

El frente de onda y su representación con polinomios de Zernike

The Wavefront and its Representation with Zernike Polynomials

YOBANI MEJÍA BARBOSA*

RESUMEN

Se presenta una revisión del concepto *frente de onda* a partir del principio de Fermat, primero de forma general y luego para la pupila de salida de un sistema óptico formador de imagen. En la pupila de salida se definen el *frente de onda de referencia* y el *frente de onda real*, y la diferencia entre estos dos define la aberración de frente de onda. Luego, se representan matemáticamente el frente de onda y la aberración de frente de onda. En principio se emplea el sistema de coordenadas cartesianas para describir el polinomio de aberración en su forma general y discutir el significado de los primeros términos (aberraciones primarias), después se pasa al sistema de coordenadas polar. En dicha representación los términos de aberración tienen una forma matemática más simple y resulta ideal para describir el frente de onda sobre una pupila circular. Finalmente, adicionando algunos términos de compensación a las expresiones matemáticas en la representación polar con el fin de minimizar el efecto de las aberraciones en la imagen final, se llega a los polinomios de Zernike. Se explica el uso de los coeficientes de Zernike y se presenta un ejemplo del frente de onda obtenido con un aberrómetro experimental tipo Hartmann-Shack cuando se emplea un modelo físico del ojo humano.

Palabras clave: frente de onda, aberración de frente de onda, polinomios de Zernike

ABSTRACT

This paper is a review of the *wavefront* concept based on Fermat's principle, first in general terms and then for the exit pupil of an optical system that creates images. The *reference wavefront* and the *real wavefront* are defined in the exit pupil and the difference between them determines the wavefront aberrations. The wavefront and the wavefront aberration are then represented mathematically. The Cartesian coordinate system is used as at the beginning to describe the aberration polynomial in general terms and discuss the meaning of the first terms (primary aberrations), and then the polar coordinate system is used. The aberration terms in this representation have a simpler mathematical form and it is ideal to describe the wavefront of a circular pupil. Finally, after adding a few compensation terms to the mathematical expressions in the polar representation in order to minimize the effect of the aberrations in the final image, we reach the Zernike polynomials. The use of Zernike coefficients is explained, and an example is presented of the wavefront obtained with an experimental Hartmann-Shack aberrometer when using a physical model of the human eye.

Keywords: Wavefront, Wavefront Aberration, Zernike Polynomials.

*Ph.D. Ciencias-Óptica. Físico. Departamento de Física, Universidad Nacional de Colombia, Bogotá, Colombia.

INTRODUCCIÓN

En los sistemas ópticos formadores de imagen las aberraciones ópticas se suelen medir en el plano de la pupila de salida o en el plano de la imagen. En el primer caso se determina el frente de onda, mientras que en el segundo se determinan curvas (sagital y tangencial) de aberración de rayo. Ambas representaciones son equivalentes y mediante un sistema de ecuaciones podemos pasar de la una a la otra. Actualmente, en ciencias de la visión, el concepto de *frente de onda* se ha venido utilizando ampliamente, ya que este resulta útil para describir la calidad de imagen que el ojo puede formar de un objeto luminoso puntual.

El concepto de frente de onda se puede definir a partir de la óptica geométrica o a partir de la óptica física. En el primer caso se define mediante el principio de Fermat y en el segundo mediante la fase de la onda; en cualquier caso el frente de onda es una superficie y esta se puede representar matemáticamente como una función de tres variables. La forma explícita de la función depende del sistema coordenado empleado para su representación.

El sistema más común y con el que quizás estamos más familiarizados es el sistema coordenado cartesiano. Dicho sistema de coordenadas en el espacio se define de la siguiente manera: consideremos tres líneas rectas perpendiculares entre sí que se interceptan en un punto O (origen de coordenadas), sobre las cuales se define una escala métrica, como se muestra en la figura 1. Los ejes x y y constituyen el plano de las abscisas y el eje vertical z el eje de las ordenadas (altura o elevación). Si las direcciones positivas de los ejes son como se muestran en la figura 1, se dice que los ejes constituyen un sistema derecho. También se define el sistema de ejes izquierdo, donde los ejes x y y se intercambian, sin embargo, en este texto emplearemos solo el sistema derecho de ejes. Una superficie en el sistema cartesiano se representa matemáticamente como una función de dos variables independientes (x, y) y una variable dependiente z , lo que se suele denotar por $z = f(x, y)$.

Aquí f denota la relación o función que nos dice cómo obtenemos el valor de z a partir de los valores dados inicialmente a x y y , en otras palabras, f representa una ecuación matemática. La gráfica de la función se construye de la siguiente manera: dados los valores x_0 y y_0 para las variables independientes, se determina mediante la función el valor de la variable dependiente, a saber: $z_0 = f(x_0, y_0)$. Las coordenadas (x_0, y_0, z_0) definen el punto P de la superficie en la figura 1. Este procedimiento se realiza para todos los valores de x ($-\infty < x < \infty$) y y ($-\infty < y < \infty$); así, si la función es continua, la gráfica será una superficie que varía suavemente.

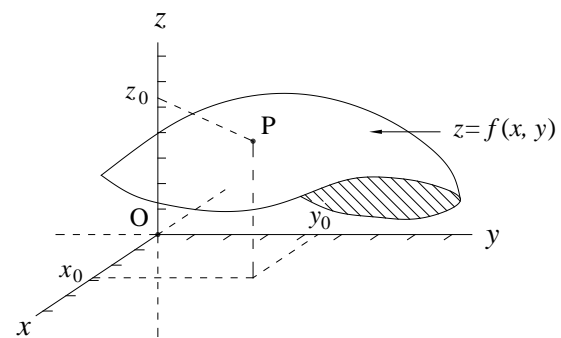


FIGURA 1. Sistema de coordenadas cartesianas

Nota. Un punto de una superficie se identifica mediante tres coordenadas lineales: x , y y z .

Otro sistema de coordenadas muy empleado es el sistema de coordenadas polares (figura 2). La ecuación de la superficie ahora tiene la forma $x = f(r, \theta)$, donde r es la distancia de la proyección de un punto de la superficie en el plano xy al origen de coordenadas O, y θ es ángulo que forma la línea que une el origen de coordenadas O con la proyección del punto en el plano xy con respecto a la horizontal (eje x). En la figura 2 las coordenadas del punto P son ahora (r_0, θ_0, z_0) . A r se le denomina la coordenada radial, a θ se le denomina coordenada polar y a z la elevación.

