

Notas de Óptica Paraxial

Leopoldo R. Gómez*
Departamento de Física,
Universidad Nacional del Sur - IFISUR - CONICET,
8000 Bahía Blanca, Argentina

(Dated: June 24, 2015)

I. INTRODUCCIÓN

La óptica paraxial es una simplificación de la óptica geométrica que se utiliza para describir, de forma fácil y eficiente, la propagación de luz para rayos que están cercanos al eje óptico. En general, los cálculos y resultados obtenidos por medio de la óptica paraxial serán una primera aproximación a un sistema óptico, siendo estrictamente válidos para sistemas en los cuales las dimensiones transversales sean mucho más chicas que las dimensiones longitudinales de objetos, imágenes, y elementos del sistema óptico.

Los rayos paraxiales son aquellos que propagan cercanos al eje óptico, tal que las desviaciones angulares respecto de este son pequeñas (ver un esquema en la figura 1). Para los rayos paraxiales valen las siguientes aproximaciones:

$$\begin{aligned}\sin(\theta) &\approx \theta \\ \cos(\theta) &\approx 1 \\ \tan(\theta) &\approx \theta\end{aligned}$$

donde θ es el ángulo que hace el rayo con el eje óptico.

Como veremos, estas aproximaciones simplifican en gran medida el trazado de rayos y cálculo de imágenes en sistemas ópticos complejos. En particular, estas aproximaciones producen la linealización de la ley de Snell $n_1 \sin(\theta_1) = n_2 \sin(\theta_2)$, tal que en la óptica paraxial la ley de Snell toma la forma:

$$n_1 \theta_1 \approx n_2 \theta_2$$

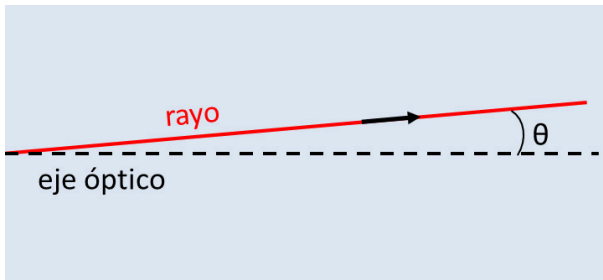


FIG. 1. Esquema de un rayo paraxial. La flecha indica la dirección de propagación.

Esta linealización de la ley de Snell va a permitir estudiar la propagación de rayos a través de un sistema por medio transformaciones lineales (matrices).

Es interesante notar que en la óptica paraxial no existe el fenómeno de aberración, ya que todos los rayos que salen de un punto convergen a otro único punto (imagen) luego de pasar por el sistema óptico. Algunos libros de texto también llaman a la óptica paraxial como óptica Gaussiana, ya que fue Gauss quien formuló sus bases alrededor del año 1800.

II. COORDENADAS Y MATRICES

Para describir un rayo que esta ubicado en un determinado plano que corta al eje óptico se utilizan dos "coordenadas", la altura del rayo en ese punto, y el ángulo θ que tiene respecto del eje óptico (ver esquema en la figura 2).

En lo que sigue vamos a ver que óptica paraxial un rayo que entra con coordenadas (y_1, θ_1) a un sistema óptico general, sale del mismo con coordenadas (y_2, θ_2) , dadas por:

$$\begin{bmatrix} y_2 \\ \theta_2 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} * \begin{bmatrix} y_1 \\ \theta_1 \end{bmatrix}$$

donde la matriz A está relacionada a los distintos elementos del sistema óptico.

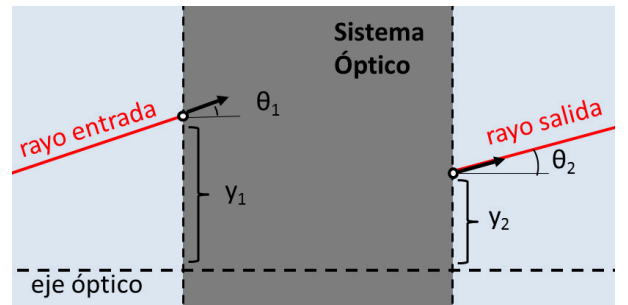


FIG. 2. Esquema de trabajo en óptica paraxial. Un rayo de entra a un sist. óptico con coordenadas (y_1, θ_1) , y sale del mismo con coordenadas (y_2, θ_2) .

