations describe the same system. The difference between Eqs. (5b) and (15) is based only on the choice of quantities used to express the matrix elements (i.e., the radii of curvatures, distances and indices of refraction, vs FL and position of the vertex planes with respect to the FP's). We can, therefore, directly compare the matrix elements. From the lower-left elements we get for the FL

$$\frac{1}{f}=[(n_2/n_1)-1] \times\{K_1+K_2-dK_1K_2[1-(n_1/n_2)]\}. \quad (19)$$

To obtain the position of the right FP, we compare the upper-left elements of Eqs. (5b) and (15); this yields

$$z_2=-f\{1-dK_1[1-(n_1/n_2)]\}. \quad (20)$$

It should be emphasized that this z_2 describes the position of the right vertex plane of the thick lens (RP_4 in Fig. 4) with respect to the right FP of the thick lens. According to our convention regarding the sign of z_2 (Fig. 9), RP_4 in Fig. 4 would lie to the left of the right-sided FP if z_2 in Eq. (20) were negative, as it would be for a positive, not-too-thick lens $[1-dK_1(1-n_1/n_2)>0]$.

An equivalent comparison of the lower right matrix elements of Eqs. (5b) and (15) gives the position of the other FP.

If we disregard the effect of apertures for the moment, the FL's and the position of the two FP's relative to an optical system describe that system completely. Despite this fact, it is customary and very practical to introduce two new concepts, namely the principal planes (PP) and nodal points (NP).

The PP's are defined as the pair of planes that have an object–image relation with a magnification $m=1$. After satisfying the requirements of Eq. (16) and setting $m=1$ in Eq. (17), we obtain for the position of the PP's with respect to the corresponding FP's [see (Fig. 10)][5]

$$z_2=-f_2 \quad \text{and} \quad z_1=-f_1. \quad (21)$$

We can, therefore, redefine the FL's as the distance between the FP's and their corresponding PP's, a definition which is often used instead of the one that we used originally. We can also use the distances a and b of two RP's from their respective PP's, instead of z_1 and z_2, to describe the positions of two RP's that have an object-image relation. By introducing $z_1=a-f$ and $z_2=b-f$ into Eq. (16), we get, for a system with $f_1=f_2=f$, the often-used relation $1/a+1/b=1/f$.

The two NP's are located on the optical axis and their position on the optical axis is such that a ray going through NP_1 with r_1', goes through NP_2 with $r_2'=r_1'$. Since the two NP's have an object–image relation, Eq. (16) has to be satisfied. Using this and setting $r_1=0$ in Eq. (15), for $r_2'=r_1'$ we obtain $z_1=-f_2$ and $z_2=-f_1$ (see Fig. 10). This, together with Eq. (21), shows that the NP's are in the PP's for systems with $f_1/f_2=n_1/n_2=1$, i.e., for all air–air systems. Besides depicting the relative positions of the FP's, PP's, and NP's for a system with positive FL's, Figure 10 also shows the self-explanatory geometrical construction of the image of r_1, using the properties of the FP's, PP's, and NP's. This construction again allows us to directly prove the two equations contained in Eq. (17), thus giving Eq. (16) again.

In some cases, the concept of NP's leads to a simple determination of PP's. The ray drawn in

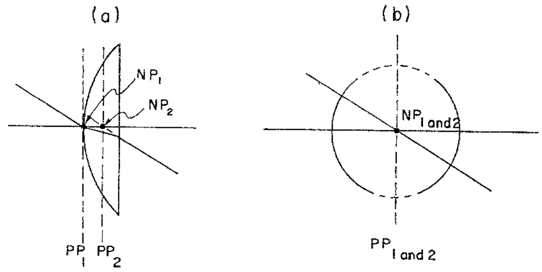

FIG. 11. Examples for simple determination of NP's and PP's.

[5] It should be noted that a set of RP's for which $m=-1$ is obtained from Eqs. (21) by changing the sign of the right sides of Eqs. (21).

Fig. 11(a) is obviously refracted twice as if it were going through a plane-parallel plate, thus giving the position of the NP's and, therefore, the PP's. Figure 11(b) shows the position of the NP's and PP's for a typical "makeshift" cylinder lens, namely a polished lucite rod. Obviously, the ray drawn in Fig. 11(b) is not refracted, showing that the two NP's are at the center of the lens, thus giving the two coinciding PP's as indicated.

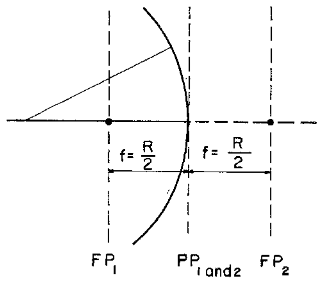

FIG. 13. Concave mirror.

D. Use of Mirrors and Significance of the Sign of FL's

When a mirror is used to connect two optical systems, this should be done in such a way that the optical axes of the two systems are reflected into each other, as shown in Fig. 12. Although the optical axis physically is then no longer a straight line, all the symmetry properties required above are still fulfilled optically. We, therefore, still draw the optical axis as a straight line in the drawings of optical systems, even though plane mirrors might actually be used, and it seems that no modifications of the whole analysis above are necessary. This is true with one exception:

In our definition of the FL's [Eq. (14)], we used the signs of r and r', belonging to different sides of the optical system. Essentially the same is true if the FL's are defined as the distance between PP's and FP's, since the definition of the PP's requires the comparison of two r's located on different sides of the optical system. While these signs are well defined when no mirrors are used, they can become ambiguous when plane mirrors are part of the optical system. If, for example, one mirror is used, the image of a right-handed coordinate system is a left-handed coordinate system; and with the use of several mirrors it is possible that the image of an object in any RP is turned by an arbitrary angle with respect to the object, although the object and image RP's might still be parallel. This makes it apparent that the signs of the r and r', on one side of such an optical system with respect to the signs of r and r' on the other side, can become a matter of convention or definition. As a consequence, the sign of the FL's then becomes a matter of definition as well. This might come as a surprise to anyone who deals mostly with single, thin, positive and negative lenses, therefore associating a positive FL with a focusing system and a negative FL with a defocusing system. If one wants to differentiate between focusing and defocusing systems, a much better criterion seems to be whether or not the focal point of interest is real or virtual, i.e., accessible or inaccessible, and we see in the discussion of the doublet (with no mirrors and, therefore, no sign difficulties) that it is quite simple to build a system with a negative FL but real FP's.