Las dos representaciones: $z = f(x, y)$ y $z = f(r, \theta)$, son equivalentes y mediante una transformación de coordenadas se obtienen relaciones matemáticas que permiten pasar de una representación a la otra.

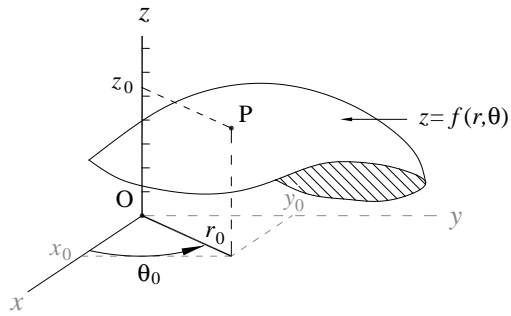


FIGURA 2. Sistema de coordenadas polares

Nota. Un punto de una superficie se identifica mediante una coordenada radial r , una coordenada angular θ y una coordenada lineal z .

PRINCIPIO DE FERMAT

El principio de Fermat se puede establecer de la siguiente manera: la luz, al ir desde un punto hasta otro, sigue la ruta que tiene la longitud de camino óptico menor (Hecht, 2000). La longitud de camino óptico o distancia óptica se define como:

(índice de refracción) \times (distancia física)

En un medio isotrópico y homogéneo el índice de refracción es constante y no depende de la dirección: por ejemplo, el vidrio o el plástico empleados usualmente en la fabricación de lentes oftálmicas son medios materiales isotrópicos y homogéneos. En un medio isotrópico no-homogéneo el índice de refracción varía de un punto a otro: por ejemplo, la lente del cristalino tiene un índice de refracción que varía con la distancia desde el centro hasta la periferia. Este tipo de variación de índice también se usa en la fabricación de fibras ópticas. Para lo que nos interesa en este texto, solo vamos a considerar medios isotrópicos (en un medio anisotrópico el índice de refracción varía con la dirección y el estado de polarización de la luz: por ejemplo, el cuarzo o la calcita son materiales anisotrópicos). En un medio homogéneo la ruta que sigue la luz entre dos puntos será una línea recta (la distancia más corta entre dos puntos en el espacio euclidiano es la longitud de la línea recta que los une); en cambio, si el medio es inhomogéneo la luz

seguirá una trayectoria diferente a la línea recta, en general, algún tipo de curva. La forma de la trayectoria depende de la manera como cambia el índice de refracción (Hecht, 2000). Lo anterior se ilustra en la figura 3: en (a) se muestra un medio homogéneo y la trayectoria recta que sigue la luz de A a B —nótese que en este caso solo hay una posible trayectoria—; en (b) el medio es inhomogéneo y la luz puede seguir varias trayectorias curvas para ir de A a B; y en (c) se muestra una aplicación usual: la formación de la imagen de A en B empleando una lente. En este caso se tienen dos medios homogéneos de diferente índice de refracción: aire-vidrio. Esta combinación de dos medios homogéneos permite que los diferentes rayos que salen de A sean enfocados en B.

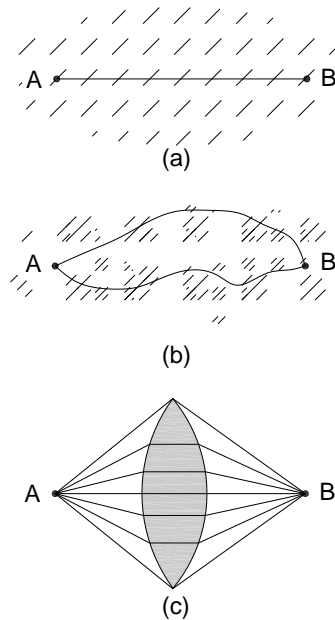


FIGURA 3. Diversas trayectorias de la luz en diferentes medios de acuerdo con el principio de Fermat

Nota: (a) medio homogéneo, (b) medio inhomogéneo, (c) dos medios homogéneos de diferente índice (aire-vidrio).

FRENTE DE ONDA

La trayectoria que sigue la luz de un punto a otro se denomina rayo de luz. Supongamos que tenemos una fuente luminosa puntual S en un

medio homogéneo de índice n ; la luz emitida por la fuente seguirá trayectorias radiales en todas las direcciones, es decir los rayos de luz en este caso son líneas rectas que salen de la fuente puntual. En la figura 4 se muestran algunos de estos rayos. Consideremos un rayo y localicemos un punto sobre ese rayo, de modo que la distancia entre S y el punto sea d ; así la distancia óptica entre S y el punto será nd . Si ahora tenemos en cuenta que de la fuente salen infinitos rayos radiales en todas las direcciones y localizamos sobre cada rayo un punto cuya distancia con respecto a S también sea d , podemos unir todos los puntos mediante una superficie. Esta superficie será una esfera y todos los rayos son perpendiculares a ella. A esta superficie que es ortogonal a los rayos de luz se le denomina frente de onda. En este ejemplo el frente de onda es una superficie esférica caracterizada por la distancia óptica nd .

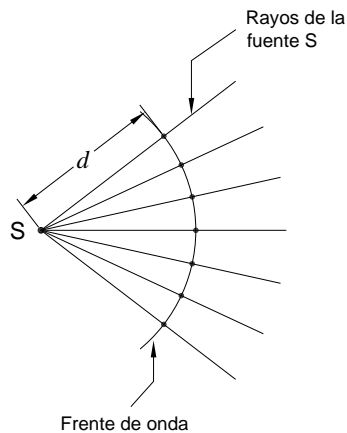


FIGURA 4. Rayos y frente de onda asociados a luz emitida por una fuente luminosa puntual en un medio homogéneo

Si la fuente luminosa está en un medio inhomogéneo, es decir, el índice de refracción n ya no es una constante, los rayos de luz emitidos por la fuente seguirán trayectorias curvas, tal y como se ilustra en la figura 5. Sin embargo, nuevamente podemos definir un frente de onda caracterizado por la distancia óptica nd , pero ahora la superficie que une todos los puntos de los rayos que están a la misma distancia óptica $nd = \text{constante}$ de S ya no será una esfera sino una superficie un poco más

compleja. Al igual que en el caso homogéneo, la superficie será ortogonal a los rayos. En este caso, aunque todos los rayos desde S hasta el frente de onda tienen la misma distancia óptica, no necesariamente tienen la misma distancia física, debido a las inhomogeneidades del medio.