* lgomez@uns.edu.ar

En lo que sigue vamos a ver que cada elemento de un sistema óptico (lentes convergentes, divergentes, medios, interfaces, etc) tienen matrices asociadas $A_1, A_2 \dots A_N$ (para un sistema de N elementos). Entonces la matriz A , asociada a todo el sistema óptico estará dada por el producto de las matrices asociadas a cada elemento:

$$A = A_1 * A_2 * \dots * A_N$$

III. ALGUNAS MATRICES IMPORTANTES ASOCIADAS A ELEMENTOS DE UN SISTEMA ÓPTICO

En lo que sigue se obtienen las matrices relacionadas a los elementos básicos de un sistema óptico

A. Traslación de un rayo

Supongamos un rayo que propaga una distancia d a lo largo del eje óptico, a través de un medio homogéneo de índice de refracción n_1 . Entonces debido a la propagación rectilínea de la luz tenemos que las coordenadas final e inicial del rayo están relacionadas por (ver figura 3):

$$y_2 = y_1 + \tan(\theta)d \Rightarrow y_2 = y_1 + d\theta$$

$$n_1\theta_2 = n_1\theta_1 \Rightarrow \theta_2 = \theta_1$$

Este sistema de ecuaciones puede ser escrito de la forma matricial:

$$\begin{bmatrix} y_2 \\ \theta_2 \end{bmatrix} = T_d * \begin{bmatrix} y_1 \\ \theta_1 \end{bmatrix}$$

donde T_d es la matriz asociada a una traslación en d a través de un medio homogéneo:

$$T_d = \begin{bmatrix} 1 & d \\ 0 & 1 \end{bmatrix}$$

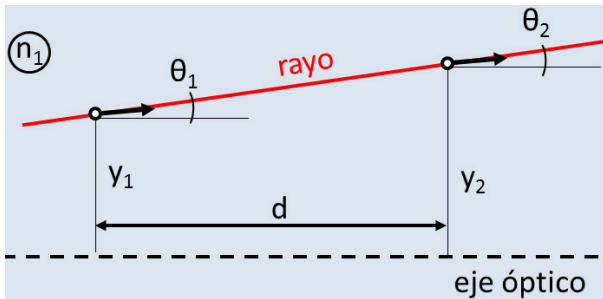


FIG. 3. Propagación rectilínea de un rayo en un medio homogéneo. En su camino, rayo se desplaza una cantidad d a lo largo del eje óptico.

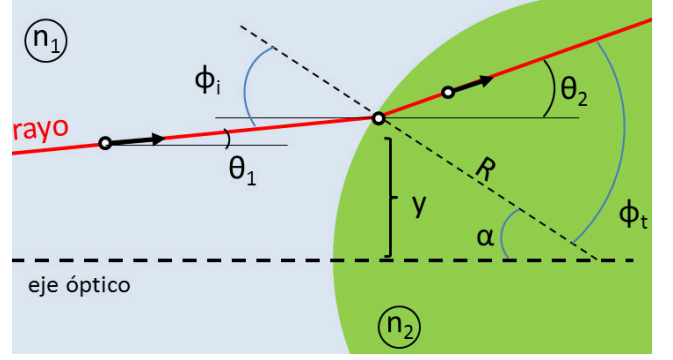


FIG. 4. Refracción de un rayo en una interfase esférica.