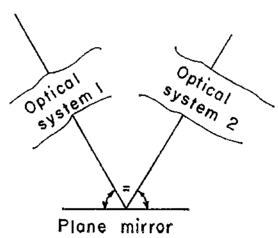

FIG. 12. Proper use of plane mirror.

For completeness we add a diagram representing a concave mirror (Fig. 13). Because of the mirror action, the optical axis is reflected back into itself, having opposite direction after reflection. We represent this part of the optical axis as an extension of the first part of the optical axis. It is easy to show that the vertex plane of the concave mirror can be chosen as the location for the two PP's and that the FL is $f = R/2$, leading to the position of the FP's as indicated in Fig. 13. With the same choice for the position of the PP's the FL of a convex mirror becomes $f = -R/2$.

V. APPLICATIONS

In the following applications we use positive lenses exclusively because the number of applications requiring negative lenses is very limited and, as a result, negative lenses are available only in a very limited number of FL's and diameters. The figures should be interpreted as schematic representations because no attempt has been made to draw them properly to scale. Individual lenses are always drawn as biconvex lenses, although one would often actually use different types, such as planoconvex lenses or, most of the time, achromats. We discuss exclusively air–air systems ($n_1=n_2$), so that the two FL's of our systems become equal.

After briefly mentioning a simple method for the determination of the FL and the position of the FP's of an optical system, we discuss three optical systems that can be used advantageously to assemble optical systems in the laboratory. We then analyze some typical problems that the experimental physicist may have to solve in his work.

A. Measurement of the FL and Position of the FP's of an Optical System

Equation (17) allows us to perform these measurements in a simple way: Using a real or virtual object in such a position that the magnification of the image with respect to the object can be measured, the magnification for this first measurement is given by Eq. (17) as

$$m_1=-z_2/f.$$

Changing the object–system distance, the distance between the system and the image has to be changed by a measurable length d and one obtains for this second measured magnification

$$m_2=-(z_2+d)/f.$$

The difference of these two equations gives for the FL

$$f=d/(m_1-m_2).$$

With this now known value for f, the equation for m_1 or m_2 then gives the position of the FP on the image side, and $z_1=f/z_2$ gives the FP on the object side of the system. An often used special case of this method consists of leaving the distance between object and image fixed and locating the system in those two positions where one obtains an image of the object.

B. The Doublet

This system is of great importance since such instruments as telescopes and microscopes are basically doublets; the doublet is, furthermore, very useful in the laboratory since it makes it possible to obtain a system with a specified FL when only lenses with a limited variety of FL's are available. According to Fig. 14, and by using Eq. (15) to calculate the matrix A_{13}, which relates RP_1 to RP_3, we obtain

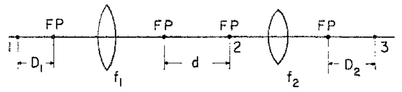

FIG. 14. The doublet.

$$A_{13}=\frac{1}{f_1f_2}\begin{pmatrix} D_2 & -f_2^2 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} d & dD_1-f_1^2 \\ 1 & D_1 \end{pmatrix}$$

and, therefore,

$$A_{13}=\frac{1}{f_1f_2}\begin{pmatrix} D_2d-f_2^2 & D_2(dD_1-f^2)-D_1f_2^2 \\ d & D_1d-f_1^2 \end{pmatrix}. \quad (22)$$

Comparing the lower-left matrix elements of Eqs. (22) and (15), we obtain for the FL of the doublet

$$f=-f_1f_2/d, \quad (23)$$

indicating that f can be varied continuously by changing d.

Setting the upper-left element of Eq. (22) equal to zero, we obtain for the position of the FP on the right side of the system

$$D_2=f_2^2/d=-ff_2/f_1,$$

leading to a real FP for a negative FL as mentioned in Sec. IVD. Because of symmetry, the same holds for the other FP.

For $d=0$, Eq. (22) describes a telescopic system with the power[6] $p=-f_1/f_2$. Focusing a telescope to a finite distance requires us to change d in such a way that the image, as seen from RP_3, appears to be at infinity. Since this is equivalent to the statement that the left-sided FP of the system has to coincide with the object plane, we obtain d by setting the lower right element of the matrix in Eq. (22) to zero, with D_1 describing the position of the object, and we have

$$d=f_1^2/D_1. \quad (24)$$

Although the "telescope" is then a focusing and

[6] In telescopes for terrestrial use, p is usually made positive with the help of prisms that act as mirrors.

not a telescopic system, the FL of the system is not a very useful concept in this case and it is customary to still call the system a telescope. Using the distance d given by Eq. (24), one gets from Eq. (22)

$$r_2' = f_1 r_1 / f_2 D_1 = -pr/D_1.$$

This means that the angle under which an object at a finite distance is seen through a telescope focused on that object is p times the angle under which the object is seen from the left-sided focal point of the objective of the telescope.

Equation (23) indicates that it is possible to realize very small FL's by using lenses with small FL's f_1, f_2 spaced by a large distance so that d becomes large compared to these FL's. This, of course, is the description of the construction of a microscope, although in reality both lenses are multiplets themselves in order to minimize aberrations.