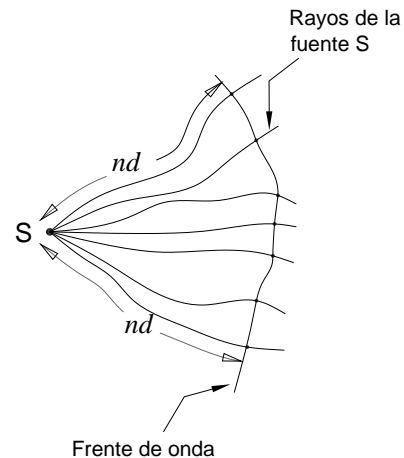


FIGURA 5. Rayos y frente de onda asociados a la luz emitida por una fuente luminosa puntual en un medio inhomogéneo

Tanto en el caso homogéneo como en el caso inhomogéneo podemos definir infinitos frentes de onda, pues basta con cambiar la distancia nd y obtendremos un nuevo frente de onda. En cualquier caso, un frente de onda es una superficie donde cualquiera de sus puntos tiene asociado el mismo valor de distancia óptica.

Entonces, si el cambio en la distancia nd representa la propagación de la luz, podemos referirnos a la propagación de un frente de onda, la cual implica, en general, un cambio en la forma de la superficie: en un medio inhomogéneo un frente de onda al propagarse cambiará continuamente su forma; en un medio homogéneo, la propagación de un frente de onda básicamente resulta en un cambio de escala, así un frente de onda esférico al propagarse solo cambiará el radio de curvatura, es decir, se mantiene la forma esférica. Como ya mencionamos anteriormente, una combinación de diferentes medios homogéneos también cambia la forma del frente de onda que se propaga debido

a la refracción de los rayos en las interfases que separan los distintos medios de acuerdo con la ley de Snell (Hecht, 2000). En la figura 6 se ilustra cómo cambia el frente de onda a medida que se propaga en diferentes medios.

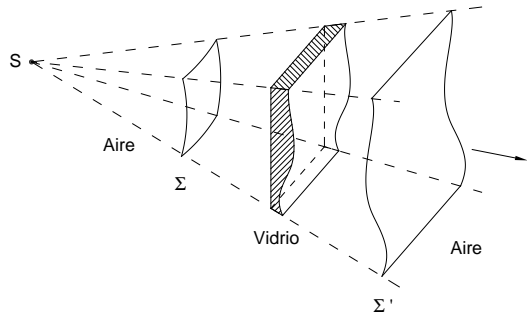


FIGURA 6. Propagación de un frente de onda

Nota. Inicialmente de la fuente puntual S emerge un frente de onda esférico Σ que va aumentando su radio de curvatura a medida que se aleja de la fuente, luego entra en un bloque de vidrio de forma irregular y al emerger de este el frente de onda cambia su forma siendo ahora un nuevo frente de onda Σ' que se sigue propagando más allá de la fuente S.

FORMACIÓN DE IMAGEN

Para formar la imagen de un objeto usualmente empleamos un sistema óptico conformado por lentes, espejos, etc. En un sistema óptico ideal como el mostrado en la figura 7, el frente de onda esférico divergente del objeto puntual es transformado en un frente de onda esférico convergente, de modo que la imagen de un punto es otro punto y la imagen de un objeto extendido es una réplica del objeto (salvo un cambio de escala). Lo anterior también implica que los medios donde se encuentran el objeto (espacio objeto) y la imagen (espacio imagen) son homogéneos.

En la práctica, aunque los espacios objeto e imagen sean homogéneos, la imagen de un punto objeto no es un punto sino más bien algo como una pequeña mancha, conocida como función de punto extendido o PSF (del inglés *Point Spread Function*) (Goodman, 1996). La forma de esta mancha depende de la geometría de los elementos que conforman el sistema óptico formador de imagen, de la homogeneidad de los medios refractores y de

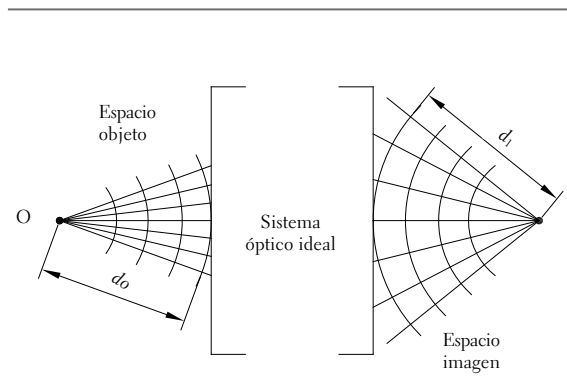


FIGURA 7. Formación de imagen de un punto objeto mediante un sistema óptico ideal

Nota. El frente de onda esférico divergente en el espacio objeto es transformado por el sistema óptico en un frente de onda esférico convergente en el espacio imagen.

la difracción de la luz en la pupila de salida del sistema óptico. Lo anterior significa que el frente de onda que emerge del sistema óptico correspondiente a un frente de onda esférico incidente ya no es esférico, sino una superficie distorsionada, tal y como se ilustra en la figura 8. El efecto sobre la imagen de un objeto extendido será un deterioro en la calidad de la imagen (Mahajan, 1998), tal como disminución de contraste y de resolución (agudeza visual).

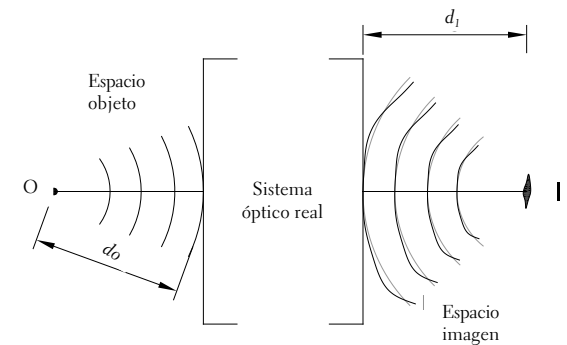


FIGURA 8. Formación de imagen de un punto objeto mediante un sistema óptico real

Nota. El frente de onda que emerge del sistema óptico está distorsionado en comparación con una superficie esférica.