B. Refracción en una interfase esférica y plana

Consideremos un rayo que se refracta en una interfase esférica como en la figura 4. La refracción en una interfase es una transformación que se lleva a cabo a una misma altura ($y_1 = y_2 = y$). Es decir, la refracción es una transformación que produce sólo un cambio en el ángulo del rayo, a través de la ley de Snell (linealizada):

$$n_1\phi_i = n_2\phi_t$$

Aquí es importante notar que la ley de Snell se aplica respecto de la normal a la interfase, y entonces se debe usar los ángulos ϕ 's en lugar de los θ 's (figura 4). Sin embargo, nosotros necesitamos las coordenadas final del rayo, es decir que necesitamos ver como es el ángulo del rayo refractado respecto al eje óptico (θ_2).

De la figura 4 es fácil ver que se puede escribir la ley de Snell como:

$$n_1(\theta + \alpha) = n_2(\theta_2 + \alpha)$$

Además, se tiene que $y = R \sin(\alpha)$, tal que en la aproximación paraxial podemos escribir $\alpha \approx \frac{y}{R}$. Entonces tenemos:

$$n_1(\theta_1 + \frac{y}{R}) = n_2(\theta_2 + \frac{y}{R})$$

tal que el ángulo del rayo refractado se escribe como:

$$\theta_2 = \frac{n_1 - n_2}{n_2 R} y + \frac{n_1}{n_2} \theta_1$$

Entonces, finalmente las coordenadas de un rayo refractado por una interfase esférica puede ser escritas en forma matricial como:

$$\begin{bmatrix} y \\ \theta_2 \end{bmatrix} = R * \begin{bmatrix} y_1 \\ \theta_1 \end{bmatrix}$$

donde R es la matriz de refracción:

$$R = \begin{bmatrix} 1 & 0 \\ \frac{n_1 - n_2}{n_2 R} & \frac{n_1}{n_2} \end{bmatrix}$$

De está misma expresión se puede obtener la matriz asociada a una refracción en una interface plana, tomando el límite $R \rightarrow \infty$, la matriz queda como:

$$R = \begin{bmatrix} 1 & 0 \\ 0 & \frac{n_1}{n_2} \end{bmatrix}$$

C. Lentes gruesas y delgadas

Podemos pensar que un razo que atraviesa una lente gruesa sufre tres transformaciones consecutivas (ver figura 5):

- Primero una refracción en una interface esférica de radio R_1 ($R_1 > 0$).
- Luego el rayo se traslada en el espesor de la lente e .
- Finalmente el rayo es refractado por una interface esférica de radio R_2 ($R_2 < 0$).

Entonces, la matriz L_g asociada a una lente gruesa puede ser escrita como el siguiente producto de matrices:

$$L_g = \begin{bmatrix} 1 & 0 \\ \frac{n_3 - n_2}{n_3 R_2} & \frac{n_2}{n_3} \end{bmatrix} * \begin{bmatrix} 1 & e \\ 0 & 1 \end{bmatrix} * \begin{bmatrix} 1 & 0 \\ \frac{n_1 - n_2}{n_2 R_1} & \frac{n_1}{n_2} \end{bmatrix}$$

Podemos obtener la matriz asociada a una lente delgada L_d , tomando el límite $e \rightarrow 0$:

$$L_d = \begin{bmatrix} 1 & 0 \\ \frac{n_3 - n_2}{n_3 R_2} & \frac{n_2}{n_3} \end{bmatrix} * \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} * \begin{bmatrix} 1 & 0 \\ \frac{n_1 - n_2}{n_2 R_1} & \frac{n_1}{n_2} \end{bmatrix}$$

Para el caso de una lente delgada que esta inmersa en un mismo medio (y entonces $n_1 = n_3$) obtenemos la reconocida expresión:

$$L_d = \begin{bmatrix} 1 & 0 \\ -\frac{1}{f} & 1 \end{bmatrix}$$

donde el foco f de la lente viene dado por $\frac{1}{f} = \frac{n_2 - n_1}{n_1} \left(\frac{1}{R_2} - \frac{1}{R_1} \right)$.