C. Combination of a Focusing with a Telescopic System

In Sec. VB we dealt with the problem of designing an optical system with an adjustable FL. It is often equally important to be able to change the distance between the FP's of an optical system without changing the FL. This can be done conveniently by combining a telescopic system with a focusing system. If in Fig. 15 RP_1

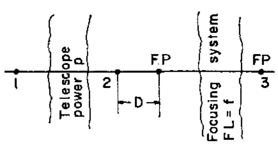

FIG. 15. Combination of a telescope and a focusing system.

and RP_2 are the two RP's that go through the entrance and exit pupil of the telescope, or any other pair of RP's that have an object–image relation, the matrix connecting $\mathbf{r}_1$ to $\mathbf{r}_3$ is given by

$$A_{13} = -\frac{1}{f}\begin{pmatrix} 0 & -f^2 \\ 1 & D \end{pmatrix}\begin{pmatrix} 1/p & 0 \\ 0 & p \end{pmatrix}$$

$$= -\frac{1}{f \cdot p}\begin{pmatrix} 0 & -f^2p^2 \\ 1 & Dp^2 \end{pmatrix}. \quad (25)$$

Comparison of Eq. (25) with Eq. (15) yields the following results:

(a) The FL of the combined system is given by $f_{\text{comb}} = p \cdot f$, and is, therefore, independent of the distance between the telescopic and focusing system. This result can also be easily obtained without the use of matrices; we simply inspect a ray that is parallel to the axis before it enters the telescopic system, making use of the properties of a telescopic system and using Eq. (14b) for the definition of the FL, we again get $f_{\text{comb}} = p \cdot f$.

(b) If the left-sided FP of the focusing system lies at a distance D to the right (left) of RP_2, the left-sided FP of the combined system lies at the distance $D \cdot p^2$ to the right (left) of RP_1. This is of

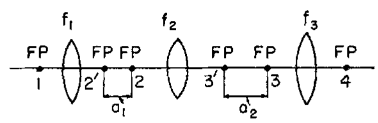

FIG. 16. Zoom-triplet ($RP_{2'}$ and $RP_{3'}$ are used in Sec. VE).

course a direct consequence of Eq. (10). It is obvious that the right-sided FP of the combined system coincides with the right-sided FP of the focusing system.

These properties of this system can be used very advantageously when two rather immovable objects, such as a spectrometer and a heavy apparatus, have to be connected optically by a system with exactly specified FL and position of the FP's. One would first build a focusing system, usually a doublet, that together with the telescope has the desired FL. The required distance between the FP's of the system can then be obtained without any change of the combined FL by properly adjusting the distance between the focusing system and the telescopic system.

D. Design of a Simple "Zoom" Lens

For some applications it is desirable to have an optical system whose FL can be varied without changing the distance between its FP's. Although this can be achieved with the system discussed in Sec. VC, we demonstrate that it is possible to design a system that is much simpler in every respect. By using Eq. (22) it can easily be shown that a doublet cannot have the desired properties. We, therefore, discuss a triplet, since it has one more distance that can be changed (Fig. 16).

To obtain a very simple system, we impose, in addition, the condition that we move only one of the three lenses in order to change the FL without changing the distance between the FP's of the system. Since moving lens 1 (or lens 3) alone is equivalent to a variation of the distance between the two components of a doublet, we, therefore, want to move lens 2 alone, thus keeping the distance between lenses 1 and 3 constant, implying $a_1+a_2=a=\text{const}$. From Fig. 16, we obtain for the matrix A_{14}, which describes the relation between $\mathbf{r}_1$ and $\mathbf{r}_4$,

$$A_{14}=A_{34}\cdot A_{23}\cdot A_{12}.$$

Using Eq. (15) to express the three matrices on the right side of this equation in terms of the FL's and a_1 and a_2, we get for A_{14}, after the multiplications are carried out,

$$A_{14}=-\frac{1}{f_1f_2f_3}\begin{pmatrix} -a_1f_3^2 & f_1^2f_3^2 \\ a_1a_2-f_2^2 & -a_2f_1^2 \end{pmatrix}. \quad (26)$$

For the FL of the system, comparison of Eq. (26) with Eq. (15) gives

$$f=f_1f_2f_3/(a_1a_2-f_2^2).$$

Since we impose the condition $a_1+a_2=a=\text{const}$, the distance between RP_1 and RP_4 is independent of a_1. Therefore, by again consulting Eq. (15), we find as the condition for a constant distance between the FP's of the system that the sum of the diagonal elements of Eq. (26), divided by the lower-left element, must be independent of the position of lens 2. By using $a_2=a-a_1$, this quantity becomes

$$[a_1f_3^2+(a-a_1)f_1^2]/[f_2^2-a_1(a-a_1)].$$

Since the denominator is of second order in a_1 and the numerator only of first order in a_1, this expression can be independent of a_1 only if the numerator vanishes for all a_1. This gives us the conditions $a=a_1+a_2=0$ and $f_3^2=f_1^2$. Since we usually work with positive lenses only, the second condition is equivalent to $f_3=f_1$. When these conditions are fulfilled, the FL of the system is given by

$$f=-f_2f_1^2/(a_1^2+f_2^2),$$

and the distance between the FP's of the system is independent of a_1 and, therefore, of f.

E. Treatment of Apertures

Up to this point we have paid no attention to the effects of apertures in an optical system. Apertures are, of course, unavoidable, since every lens mount represents an aperture. As we see in Sec. VF and VG, apertures can be very important or even be the main concern in the design of an optical system. It is, of course, practically impossible to develop a general theory that describes the effects of all apertures in an optical system, since the detailed effects of the apertures depend too much on the problem to be solved with the optical system. We can, however, give a method that usually makes a discussion of the effects of the apertures fairly simple: Instead of dealing with the apertures themselves we project all of them into the space to the left or the right of the optical system, and we see that this can be done with very little additional computational effort. That this procedure achieves exactly the intended purpose can easily be seen from Fig. 17: If a light ray, originating from P, intersects the

FIG. 17. Effect of an aperture.

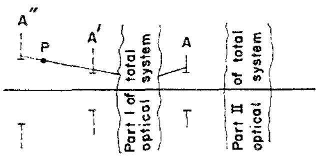

image A' of the aperture A, it must also intersect A and is therefore not transmitted through the whole optical system. The same is obviously also true if A'' were to be the image of A and it the backward extension of the ray were to intersect A'' as indicated in Fig. 17.