Una manera práctica de modelar un sistema óptico es mediante sus pupilas de entrada y de salida. Las pupilas son las imágenes del diafragma de apertura del sistema óptico. Veamos estos conceptos con la

ayuda de la figura 9, en la que se muestra un sistema óptico formado por dos lentes simples y un diafragma circular en medio de las dos lentes. La tarea de este diafragma es controlar la cantidad de luz que entra al sistema para formar la imagen; si en un sistema no se encuentra el diafragma como un elemento separado, entonces el borde de alguna de las lentes será el elemento que controla la cantidad de luz. Al diafragma que controla la cantidad de luz se le denomina diafragma de apertura (DA). Suponiendo que se forma una imagen ideal, la imagen del punto objeto O será el punto imagen I. La pupila de entrada es la imagen del diafragma de apertura vista desde el punto objeto a través de los elementos ópticos que están antes del diafragma (en este caso la lente L1). La pupila de salida es la imagen del diafragma de apertura vista desde el punto imagen a través de los elementos ópticos que están después del diafragma (en este caso la lente L2). Entonces, en la práctica, desde el punto objeto el tamaño del cono de luz que entra al sistema óptico está determinado por el diámetro de la pupila de entrada, y el tamaño del cono de luz que sale del sistema óptico para formar la imagen está determinado por el diámetro de la pupila de salida, como se muestra en la figura 10; en otras palabras, las pupilas determinan la cantidad de luz que efectivamente entra o sale del sistema óptico.

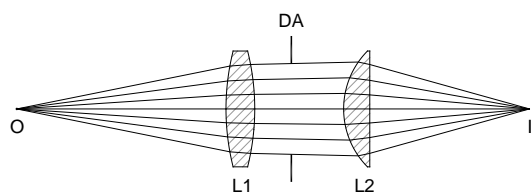


FIGURA 9. Sistema de dos lentes que forma una imagen perfecta

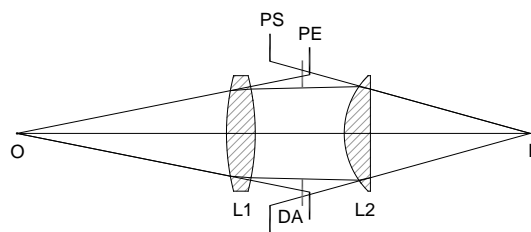


FIGURA 10. Pupilas de entrada (PE) y salida (PS) y conos de luz en el sistema de la figura 9

Las pupilas nos permiten modelar cualquier sistema óptico formador de imagen, tal como se muestra en la figura 11. Este modelo generalizado supone que: 1) entre las pupilas de entrada y de salida el paso de la luz ocurre de acuerdo con las leyes de la óptica geométrica; 2) la pupila de entrada no distorsiona el frente de onda esférico a su paso; y 3) la distorsión del frente de onda en el espacio imagen se genera en la pupila de salida. En otras palabras, este modelo supone que los defectos ópticos del sistema se condensan en la pupila de salida. Lo anterior se representa matemáticamente mediante una función de transmitancia en el plano de la pupila de salida, entonces, el frente de onda que emerge de la pupila de salida es el resultado de multiplicar el frente de onda esférico que llega a la pupila de salida por la función de transmitancia que contiene los defectos ópticos del sistema (Goodman, 1996). Con este modelo el estudio de la calidad de imagen generada por un sistema óptico se limita a medir el frente de onda en la pupila de salida. Una vez que se conoce el frente de onda en la pupila de salida se puede evaluar la PSF u otras funciones comúnmente empleadas para caracterizar un sistema óptico. Un ejemplo de la aplicación de este modelo generalizado en ciencias de la visión lo tenemos en los aberrómetros oculares (Thibos, 2000). Estos instrumentos primero generan una fuente puntual luminosa en la retina y luego miden (en forma indirecta) en la pupila de salida del ojo el frente de onda emergente. En este caso, se tiene que la luz emerge de la retina y sale del ojo atravesando los medios transparentes del ojo, por lo que la pupila de salida será la imagen del iris que un observador externo ve. Los resultados suelen ser presentados mediante un histograma de coeficientes de polinomios de Zernike.

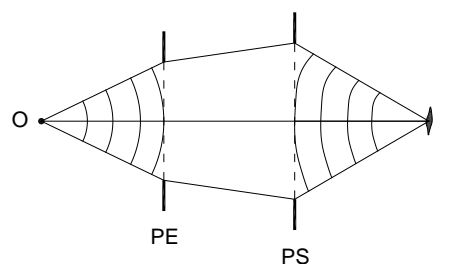


FIGURA 11. Modelo generalizado de un sistema óptico a partir de las pupilas de entrada y de salida

Las pupilas también nos permiten definir un par de rayos de gran relevancia en óptica geométrica: el rayo principal y el rayo marginal. El rayo principal sale de un punto objeto fuera de eje, entra al sistema óptico dirigido al centro de la pupila de entrada, pasa por el centro del diafragma de apertura y llega al punto imagen correspondiente en una dirección tal como si viniera del centro de la pupila de salida. El rayo marginal sale de un punto objeto en eje, entra al sistema óptico dirigido al borde de la pupila de entrada, pasa por el borde del diafragma de apertura y llega al punto imagen correspondiente en una dirección tal como si viniera del borde de la pupila de salida. En un sistema óptico, un plano que contenga al eje óptico y al rayo principal se denomina plano meridional o tangencial, y el plano ortogonal al plano meridional se denomina plano sagital.

REPRESENTACIÓN MATEMÁTICA DEL FRENTE DE ONDA

De acuerdo con lo dicho en la sección anterior, para caracterizar el sistema óptico (calidad de imagen) se requiere conocer el frente de onda en la pupila de salida. Primero vamos a suponer que el sistema forma una imagen perfecta, es decir, la imagen será un punto. Entonces, un observador en el punto imagen verá que de la pupila de salida sale un frente de onda esférico que converge en el punto imagen; el radio del frente de onda esférico en la pupila lo denotaremos por R . La cuestión ahora es cómo describimos matemáticamente este frente de onda. Comencemos por introducir un sistema de coordenadas cartesianas (x, y, z) cuyo origen esté en el plano de la pupila de salida y cuyo eje z coincida con el eje óptico del sistema, como se ilustra en la figura 12: x y y denotan las coordenadas transversales en el plano de la pupila de salida, mientras que x_i y y_i denotan las coordenadas transversales en el plano imagen gaussiano (donde se forma la imagen perfecta). La porción de la esfera que describe el frente de onda en la pupila de salida estará representada por la siguiente ecuación:

$$z = R - \sqrt{R^2 - (x^2 + y^2)} \quad (1)$$

tal que $x \leq r_0$ y $y \leq r_0$, siendo r_0 el radio de la pupila de salida. Esta es la ecuación de una esfera de radio R con centro en $(0, 0, R)$. Note que si $x = 0$ y $y = 0$ (vértice de la esfera), entonces $z = 0$. Las coordenadas del punto imagen serán las mismas que las del centro de la esfera. Como esta esfera representa el frente de onda en la pupila de salida cuando la formación de la imagen es perfecta, a esta esfera la denominamos esfera de referencia.