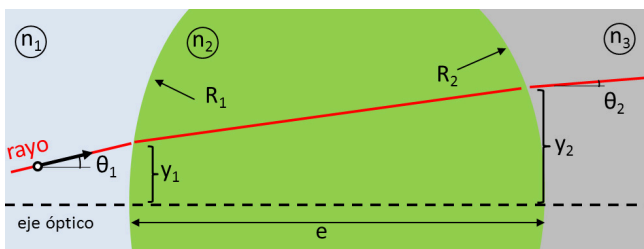


FIG. 5. Propagación de un rayo a través de una lente gruesa. Podemos pensar que esta propagación consta de 3 transformaciones sucesivas: una refracción en una interface esférica.

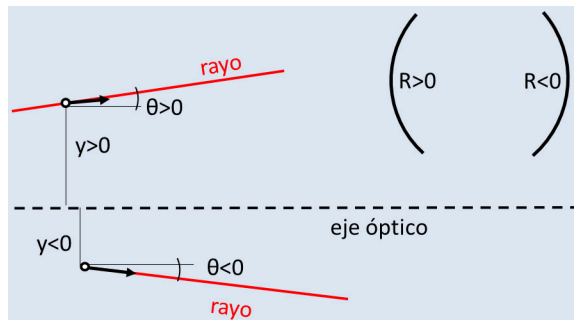


FIG. 6. Convención de signos en óptica paraxial.

D. Nota sobre la convención de signos

En general, similar a lo que ocurre con la óptica geométrica, en óptica paraxial se utiliza la siguiente convención de signos (ver figura 6):

- Para distancias de propagación positivas d , los rayos viajan de izquierda a derecha.
- Las alturas y de rayos se consideran positivas (negativas) si estan por arriba (por debajo) del eje óptico.
- En general se consideran todos los ángulos como agudos, con su respectivo signo.
- Los radios de curvatura de interfaces o superficies son positivos si el centro de curvatura está a la derecha del vertice de la superficie.

IV. SISTEMAS ÓPTICOS GENERALES

A. Puntos cardinales de un sistema óptico

Los planos y puntos cardinales ayudan a realizar el trazado de rayos de un sistema óptico general y a determinar las imágenes de objetos. Estos planos y puntos, característicos de cada sistema óptico, son:

- Planos y puntos *focales* objeto e imagen.
- Planos y puntos *principales* objeto e imagen
- Planos y puntos *conjugados*

En general, para ubicar la posición de estos planos y puntos característicos de un sistema óptico, es muy útil considerar que el objeto se ubica a una distancia D del sistema óptico, y luego de atravesar el sistema los rayos recorren otra distancia D_p (ver figura 7). Entonces las coordenadas inicial y final de los rayos se escriben como:

$$\begin{bmatrix} y_2 \\ \theta_2 \end{bmatrix} = \begin{bmatrix} 1 & D_p \\ 0 & 1 \end{bmatrix} * \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} * \begin{bmatrix} 1 & D \\ 0 & 1 \end{bmatrix} * \begin{bmatrix} y_1 \\ \theta_1 \end{bmatrix}$$

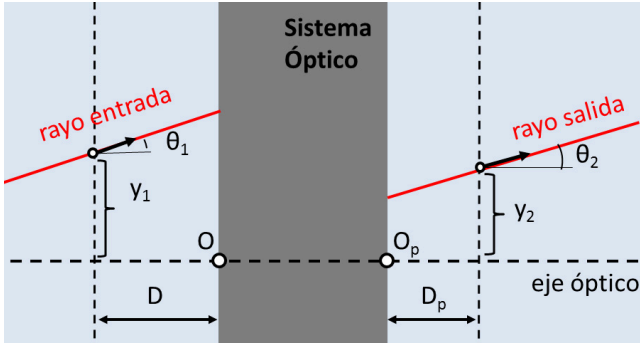


FIG. 7. Esquema de trabajo para obtener los puntos cardinales de un sistema óptico general. Un rayo sale de un objeto que se encuentra a una distancia D del sistema óptico, y luego de atravesar el mismo, propaga otra distancia D_p .