For the cases where the location and size of the projection of the apertures are not trivial, we use the example of the triplet, Fig. 16, to demonstrate how the images of all apertures can be obtained with very little more work than is already necessary to obtain the matrix describing a system. If we want to project all apertures into the space to the left of the optical system, we should calculate the matrix A_{14} in the following way: We first write down A_{12}, then multiply this matrix from the left by A_{23} to obtain A_{13}, then multiply this matrix from the left by A_{34} to obtain A_{14}, and so on, if there are more lenses. For example, if we want to find the image of an aperture that lies between lens 2 and 3 (this could, of course, be a lens mount of lens 2 or 3!), we use the matrix A_{13} and insert for a_2 such a value that, for the purpose of this calculation, RP_3 coincides with the plane of the aperture. By using the intermediate results of the calculation of the matrix A_{14} in this way, we obtain without any additional computations the matrices connecting all aperture planes with RP_1. If we want to project all apertures into the space to the right of the optical system, we calculate analogously $A_{3'4}$, $A_{2'4}$, A_{14}, in that order (see Fig. 16).

To actually obtain the image of an aperture from the matrix that connects the aperture plane with a RP outside of the optical system, we use the notation of Eq. (4) with RP_1 and RP_2 representing these RP's. Referring to Fig. 6 as well as to Eq. (13) and its derivation (with D_1 and D_2 replaced by d_1 and d_2), we find the location of the image of RP_2 by setting the upper-right element of the matrix in Eq. (13) equal to zero for $d_2=0$, obtaining

$$d_1 = -a_{12}/a_{11} \tag{27a}$$

for the location. The size of the image of RP_2 is given by the upper-left element of the matrix in Eq. (13) and becomes

$$r_0 = r_2/a_{11}. \tag{27b}$$

Analogously we find that the image of RP_1 is given by

$$d_2 = -a_{12}/a_{22}, \tag{27c}$$

and

$$r_3 = r_1\left(a_{11} - \frac{a_{21}a_{12}}{a_{22}}\right) = r_1 \cdot \frac{\|A_{12}\|}{a_{22}} = r_1 \cdot \frac{n_1/n_2}{a_{22}}. \tag{27d}$$

These results are schematically represented for the usual case $n_1/n_2=1$ in Fig. 18, which also shows the angles extended by the images of $r_1(r_2)$ as seen from the intersection between the axis and $RP_2(RP_1)$.

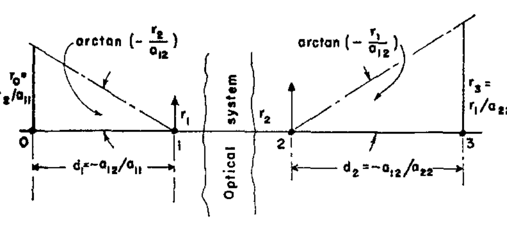

FIG. 18. Projection of RP's (or apertures) for $n_1/n_2=1$ (d_1 and d_2 are positive in this figure).

F. Parallax-Free Photography

When we take a photograph with a camera consisting of a lens system and a film (which we assume to be perpendicular to the axis of the lens system), there is obviously only one plane in the object space that has an object–image relation with the film plane. We call this plane the main object plane (MOP). Because of the finite size of the aperture of the lens, any other plane is reproduced on the film with a resolution that decreases with increasing distance between the plane under consideration and the MOP. When we photograph an object that has a considerable length in the axial direction, we can, under some circumstances, have an additional loss of resolution because of the parallax. With this term we describe the fact that, in normal photographic techniques, the size of the reproduction of an object plane on the film depends on the distance between the object plane and the camera, becoming larger when that distance decreases. An

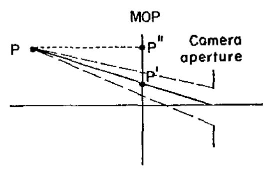

FIG. 19. Limitations of resolution because of parallax and the size of the camera aperture.

example in which parallax can diminish or limit the resolution is a plasma column that exhibits the same luminosity pattern in all planes perpendicular to the axis of the column. To better understand both the limited resolution caused by the finite camera aperture and the parallax, we refer to Fig. 19, which illustrates the reproduction of a point source P that is not located in the MOP. The dashed lines indicate the light cone that is accepted by the camera aperture,[7] and the solid line represents the center of this cone. The intersection between the light cone and the MOP is a circle with its center at P'. If point P lies between the MOP and the camera, we work with the backward extension of the light cone and obtain an equivalent result. Since the MOP has an object–image relation with the film plane, the reproduction of the point P on the film plane

[7] In cases where the physical camera aperture is not in front of the lens system, we mean by "camera aperture" the projection of the physical aperture into the object space, i.e., the aperture as it appears from object space.

is identical with this circle, except for the magnification with which the image of the MOP is formed on the film plane. We draw from this the conclusion that the parallax, which in this case amounts to the distance $P'-P''$, depends only on the location of the aperture; therefore, the parallax can be completely eliminated if it is possible to move the aperture, as seen from the object space, to infinity. This requires that the lens mount of the first lens of the optical system that achieves this does not act as an aperture, which implies that this first lens has to be somewhat larger in the radial direction than the object. Since highly corrected systems with free apertures of the order of 15 cm are available as war surplus material, this requirement is not prohibitive for many experiments.