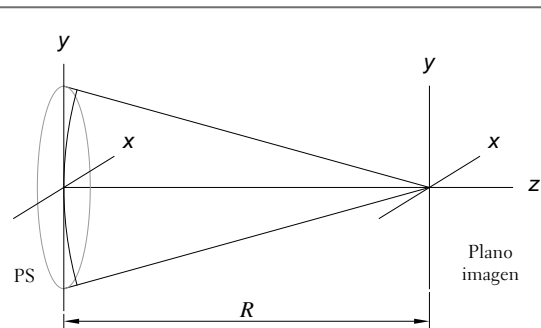


FIGURA 12. Coordenadas de la pupila de salida (PS) y del plano imagen gaussiano

Nota. La esfera de radio R en la pupila de salida se denomina esfera de referencia.

Volvamos al sistema óptico de la figura 9, pero ahora consideremos la formación de la imagen real. En este caso, los rayos en el espacio imagen no se cruzan todos en el mismo punto, así que la imagen será una mancha extendida, como se muestra en la figura 13, y en consecuencia el frente de onda en la pupila de salida no puede ser esférico. La forma del frente de onda en este caso será una superficie más compleja en comparación con la forma esférica, y su forma explícita dependerá de cada sistema óptico en particular.

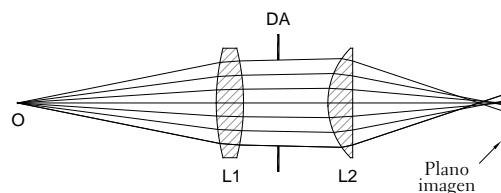


FIGURA 13. Formación real de la imagen del punto objeto

Nota. La imagen del punto objeto es una mancha extendida.

Vamos a denotar a la esfera de referencia como $z = S(x, y)$, y al frente de onda real como $z = S'(x, y)$. Decimos que $z = S(x, y)$ es la ecuación implícita de la esfera de referencia, mientras que la ecuación 1 es la ecuación explícita de la esfera de referencia. En forma análoga, $z = S'(x, y)$ es la ecuación implícita del frente de onda real, pero por ahora no tenemos una ecuación explícita de esta superficie como en el caso de la esfera de referencia. En la práctica, la forma del frente de onda real se analiza con respecto a la esfera de referencia, calculando la diferencia entre las dos superficies; es lo que se denomina aberración de frente de onda. En general, la aberración de frente de onda es pequeña en comparación con el radio de la esfera de referencia, por lo que la forma del frente de onda real se puede calcular como

$$S'(x, y) = S(x, y) + W(x, y) \quad (2)$$

siendo $W(x, y)$ la aberración de frente de onda.

ABERRACIÓN DE FRENTE DE ONDA

Según lo dicho en la sección anterior, medir la aberración de frente de onda es equivalente a medir la forma del frente de onda real; por otra parte, usualmente, lo que medimos experimentalmente es la aberración de frente de onda. En la figura 14 se representa el frente de onda real junto con la esfera de referencia. El rayo 1 es el rayo correspondiente a la esfera de referencia en el punto P, el cual cruza al frente de onda real en el punto Q; el rayo 2 es el rayo correspondiente al frente de onda real en el punto Q; la aberración de frente de onda, que denotamos por W , se define como la distancia óptica entre P y Q, a lo largo del rayo 1 (Rayces, 1964). Otros autores (Mahajan, 1998) definen la aberración de frente de onda como la distancia óptica entre los frentes de onda a lo largo del rayo 2. En la práctica, el ángulo entre los rayos 1 y 2 es pequeño y se llega al mismo resultado, dado por las ecuaciones:

$$\Delta x_i = -\frac{R}{n'} \frac{\partial W}{\partial x} \quad (3a)$$

$$\Delta y_i = -\frac{R}{n'} \frac{\partial W}{\partial y} \quad (3b)$$

donde Δx_i y Δy_i miden la diferencia entre las coordenadas x_i y y_i de la intersección del rayo 1 con el plano imagen (punto imagen ideal) y de la intersección del rayo 2 con el plano imagen. En otras palabras, Δx_i y Δy_i son las desviaciones de los rayos reales con respecto a los rayos ideales medidas en el plano imagen; a estas desviaciones se les denomina aberraciones de rayo: $\partial W / \partial x$ y $\partial W / \partial y$ representan las derivadas de la aberración de frente de onda en las direcciones x y y . Geométricamente, la derivada es la pendiente. El índice de refracción en el espacio imagen se denota por n' .

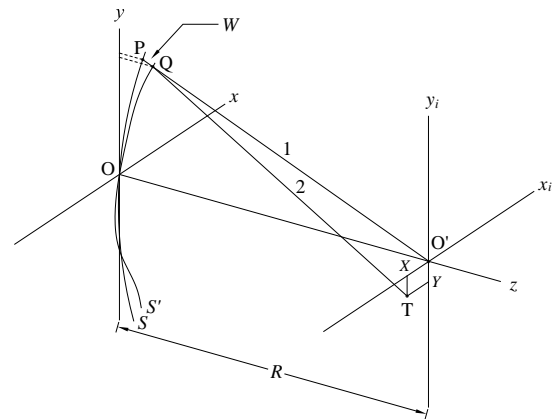


FIGURA 14. La longitud entre los puntos Q y P a lo largo del rayo 1 es la distorsión del frente de onda S' en comparación con la esfera de referencia S

Nota. A esta longitud se le denomina aberración de frente de onda.