Realizando los cálculos matriciales obtenemos la relación explícita entre los rayos inicial y final:

$$\begin{bmatrix} y_2 \\ \theta_2 \end{bmatrix} = \begin{bmatrix} a_{11} + D_p a_{21} & D(a_{11} + D_p a_{21}) + a_{12} + D_p a_{22} \\ a_{21} & a_{22} + D a_{21} \end{bmatrix} \begin{bmatrix} y_1 \\ \theta_1 \end{bmatrix}$$

En lo que sigue vamos a utilizar esta relación para determinar en que posiciones del eje óptico se localizan los distintos puntos cardinales de un sistema óptico general.

B. Planos y puntos focales

Para determinar la ubicación F_p del plano focal imagen en un sistema óptico general hay que considerar que rayos que entran paralelos al sistema óptico, convergerán en un punto perteneciente al plano focal imagen (ver un esquema en la figura 8a).

La altura de este punto donde convergen rayos paralelos no depende de las alturas de los rayos iniciales y_1 , sino sólo del ángulo de ingreso al sistema óptico, tal que $y_2 = y_2(\theta_1)$. Entonces para determinar la posición del plano focal imagen consideremos que el plano de referencia a la salida del sistema óptico coincide con el plano focal imagen, tal que $D_p = \overline{O_p F_p}$, luego:

$$y_2 = (\text{constante}) * \theta_2 \quad (1)$$

Entonces el primer elemento de la matriz conectando los rayos de entrada y salida debe anularse:

$$a_{11} + D_p a_{21} = 0$$

tal que:

$$D_p = \overline{O_p F_p} = -a_{11}/a_{21} \quad (2)$$

De esta forma, el valor de $-a_{11}/a_{21}$ nos determina la posición del plano focal imagen, **medida respecto del último vértice del sistema óptico O_p** . Hay que notar

que en este cálculo supusimos que el plano focal imagen estaba ubicado a la derecha del vértice O_p . Entonces si el valor de $-a_{11}/a_{21}$ es un número positivo significa que el plano focal está efectivamente ubicado a la derecha de O_p . En el caso contrario, si para algún sistema óptico este cálculo da un valor negativo, entonces esto significaría que el plano focal imagen estará ubicado a la izquierda de O_p .

La determinación del plano focal objeto se realiza por medio de un razonamiento similar (ver figura 8b). Si tuvieramos rayos con distintos ángulos θ_1 que salen de un mismo punto de altura y_1 que estuviera ubicado sobre el plano focal objeto, entonces luego de pasar por el sistema óptico los rayos saldrían paralelos, con ángulo θ_2 que sería sólo función de la altura de los rayos iniciales $\theta_2 = \theta_2(y_1)$.

Entonces para determinar la ubicación del plano focal objeto, decimos que el objeto está ubicado en el plano focal objeto tal que $D = \overline{F O}$, tal que:

$$\theta_2 = (\text{constante}) * y_1 \quad (3)$$

y el elemento de matriz que conecta a θ_1 y θ_2 debe anularse $a_{22} + D a_{21} = 0$, tal que:

$$D = \overline{F O} = -a_{22}/a_{21} \quad (4)$$

Entonces el valor de $-a_{22}/a_{21}$ nos da la posición del plano focal objeto, **ubicado respecto al primer vértice del**