Before we go into the discussion of a system that eliminates the parallax, we have to determine what should be the magnification m (=image size: object size) and the full opening angle α of the light cone of the system. If we keep the exposure time constant, these two quantities cannot be chosen independently if we want to obtain a negative of a certain density; if a circle with radius r_1 in the MOP radiates with a given power density σ, the energy accepted by the optical system during the exposure time t is proportional to $\sigma t \cdot r_1^2 \cdot \alpha^2$. Since, with the magnification m, the image of that circle has the radius $r_2 = mr_1$, the energy density on the film plane becomes $\sigma t r_1^2 \alpha^2 / r_2^2 = \sigma t \cdot \alpha^2/m^2$. Since we keep the exposure time t constant α^2/m^2 has to be constant. To relate α^2/m^2 to something familiar, we note that if a photograph is taken with a camera from a distance D that is large compared to the FL f of the camera lens, which shall have the diameter d, we have $|\alpha| = d/D$ and $|m| = f/D$. If we introduce the f number $N = f/d$ of the lens, we obtain[8]

$$\alpha/m = 1/N. \tag{28}$$

To determine the best choice for m, we consider the resolution that can be obtained if both the resolution of the film and the opening angle of the light cone are taken into account. To calculate the resolution resulting from several sources, we use the square root of the sum of the squares of the individual resolutions; this is usually a good approximation. If we refer all resolutions to the MOP, the resolution ϵ of the film leads to

[8] For simplicity we omit the absolute value signs | | from here on.

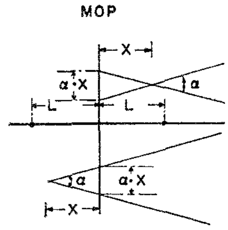

FIG. 20. Resolution because of the opening of the light cone.

$$R_\epsilon = \epsilon/m. \tag{29a}$$

The resolution with which a plane at distance x from the MOP is reproduced because of the opening angle α of the light cone is given by $\alpha \cdot x$ (Fig. 20). We assume for simplicity that the luminosity does not depend on the axial location of the plane (as long as the plane is within the boundaries of the object of length $2L$) and that the MOP is in the middle of the object; we then obtain as the resulting resolution from all planes by square superposition

$$R_\alpha^2 = \int_0^L \alpha^2 x^2 \frac{dx}{L},$$

and

$$R_\alpha = \alpha L/\sqrt{3} = mL/N\sqrt{3}. \tag{29b}$$

The combined resolution $R_{\epsilon\alpha} = (R_\epsilon^2 + R_\alpha^2)^{\frac{1}{2}}$ becomes a minimum for

$$m = (\sqrt{3}\epsilon N/L)^{\frac{1}{2}}, \tag{30a}$$

and for this value we obtain

$$R_{\epsilon\alpha} = (2\epsilon L/N\sqrt{3})^{\frac{1}{2}} \tag{30b}$$

and

$$\alpha = m/N = (\epsilon \cdot \sqrt{3}/NL)^{\frac{1}{2}}. \tag{30c}$$

In the derivation of Eqs. (30) we kept the exposure time t constant, which is appropriate in many cases. If, however, the object exhibits some motion, characterized by a velocity v, this velocity can contribute to the over-all resolution and should be taken into account, provided t is not fixed for some other reason. If we get a properly exposed negative with the exposure time t_0 and with $\alpha_0/m_0 = 1/N_0$, the exposure time has to be $t = t_0 N^2/N_0^2$ for $\alpha/m = N$, provided the reciprocity law holds. Associated with t is the resolution $R_v = vt = vt_0 N^2/N_0^2$. Combining R_v with $R_{\epsilon\alpha}$ [Eq. (30b)] through $R_{\epsilon\alpha v} = (R_v^2 + R_{\epsilon\alpha}^2)^{\frac{1}{2}}$, we obtain a minimum of $R_{\epsilon\alpha v}$ for

$$N = N_0(\epsilon L/2\sqrt{3}v^2 t_0^2 N_0)^{1/5}, \tag{31a}$$

and $R_{\epsilon\alpha}$ becomes

$$R_{\epsilon\alpha v} = (2.5 \cdot \epsilon L/N\sqrt{3})^{\frac{1}{2}}. \tag{31b}$$

The Eqs. (30a) and (30c) are of course still valid.

To realize a parallax-free photographic system, we need, as mentioned above, a first lens that has a somewhat larger diameter than the object. The simplest system is obtained by placing the camera behind this first lens so that the camera aperture is in the FP of the first lens. Because of its location, the aperture (of diameter a) appears to be at infinity as seen from the object, and extends the acceptance angle $\alpha = a/f_1$, where f_1 is the FL of the first lens. Since the proposed lens system is a doublet, we can use Fig. 14 and

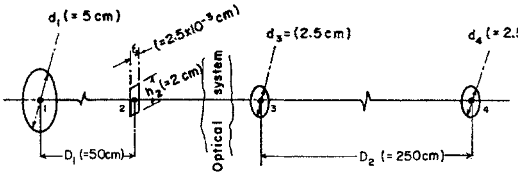

FIG. 21. Positions of the apertures in a spectrometer setup.

Eq. (22) for its description. In Fig. 14, RP_1 corresponds to the MOP, and RP_3 corresponds to the film plane. We use for the distance d the value $d = -f_2$ because the aperture of the camera lens is, in general, located fairly close to a PP of that lens. Since the MOP and the film plane have an object–image relation, we obtain the distance D_2 between the FP of the camera lens and the film plane by setting the upper-right matrix element in Eq. (22) equal to zero, yielding

$$D_2 = -D_1 f_2^2/(f_2 D_1 + f_1^2). \tag{32}$$

Using this value for D_2 in the upper-left element of Eq. (22), we obtain for the magnification m

$$m = -(f_2/f_1)[f_1^2/(f_2 D_1 + f_1^2)]. \tag{33}$$

Because one usually has only one lens big enough to be used as lens 1, f_1 is given. Since D_1, in most cases, is fixed within a fairly narrow limit by the experimental conditions and enters in Eq. (33) only weakly, Eqs. (33) and (30a) essentially determine the FL of the camera. We conclude from Eq. (32) that focusing on the MOP is simple for $D_1 < 0$, i.e., when the MOP lies between the first lens and its front FP. When $D_1 > 0$, focusing can become impossible with cameras that do not allow us to bring the film plane much closer to the lens than its back FP. Under these circumstances one has to use an auxiliary lens between the first lens and the camera. For conceptual simplicity it is advisable to place that lens in such a way that it forms together with the first lens a telescope. Using the properties of the telescope as derived in Sec. IVB, the design of the system is straightforward and is omitted here.