En los aberrómetros tipo Hartmann-Shack (Thibos, 2000) se forma la imagen de la pupila de salida del ojo (suponiendo que la luz sale de la retina) en el plano del arreglo de microlentes. En consecuencia, el frente de onda que incide sobre el arreglo de microlentes es el conjugado (la imagen) del frente de onda en la pupila de salida del ojo. En este caso, el frente de onda de referencia en el arreglo de microlentes es plano ($R \rightarrow \infty$), es decir $S(x, y) = 0$, de modo que en esta configuración la aberración de frente de onda es idéntica a la forma del frente de onda real. Las ecuaciones que relacionan las aberraciones de rayo Δx_i y Δy_i con las pendientes

del frente de onda tienen la misma forma que las ecuaciones 3a y 3b, pero en lugar de R se tiene la distancia focal f de las microlentes del arreglo y $n' = 1$ (aire), de modo que en un aberrómetro tipo Hartmann-Shack, las ecuaciones que relacionan las aberraciones de rayo con las pendientes son:

$$\Delta x_i = -f \frac{\partial W}{\partial x} \quad (4a)$$

$$\Delta y_i = -f \frac{\partial W}{\partial y} \quad (4b)$$

Cada microlente del arreglo enfoca una pequeña región del frente de onda en el plano focal de las microlentes; el tamaño de la región enfocada será del mismo tamaño de la microlente y el resultado final es un patrón de puntos, conocido como patrón de Hartmann. Por lo tanto, con un aberrómetro tipo Hartmann-Shack, obtenemos directamente las aberraciones de rayo Δx_i y Δy_i , restando las posiciones de los puntos del patrón de Hartmann generado por el ojo bajo análisis de las posiciones de los puntos del patrón de Hartmann, generado a su vez por un frente de onda plano. Usualmente se dice que el aberrómetro mide las pendientes del frente de onda, pero estrictamente lo que mide son las aberraciones de rayo y a partir de las ecuaciones 4a y 4b se deducen las pendientes locales en las direcciones transversales x y y . Por último, para obtener $W(x, y)$ de las ecuaciones 4a y 4b se requiere integrar, por lo que la precisión en la reconstrucción del frente de onda también depende de los algoritmos empleados para integrar dichas ecuaciones.

Desde el punto de vista matemático, la forma explícita de la aberración de frente de onda es una ecuación que puede contener muchos términos (en principio infinitos términos); esta en el sistema de coordenadas cartesiano tiene la forma general:

$$\begin{aligned} W(x, y) = & a_{00} + a_{10} x + a_{01} y \\ & + a_{11} xy + a_{20} x^2 + a_{02} y^2 \\ & + a_{21} x^2 y + a_{12} xy^2 + a_{30} x^3 + a_{03} y^3 \\ & + \dots a_{mn} x^m y^n + \dots \end{aligned} \quad (5)$$

donde los coeficientes a_{mn} , con $m = 1, 2, 3 \dots$ y $n = 1, 2, 3 \dots$, son números reales, cuyo valor determina la relevancia que cada uno de los monomios $x^m y^n$ tiene dentro de la suma. Cada monomio en sí representa una superficie, por ejemplo: $f(x, y) = x$ representa un plano inclinado en la dirección x (*tilt about y*), $f(x, y) = y$ representa un plano inclinado en la dirección y (*tilt about x*). El ángulo de inclinación dependerá del valor del coeficiente correspondiente, así si $a_{10} = 1$, entonces tendremos un plano inclinado en la dirección x que forma un ángulo de 45° . El término a_{00} representa un plano horizontal (paralelo al plano definido por los ejes x y y) de altura igual a $z = a_{00}$ (pistón). El coeficiente a_{00} permite desplazar axialmente el resto de los términos de la ecuación 5 con respecto al plano formado por los ejes x y y . Por esta razón, al término a_{00} se le denomina pistón, y en sí no representa ningún tipo de aberración. Estos tres monomios, en la primera línea de la ecuación 5, se denominan monomios de orden 1 y a las figuras correspondientes superficies de orden 1. Algunos autores se refieren a la dirección de la inclinación de los planos de manera diferente, considerando el eje que se debe girar para obtener la inclinación, por ejemplo, el plano inclinado $f(x, y) = x$ se obtiene rotando un plano horizontal alrededor del eje y (un ángulo de 45°). Por eso se refieren a este plano como una inclinación alrededor del eje y (*tilt about y*); sin embargo, en este texto nos parece más claro decir que este plano está inclinado en la dirección x . Por otra parte, el monomio x^2 representa un paraboloide cilíndrico, es decir, el perfil de la superficie en cualquier plano paralelo al plano formado por los ejes x y z es una parábola; análogamente, el monomio y^2 representa un paraboloide cilíndrico, pero ahora el perfil parabólico de la superficie ocurre en cualquier plano paralelo al plano formado por los ejes y y z . Por último, el monomio xy representa una superficie un poco más compleja, pues en las direcciones diagonales el perfil de la superficie también son parábolas, en un caso orientada hacia arriba y en el otro orientada hacia abajo, mientras que el perfil en las direcciones x o y son líneas rectas; a esta superficie, debido a su forma, se le llama silla de

montar. En la figura 15 se muestran las gráficas de las tres superficies descritas por los monomios de la línea 2 de la ecuación 5 para una región cuadrada de 8×8 mm. A estos tres monomios se les denomina monomios de orden 2 y a las figuras correspondientes superficies de orden 2.

El orden del monomio es igual a la suma de los índices m y n , así, en la ecuación 5 los monomios están ordenados por líneas según su orden, por ejemplo, en la línea 3, los cuatro términos tienen en común que la suma de sus índices es 3. En conclusión, el orden del monomio $x^m y^n$ es $m + n$. A medida que avancemos en el orden del monomio, la superficie descrita es cada vez más compleja.

La suma de dos o más monomios se denomina polinomio, y el orden del polinomio estará determinado por el del monomio de más alto orden. El polinomio dado por la ecuación 5 se denomina polinomio de Taylor, o simplemente polinomio en coordenadas cartesianas.

El polinomio de orden 2, $f(x, y) = a_{11} xy + a_{20} x^2 + a_{02} y^2$, resulta de gran importancia en ciencias de la visión, ya que con este polinomio se representan matemáticamente el cilindro o astigmatismo

y la esfera media o el defoco, como veremos más adelante.