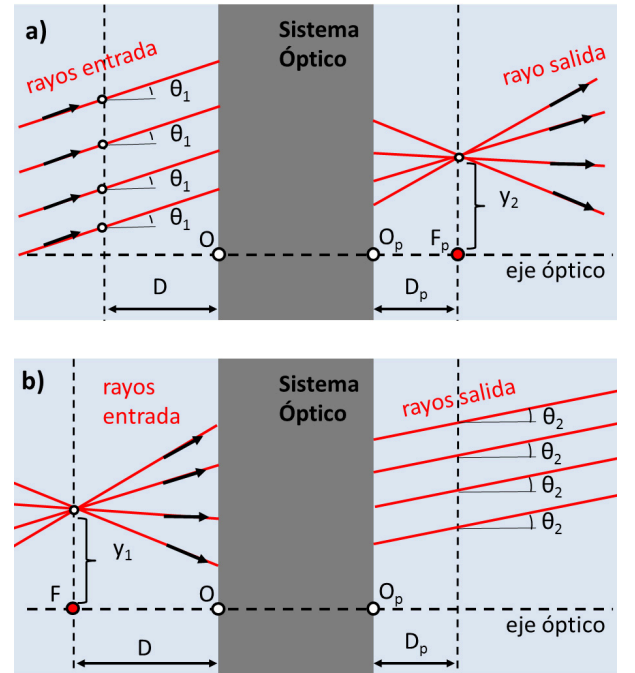


FIG. 8. Determinación de planos focales. a) Rayos que entran paralelos al sist. óptico convergen en un punto perteneciente al plano focal F_p . b) Rayos que salen de un mismo punto perteneciente al plano focal objeto F salen paralelos luego de atravesar el sistema óptico.

sistema óptico O . Notemos que para hacer el cálculo supusimos que el plano focal objeto estaba ubicado a la izquierda de O . Entonces, si el valor de $-a_{22}/a_{21}$ es positivo el plano focal objeto esta efectivamente ubicado a la izquierda de O . Si este valor es negativo, entonces el plano focal objeto estará ubicado a la derecha de O .

Matriz del sistema óptico entre planos focales: Si en la expresión que utilizamos para determinar los planos focales ponemos que las distancias D y D_p corresponden a los planos focales objeto ($D = \overline{FO} = -a_{22}/a_{21}$) e imagen ($D_p = \overline{O_p F_p} = -a_{11}/a_{21}$), obtenemos lo que se conoce como matriz del sistema óptico entre planos focales:

$$A = \begin{bmatrix} 0 & f \\ -1/f_p & 0 \end{bmatrix}$$

donde $f = -\det(A)/a_{21}$ y $f_p = -1/a_{21}$. Esta matriz será de gran utilidad a la hora de calcular la posición de imágenes de objetos.

C. Formación de imágenes

En lo que sigue calculamos la imagen de un objeto por un sistema óptico general. Supongamos que el objeto se ubica a una **distancia** x a la izquierda del **plano focal objeto** F , y la imagen se forma a una **distancia** x_p a la derecha del **plano focal imagen** F_p (ver figura 9). Entonces buscamos ver como son x y x_p , y el aumento del objeto por el sist. óptico.

Para hacer el cálculo de la imagen del objeto, es muy útil considerar la matriz del sistema entre planos focales, tal que la ecuación matricial de rayos que salen del objeto y terminan en su imagen toma la forma:

$$\begin{bmatrix} y_p \\ \theta_2 \end{bmatrix} = \begin{bmatrix} 1 & x_p \\ 0 & 1 \end{bmatrix} * \begin{bmatrix} 0 & f \\ -1/f_p & 0 \end{bmatrix} * \begin{bmatrix} 1 & x \\ 0 & 1 \end{bmatrix} * \begin{bmatrix} y \\ \theta_1 \end{bmatrix}$$

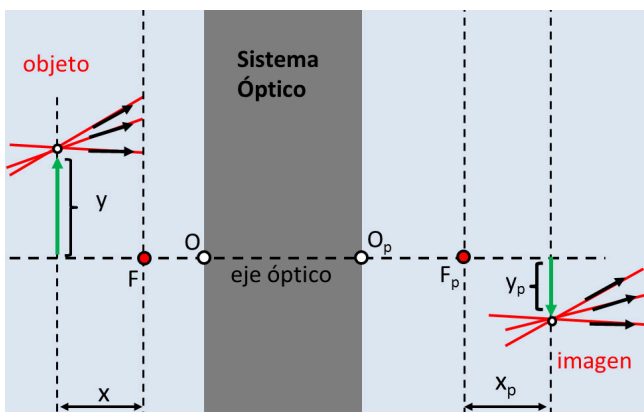


FIG. 9. Formación de la imagen de un objeto.