G. Spectrometer Setup

Spectroscopy is a very important tool for many scientific investigations and in many cases the design of the optical system that connects the spectrometer with the apparatus or specimen to be diagnosed is by no means trivial. Because of the tremendous variety of experimental conditions it seems impossible to give a general theory that takes into account all possible circumstances of an experiment. We, therefore, discuss here only one problem, which, however, is rather typical for many experiments and requires the discussion of many of the methods and techniques that would also be useful for the analysis of other experimental conditions. Since it becomes obvious that we should distinguish between several variants of this problem, depending upon the numerical values of apertures and distances, we use the values as they were encountered in an actual experiment. This makes the method clear enough to enable the reader to analyze the other variants quite easily.

In the experiment, the object to be diagnosed was a uniform plasma column $D_2 = 250$ cm long, accessible through a window of diameter $d_3 = 2.5$ cm (schematically represented in RP_3 of Fig. 21). In order to avoid the measurement of light that is emitted from the walls, light should only be admitted to the spectrometer that originates from a cylinder of 2.5-cm diam. This can easily be accomplished by introducing an aperture in the optical system so that its projection is a virtual aperture of $d_4 = 2.5$-cm diam at the other end of the plasma column, which lies in RP_4 of Fig. 21. The dimensions of the rectangular entrance slit of the spectrometer were $h_2=2$ cm and $\epsilon_2=2.5\times10^{-3}$ cm (schematically represented in RP_2 of Fig. 21). The light entering the spectrometer is further restricted by a mirror that reflects the light onto a diffraction grating. The mirror diameter d_1 is 5 cm and is separated by $D_1=50$ cm from the entrance slit. The object of this discussion is to find the optical system that, with the given apertures, gives the maximal amount of light for spectroscopic diagnosis.

For the determination of the light transmission through a set of apertures, such as in this problem, it is important to realize that the amount of light transmitted through all apertures from an infinite uniformly-illuminated plane is independent of the distance between that plane and the first physical aperture; this aperture is in our case located in RP_3. This independence can easily be seen as follows: if we select a small surface element in the first aperture and calculate the angular distribution of the radiation going through that surface element, neglecting all the other apertures, we see that this radiation is independent of the distance between that aperture and the radiating plane. Since this holds for all surface elements, we conclude that the amount as well as the angular distribution of the radiation going through the whole first aperture is also independent of the distance between that aperture and the radiating plane; thus, it follows that the radiation going through the whole set of apertures is independent of that distance too. If a whole volume is radiating uniformly, we therefore have to consider only the transmission from one plane or luminous slab; we can locate this plane so that the calculation becomes as simple as possible. If we have two apertures in RP_3 and RP_4 of Fig. 21, spaced by a distance D_2, the obvious choice for the location of the radiating plane is either RP_3 or RP_4. If we locate the plane in RP_3, and the area of the aperture opening is A_3, only the radiation from that area is transmitted through the aperture in RP_3. The amount of radiation transmitted through both apertures is therefore proportional to A_3 and the solid angle that the aperture in RP_4 extends as seen from RP_3. If the area of the aperture in RP_4 is A_4, the light transmission from this particular radiating plane, and therefore from any other plane or the whole volume, is consequently proportional to

$$T=A_3A_4/D_2^2. \qquad (34)$$

If apertures are in more than two RP's, the calculation of the transmission can become much more complicated, since the solid angle defined by two apertures as seen from the third aperture is, in general, different for different surface elements of the third aperture.

In order to allow a simple determination of the amount of light that is transmitted through the given set of four apertures in RP_1 through RP_4, we project the apertures in RP_1 and RP_2 into the space to the right of the optical system. We particularly require that the image of RP_1 falls on RP_3 and that the image of RP_2 coincides with RP_4, since these seem to be the logical positions for the images of RP_1 and RP_2 and because this allows the application of the very simple Eq. (34).

Having the distance between the images of RP_1 and RP_2 fixed in this way, we expect that this establishes some correlation between the possible magnifications m_1 and m_2 with which the images of RP_1 and RP_2 are formed. To determine this correlation, we could directly apply the expressions derived in Sec. VE. We prefer not to use these results but rather derive the relation betwen m_1 and m_2 in a straightforward way as one would do if the results of Sec. VE were not available.

Using the magnification m_1 with which the image of RP_1 is formed at the location of RP_3, the matrix A_{13} connecting $\mathbf{r}_1$ and $\mathbf{r}_3$ has to have the form

$$A_{13}=\begin{pmatrix} m_1 & 0 \\ a_{21} & 1/m_1 \end{pmatrix}, \qquad (35)$$

with a_{21} still undetermined.

To obtain the matrix A_{24} connecting $\mathbf{r}_2$ and $\mathbf{r}_4$, we have to multiply A_{13} by $\begin{pmatrix} 1 & -D_1 \\ 0 & 1 \end{pmatrix}$ from the right and by $\begin{pmatrix} 1 & D_2 \\ 0 & 1 \end{pmatrix}$ from the left, yielding

$$A_{24}=\begin{pmatrix} m_1+D_2a_{21} & D_2/m_1-D_1(m_1+D_2a_{21}) \\ a_{21} & 1/m_1-D_1a_{21} \end{pmatrix}. \qquad (36)$$

If we introduce the magnification m_2 with which the image of RP_2 is formed in RP_4, m_2 has to be equal to the upper-left element of A_{24},

$$m_2 = m_1 + D_2 a_{21} \tag{37a}$$

and the upper-right element of A_{24} has to vanish, yielding the following relation between m_1, m_2, D_1 and D_2:

$$m_1 m_2 = D_2/D_1. \tag{38}$$

We now use Eqs. (38) and (34) to determine the magnifications m_1 and m_2 for optimal light transmission. We choose $|m_2|$ as an independent variable and see how T [Eq. (34)] depends on $|m_2|$. For $|m_2| < d_4/h_2 = 1.25$, the image of the slit in RP_2 is smaller than the aperture in RP_4, giving $A_4 = \epsilon_2 h_2 \cdot m_2^2$ for the effective aperture area in RP_4. Since the image of the aperture in RP_1 is larger than the aperture in RP_3, the effective aperture area in RP_3 is $A_3 = (\pi/4)d_3^2$. Using Eq. (34) we see, therefore, that for $|m_2| < 1.25$,

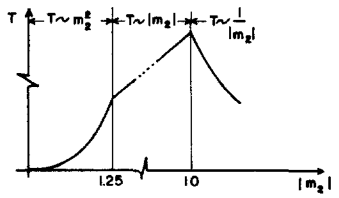

FIG. 22. Light transmission T as a function of $|m_2|$ for a spectrometer setup.