Geoméricamente diremos entonces que un polinomio representa una superficie. La forma de esta superficie puede ser simple o compleja, dependiendo del número de monomios que componen el polinomio. A su vez, podemos decir que la superficie que describe el polinomio es la suma de un conjunto de superficies más simples, por ejemplo, planos inclinados, paraboloides, etc. Otro hecho importante es que diferentes superficies pueden ser generadas por polinomios conformados por el mismo tipo de monomios; lo que hace la diferencia aquí es el valor de los coeficientes en cada caso. Por esta razón, es común describir la superficie simplemente por sus coeficientes (ordenados en algún tipo de histograma) en lugar de escribir la ecuación. Por ejemplo, la figura 16 muestra la representación de dos superficies de orden 2 (parte inferior) mediante el valor de los coeficientes de sus ecuaciones. El valor de los coeficientes se dibuja en un histograma (parte superior). Los coeficientes de los polinomios que describen estas superficies son: en el primer caso (figura 16 [a]), $a_{00} = 0$, $a_{10} = 1,2$, $a_{01} = -0,5$, $a_{11} = 0,1$, $a_{20} = 0,7$, $a_{02} = -0,7$; en el segundo caso (figura 16 [b]),

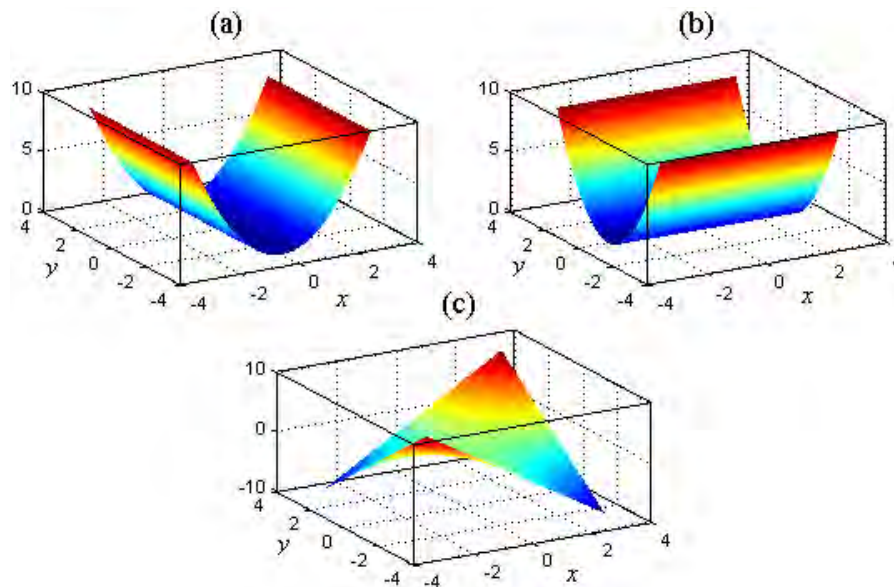


FIGURA 15. Superficies de segundo orden

Nota. (a) paraboloides cilíndrico en x ; (b) paraboloides cilíndrico en y ; (c) silla de montar

$a_{00} = 0$, $a_{10} = 0,1$, $a_{01} = -0,1$, $a_{11} = -0,1$, $a_{20} = 1,0$, $a_{02} = 1,0$. En ambos casos no hay término pistón (el primer coeficiente es cero), y las superficies son el resultado de sumar planos inclinados en las direcciones x y y , la silla de montar y los paraboloides cilíndricos en las direcciones x y y . Sin embargo, las dos superficies son diferentes debido al peso relativo de los coeficientes en cada una de las superficies. El histograma de la figura 16 (a) muestra que sobresale la contribución del plano inclinado en la dirección x , hay una pequeña contribución de la silla de montar dada por el coeficiente a_{11} , y los dos paraboloides cilíndricos tienen un peso medio y además signo contrario, lo que genera una nueva silla de montar, así que el aspecto general de la superficie es una silla de montar inclinada positivamente en la dirección x . El histograma de la figura 16 (b) muestra que las componentes de inclinación y de silla de montar son despreciables frente a los paraboloides cilíndricos, que en este caso tienen el mismo signo dando origen a un paraboloides de revolución, así que el aspecto general de esta superficie está determinado por el paraboloides de revolución.

ABERRACIONES DE ORDEN 2

En esta sección vamos a mostrar cómo se obtiene la expresión usual de aberración de frente de onda para el desenfoque y el astigmatismo en la aproximación de orden 2. Estas aberraciones permiten describir la miopía, la hipermetropía y el astigmatismo ocular.

De foco (aberración de desenfoque)

Esta es la aberración más simple y tiene su origen en la incapacidad del sistema óptico de enfocar la imagen en el plano imagen gaussiano. En la figura 17 se ilustra una configuración posible.

Supongamos que el sistema óptico es capaz de formar una imagen ideal, pero el punto imagen no se localiza en el plano imagen gaussiano correspondiente a la esfera de referencia, sino en un plano anterior. Esto puede ocurrir porque se ha desplazado axialmente alguna de las lentes del sistema o porque simplemente el plano imagen se ha colocado en una posición diferente. En cualquier

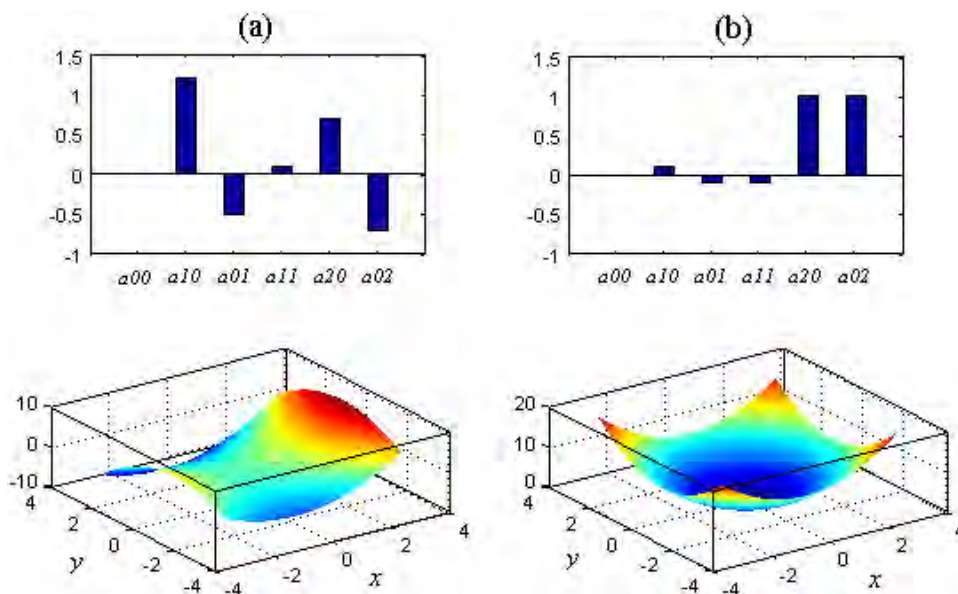


FIGURA 16. Representación de una superficie mediante un histograma de coeficientes

Nota. Las dos superficies están compuestas por el mismo tipo de monomios pero difieren en el valor de sus coeficientes.