y haciendo los cálculos matriciales tenemos:

$$\begin{bmatrix} y_p \\ \theta_2 \end{bmatrix} = \begin{bmatrix} -x_p/f_p & -xx_p/f_p + f \\ -1/f_p & -x/f_p \end{bmatrix} * \begin{bmatrix} y \\ \theta_1 \end{bmatrix}$$

Entonces para que sea la imagen del objeto la altura de la imagen y_p no debería depender del ángulo inicial θ_1 . Entonces el elemento de matriz que conecta a θ_1 y y_p debería anularse

$$-xx_p/f_p + f = 0$$

con lo que se obtiene la localización de la imagen, medida respecto al plano focal imagen F_p :

$$x_p = \frac{ff_p}{x} \quad (5)$$

de nuevo, notemos que para hacer el cálculo supusimos que la imagen se forma a la derecha del plano focal imagen F_p . Entonces, si el valor de $\frac{ff_p}{x}$ es positivo, la imagen del objeto esta efectivamente ubicada a la derecha de F_p . Si este valor es negativo, entonces la imagen estará ubicado a la izquierda de F_p .

Luego la altura de la imagen viene dada por la expresión:

$$y_p = \frac{-x_p}{f_p} y \quad (6)$$

D. Planos principales

Los planos principales objeto H e imagen H_p se definen como aquellos planos de magnificación 1. De esta forma, todo rayo que corte al plano principal objeto H saldrá del plano principal H_p a la misma altura.

Para determinar la ubicación de estos planos podemos utilizar la ecuación para la formación de imágenes, suponiendo que el objeto se ubica en el plano principal objeto, y la imagen se forma en el plano principal imagen. Entonces de las ecuaciones de Newton tenemos:

$$y_2 = \frac{-x_p}{f_p} y_1 + \left(\frac{-xx_p}{f_p} + f \right) \theta_1 \quad (7)$$

Entonces como los planos principales tienen magnificación 1 tenemos $y_2 = y_1$, tal que:

$$\begin{aligned} \frac{-x_p}{f_p} &= 1 \Rightarrow x_p = -f_p \\ \frac{-xx_p}{f_p} + f &= 0 \Rightarrow x = -f \end{aligned}$$

De esta forma el plano principal imagen H_p estará ubicado a una distancia f_p del plano focal imagen. Debido al signo $-$, el plano principal imagen se ubica a la derecha del plano focal imagen. Similarmente, el plano principal objeto se ubica a la derecha del plano focal objeto, en el valor de f (ver figura 10a).

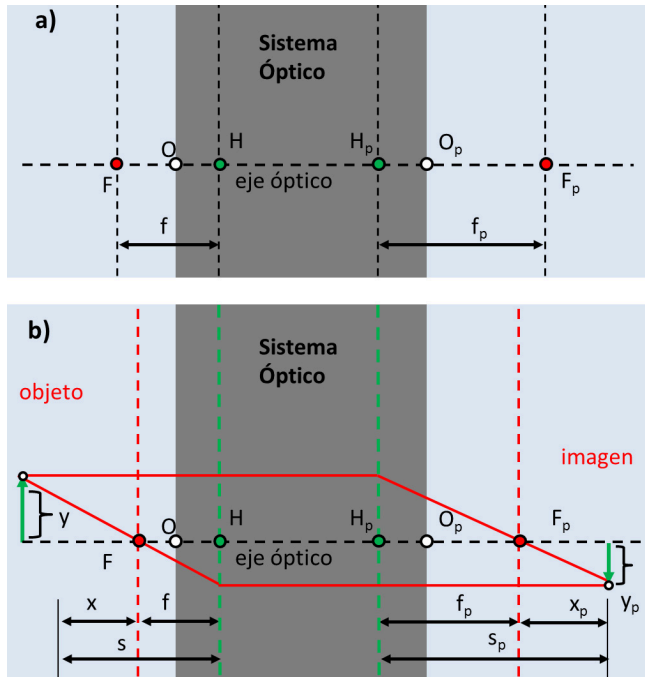


FIG. 10. Planos principales de un sistema óptico. a) Localización de planos focales F , F_p y principales H , H_p . Aquí los puntos O , O_p delimitan el sistema óptico. b) Determinación de la imagen utilizando los planos principales.