T is proportional to m_2^2. If $|m_2|$ increases beyond 1.25, $|m_2| \cdot h_2 > d_4$ and A_4 grows only proportional to $|m_2|$ whereas[9] A_3 is still given by $(\pi/4)d_3^2$ as long as the image of the aperture in RP_1 is larger than the aperture in RP_3. In this range of values of m_2, T is, therefore, proportional to $|m_2|$. If $|m_2|$ is increased beyond the point where the image of the aperture in RP_1 just fits the aperture in $RP_3(m_1 d_1 < d_3)$, the effective area of the aperture in RP_3 becomes

$$A_3 = \tfrac{1}{4}\pi m_1^2 d_1^2 = \tfrac{1}{4}\pi d_1^2 \cdot D_2^2/D_1^2 m_2^2.$$

Since A_4 is still proportional to $|m_2|$, T is from there on proportional to $1/|m_2|$. From this behavior of $T(|m_2|)$, which is schematically represented in Fig. 22, we see that we obtain a very sharp and sensitive maximum of T for that value of $|m_2|$ for which the image of the aperture in RP_1 just fits the aperture in RP_3. By using the numerical values given above, this is the case for $|m_1| = 0.5$ and $|m_2| = 10$. [It should be noted that $T(|m_2|)$ behaves differently when the numerical values are such that the image of the slit in RP_2 is still smaller than the aperture in RP_4 when the image of the aperture in RP_1 fits the aperture in RP_3]. For the optimum value of $|m_2|$, T becomes

$$T_{\max} = \tfrac{1}{4}\pi \cdot d_3^2 \cdot d_4 \cdot \epsilon_2 |m_2| / D_2^2.$$

Using $d_3 = d_4 = |m_1| \cdot d_1$ and Eq. (38), we finally get

$$T_{\max} = \tfrac{1}{4}\pi d_3^2 d_1 \epsilon_2 / D_1 D_2.$$

The best value of $|m_2|$ determines a_{21} and therefore determines the FL of the optical system by means of Eq. (37a), which can be rewritten as

$$a_{21} = (m_2/D_2)[1 - (m_1/m_2)]. \tag{37b}$$

We expressed a_{21} intentionally in this way to make the following point: While the sign of m_2 and, therefore, a_{21} is arbitrary and of no practical significance, the sign of m_1/m_2 has to be the same as the sign of $m_1 m_2$, which is determined by Eq. (38). Of course, D_2/D_1 is positive if the RP's actually have the relative position as indicated in Fig. 21. It can, however, be practical to throw the image of RP_1 on the RP farthest to the right (where RP_4 is in Fig. 21) and RP_2 on the RP close to the optical system (where RP_3 is in Fig. 21). Since this represents a reversal of the positions of RP_3 and RP_4, it follows that, in this case, D_2, and consequently $m_1 m_2$ and m_1/m_2, has a negative sign, leading to a slightly different value of $|a_{21}|$. With the RP's in the positions as in Fig. 21 we obtain from Eq. (37b), after introduction of the numerical values introduced above, $|1/a_{21}| = 25/0.95 \approx 26.3$ cm.

Comparison of Eq. (35) with Eq. (15) gives, for the position of the rightsided FP with respect to RP_3,

$$m_1/a_{21} = D_2\{(m_1/m_2)/[1 - (m_1/m_2)]\}. \tag{39a}$$

Since the sign of m_1/m_2 is always the same as the sign of D_2, the right-sided FP of the optical system lies always to the left of RP_3 (if $1 - m_1/m_2 > 0$). Analogously, for the position of the leftsided FP of the optical system with

[9] Here and later we use the fact that ϵ is very small compared to d_4/m_2.

respect to RP_1 we obtain

$$\frac{1}{m_1a_{21}}=\frac{D_2/m_1m_2}{1-(m_1/m_2)}=\frac{D_1}{1-(m_1/m_2)}$$
$$=D_1\left(1+\frac{m_1/m_2}{1-(m_1/m_2)}\right). \quad (39b)$$

With Eqs. (37b) and (39), the optical system for the attainment of maximal light transmission is completely specified. If RP_3 and RP_4 are located as in Fig. 21 and if the spectrometer can be mounted close enough to the apparatus, a single lens or a doublet can be sufficient to satisfy Eqs. (37b) and (39). If either one of these conditions is not fulfilled, a system as described in Sec. VC is very practical, particularly if one wants to reverse the positions of RP_3 and RP_4.

In this whole discussion we assumed that the images of RP_1 and RP_2 coincide with RP_3 and RP_4. Although it would be too involved to describe the details here, it is fairly easy to show by dislocating the images of RP_1 and RP_2 from RP_3 and RP_4 that the optical system described by Eqs. (37b) and (39) does give a maximum for T. It can, furthermore, be shown that through the use of cylinder lenses, T cannot be improved either.

H. Huygens' Eyepiece

To demonstrate the use of matrices for the discussion of chromatic aberrations, in this section we derive the basic design of Huygens' eyepiece, at least as far as chromatic aberrations are concerned. This eyepiece consists, for economic reasons, of two thin lenses that are made of the same glass; we have to choose these lenses and the distance between them so that chromatic aberrations are least objectionable.