Trazado de rayos utilizando los planos principales: Los planos principales son útiles a la hora de realizar el trazado de rayos para la determinación de la imagen de un objeto. para ver esto, escribamos la matriz del sistema entre planos principales:

$$\begin{bmatrix} 1 & 0 \\ -1/f_p & f/f_p \end{bmatrix}$$

Entonces, ahora supongamos que un objeto se ubica a una distancia s del **plano principal objeto** y se quiere determinar en que posición s_p se forma la imagen, **respecto del plano principal imagen** (ver figura 10b).

Entonces si utilizamos la matriz del sistema entre puntos principales tenemos:

$$\begin{bmatrix} y_p \\ \theta_2 \end{bmatrix} = \begin{bmatrix} 1 & s_p \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -1/f_p & f/f_p \end{bmatrix} \begin{bmatrix} 1 & s \\ 0 & 1 \end{bmatrix} \begin{bmatrix} y \\ \theta_1 \end{bmatrix}$$

y realizando las multiplicaciones matriciales obtenemos

la expresión:

$$\begin{bmatrix} y_p \\ \theta_2 \end{bmatrix} = \begin{bmatrix} 1 - \frac{s_p}{f_p} & \frac{s(f_p - s_p) + s_p f}{f_p} \\ -1/f_p & \frac{-s + f}{f_p} \end{bmatrix} \begin{bmatrix} y \\ \theta_1 \end{bmatrix}$$

Como queremos que en s_p se forme la imagen del objeto, entonces debe ocurrir que el elemento de matriz

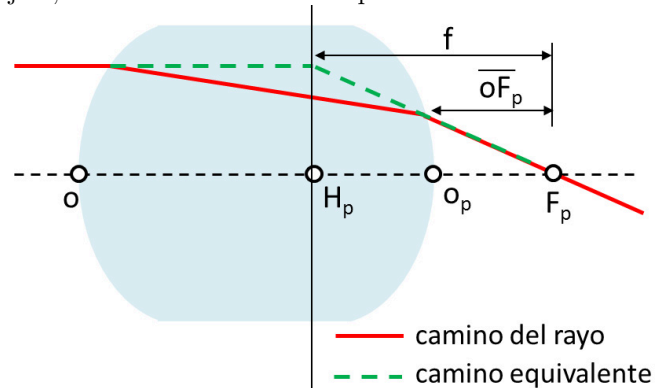


FIG. 11. Trazado de rayos equivalentes por medio de la utilización de los planos principales.

que conecta a y_p con θ_2 se anule, tal que:

$$\begin{aligned} \frac{s(f_p - s_p) + s_p f}{f_p} &= 0 \\ \Rightarrow \frac{s_p}{f_p} + \frac{s}{f} &= 1 \end{aligned} \quad (8)$$

La última de estas ecuaciones define la posición de la imagen s_p . Además la magnitud de la imagen vendrá dada por $y_2 = (1 - \frac{s_p}{f_p}) y_1$.

De esta forma, los planos conjugados proveen una gran simplificación del sistema óptico: **Podemos tratar un sistema óptico complejo como una lente delgada equivalente, siempre que consideremos las distancias focales respecto de los planos principales del sistema.** Gracias a esto, se puede realizar el trazado de rayos equivalentes utilizando los planos principales objeto e imagen (ver esquema en las figuras 10b y 11). Esto simplifica en gran medida el trazado de rayos en sistemas ópticos complejos, o de un gran número de elementos.