A completely chromatically-corrected optical system would be describable by a matrix whose elements would be entirely independent of the wavelength λ of the light. Since the refractive index n of all materials varies with λ, this is, of course, impossible. One, therefore, has to be satisfied if the first derivative of the matrix elements with respect to λ vanishes not everywhere but for at least one or preferably more wavelengths. Using Eq. (22) it is, however, easy to show that it is impossible to design a doublet of the kind described above in such a way that the first derivatives of all matrix elements disappear for a given wavelength. We, therefore, have to investigate which chromatic aberrations are most significant in the use of such an eyepiece and then design the eyepiece accordingly.

An eyepiece is used to visually observe an object that is projected into one FP of the eyepiece by means of a chromatically well-corrected objective. The image of an off-axis object point can have two kinds of chromatic aberration: The angle α_2 between the axis and the image point, as seen with the eye, can depend on λ, as well as can the distance d_2 between the image and the eye. If we identify the object plane and the plane that goes through the front NP of the eye as RP_1 and RP_2, and if we describe the connection between $\mathbf{r}_1$ and $\mathbf{r}_2$ by a matrix as in Eq. (4) (of course with $n_1/n_2=1$), from Fig. 18 we can directly obtain

$$\alpha_2=-r_1/a_{12} \quad (40a)$$

and

$$d_2=-a_{12}/a_{22}. \quad (40b)$$

We can draw from these relations the conclusion that, for satisfactory chromatic correction of an optical system used as an eyepiece, only $a_{22}'=0$ and $a_{12}'=0$ have to be fulfilled. (The prime indicates differentiation with respect to λ.) For the eyepiece discussed here, it is possible to show with the help of Eq. (22) that these two conditions cannot be satisfied simultaneously with $a_{22}=0$, which also has to be fulfilled at the reference wavelength λ_0 since the object is located in the leftsided FP of the eyepiece. Since we can achromatize either d_2 or α_2, we obviously choose α_2 for achromatization; we make this choice because every off-axis object point would appear to the eye as a line showing all colors of the spectrum if $\alpha_2'\neq 0$ for all wavelengths of interest. From Eq. (40a) it follows that $\alpha_2'=0$ for $a_{12}'=0$. Although we could try to satisfy $a_{12}'=0$ directly, it is more convenient to use the fact that the eye is always very close to the back FP of the eyepiece, so that we can assume in very good approximation that not only $a_{22}=0$ but also $a_{11}=0$ for $\lambda=\lambda_0$. If we differentiate the identity $a_{11}a_{22}-a_{12}a_{21}=1$ with respect to λ, we find that $(a_{11}a_{22})'=0$, although $a_{11}'\neq 0$ and $a_{22}'\neq 0$, and we obtain the result that $a_{12}'=0$ requires also that $a_{21}'=0$. Referring to Fig. 14 and Eq. (22), we see that $a_{21}=d/f_1f_2$. When we differentiate this equation, we have to realize that the distance d between the FP's of the two lenses depends on λ. However, since the two lenses are assumed to be thin, the positions of their PP's are in very good approximation independent of λ, and the distance between the PP's of the two lenses is practically the same as the distance D between the lenses. With $d=D-f_1-f_2$ we obtain, therefore,

$$a_{21}=(D-f_1-f_2)/f_1f_2,$$

with $D'=0$. Moreover, because the two thin lenses are made of the same glass, it is evident from Eq. (5b) that $u=f_2/f_1$ is also independent of λ. Using this when we differentiate the equation

$$a_{21}=(D-f_1-f_2)/f_1f_2 = (1/u)\{(D/f_1^2)-[(1+u)/f_1]\},$$

we directly obtain from $a_{21}'=0$ that D has to be

$$D=(f_1+f_2)/2. \tag{41a}$$

This equation can obviously only be fulfilled for one wavelength λ_0, for which one chooses the wavelength of maximum sensitivity of the eye. With the spacing given by Eq. (41a), the FL of the eyepiece becomes

$$f=2f_1f_2/(f_1+f_2). \tag{41b}$$

With Eq. (22) the following statements can easily be proven: If $f_1>0$ and $f_2>0$, only one of the FP's of the eyepiece is real. Since the eye should be in or close to the back FP of the eyepiece, this is the FP that should be accessible, which is only and always the case when $f_1>f_2$. Huygens' original design is characterized by $f_1=3f_2$ and consequently $D=2f_2$ and $f=(3/2)f_2$.

When we calculate how d_2 depends on λ, we should work with $1/d_2$ instead of d_2 directly because $d_2(\lambda_0)=\infty$. If we expand $1/d_2$ into a Taylor series and take only the first nonvanishing term, we obtain, with Eq. (40b),

$$(1/d_2)_{\lambda=\lambda_0+\Delta\lambda}=(1/d_2)_{\lambda=\lambda_0}'\cdot\Delta\lambda = -\Delta\lambda\cdot a_{22}'/a_{12}. \tag{42a}$$

From Eq. (22) we see that a_{22} can be written as

$$a_{22}=D_1a_{21}-f_1/f_2.$$

Because $a_{21}'=0$ and $(f_1/f_2)'=0$, we get

$$a_{22}'=a_{21}\cdot D_1'. \tag{43}$$

With the same consideration that led to the introduction of D earlier, we see that $D_1'=-f_1'$ Since we are dealing with a thin lens, this can, with the use of Eq. (5b), be expressed as $D_1'=f_1n'/(n-1)$. Using this and $a_{21}(\lambda_0)\cdot a_{12}(\lambda_0)=-1$ in Eq. (43) and Eq. (42a), we obtain

$$(1/d_2)_{\lambda=\lambda_0+\Delta\lambda}=a_{21}^2f_1n'\Delta\lambda/(n-1),$$

or, with $a_{21}^2=1/f^2$, we finally get

$$d_2(\lambda_0+\Delta\lambda)=f\cdot(f/f_1)\cdot(n-)1/n'\Delta\lambda. \tag{42b}$$

Without going into details here it might be worthwhile to mention that, to the eye, the apparent chromatic aberration resulting from the effect described by Eq. (42b) is proportional to the size of the exit pupil (depth of field consideration), whereas the residual higher-order chromatism in α_2 is independent of the size of the exit pupil.