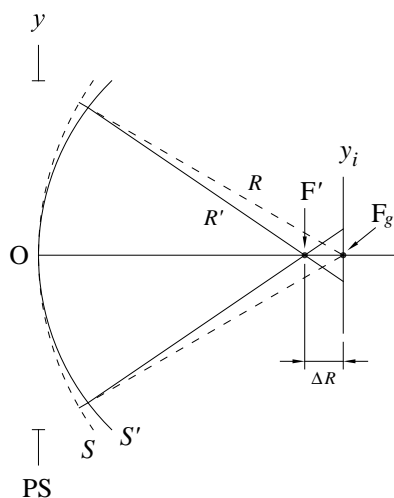


FIGURA 17. Frente de onda esférico desenfocado de radio R' y frente de onda de referencia R

caso, el radio de la esfera de referencia será igual a la distancia que separa el plano de la pupila de salida y el plano imagen. El frente de onda de referencia se enfocará en el punto F_g . Por otra parte, como el sistema puede formar una imagen ideal, el frente de onda real será una esfera de radio igual a R' , y se enfocará en el punto F' . Decimos entonces que la imagen está desenfocada y la longitud axial de desenfoco es igual a $\Delta R = R' - R$. En consecuencia, en el plano imagen gaussiano se observará una mancha circular cuyo diámetro es proporcional a la longitud de desenfoco (Atchison y Smith, 2000). Ahora bien, para calcular la aberración de frente de onda para el desenfoco debemos restar las dos esferas, de radios R' y R . De acuerdo con la definición de la aberración de frente de onda, esta se evalúa a lo largo del rayo 1 (figura 14), lo que puede resultar muy complejo aunque las superficies sean relativamente simples. En lugar de hacer la resta a lo largo del rayo 1, esta se hace a lo largo de una línea paralela al eje z (eje óptico) que pase por el punto P (tal y como lo indica la ecuación 2). Esta aproximación en general se satisface muy bien ya que el $F/\#$ efectivo del sistema (cociente entre el radio de la esfera de referencia y el diámetro de la pupila de salida) es mucho mayor que 1; en el caso del aberrómetro tipo Hartmann-Shack, esto siempre se cumple ya

que la dirección del rayo 1 (figura 14) es paralela al eje óptico.

A la distancia entre un punto del frente de onda y el plano de la pupila de salida se le denomina sagita. De acuerdo con esta definición, la altura z dada en la ecuación 1 en sí es la sagita, debido a que hemos colocado nuestro origen de coordenadas en el vértice de la esfera. Entonces, la aberración de defoco se define matemáticamente como la diferencia de sagitas entre el frente de onda y la esfera de referencia en aproximación de segundo orden. Para aclarar esto, vamos a escribir la diferencia de las sagitas de las dos esferas:

$$W(x, y) = \left(R' - \sqrt{R'^2 - (x^2 + y^2)} \right) - \left(R - \sqrt{R^2 - (x^2 + y^2)} \right) \quad (6)$$

Como ya lo mencionamos, esta ecuación es más bien compleja, ya que no podemos reducirla a una forma más simple; pero si aproximamos cada uno de los radicales por sus términos de segundo orden, es decir, si cambiamos las esferas por superficies parabólicas de revolución, obtenemos:

$$W(x, y) = \frac{x^2 + y^2}{2R'} - \frac{x^2 + y^2}{2R} = -\frac{(R' - R)}{2RR'}(x^2 + y^2) \quad (7)$$

lo que es mucho más simple de entender (los lectores interesados en los detalles de estos cálculos pueden consultar un texto de cálculo que explique la expansión en series de Taylor de una función. En este texto, en aras de la simplicidad, no mostraremos la manera como se obtiene el término de segundo orden de una raíz cuadrada). Note que la ecuación 7 es un polinomio de orden 2, y si se compara con la línea 2 de la ecuación 5, lo que tenemos es la suma de dos monomios con coeficientes a_{20} y a_{02} , los cuales resultan ser iguales, a saber: $a_{20} = a_{02} = -(R' - R)/2RR'$. Si denotamos simplemente estos coeficientes por a_d , podemos escribir la ecuación para el defoco como:

$$W(x, y) = a_d(x^2 + y^2) \quad (8)$$

donde el valor de $a_d = -\Delta R / 2R^2$. En este último resultado se ha aproximado $R'R$ por R^2 , ya que en la práctica el defoco longitudinal es mucho menor que el radio de la esfera de referencia. Así, el coeficiente de defoco queda determinado por el radio de la esfera de referencia y el defoco longitudinal. La forma del defoco es un paraboloide de revolución, como se ilustra en la figura 19a.

En óptica geométrica, así como el pistón, el defoco tampoco se considera una aberración, ya que este se puede corregir simplemente desplazando el plano imagen gaussiano a la posición donde realmente se forma la imagen. Sin embargo, en ciencias de la visión, el defoco sí se considera como un defecto visual. El término 'defoco' permite describir la miopía ($\Delta R < 0$, $a_d > 0$) o la hipermetropía ($\Delta R > 0$, $a_d < 0$), por lo que a estas también se les denomina aberraciones oculares de segundo orden o primarias.

ASTIGMATISMO

El astigmatismo resulta de una asimetría en el frente de onda en forma de superficie tórica. Esta asimetría puede ocurrir en un sistema óptico con simetría rotacional cuando el punto objeto está fuera de eje o para un punto objeto en eje cuando en el sistema óptico alguna de las superficies de las lentes es de forma tórica. En una superficie tórica los meridianos principales son círculos. Supongamos que el meridiano S_1 en la dirección y ($90^\circ - 270^\circ$), de ahora en adelante denominado meridiano tangencial, tiene un radio de curvatura igual a R_1 , y que el meridiano S_2 en la dirección x ($0^\circ - 180^\circ$), de ahora en adelante denominado meridiano sagital, tiene un radio de curvatura R_2 . Los rayos correspondientes a estos dos meridianos se denominan rayos tangenciales y rayos sagitales, y los planos que contienen estos rayos se denominan planos tangencial y sagital, respectivamente. Los rayos tangenciales se enfocan en el punto F_T (centro de curvatura de S_1), y los rayos sagitales se enfocan en el punto F_S (centro de curvatura de S_2), como se muestra en la figura 18.

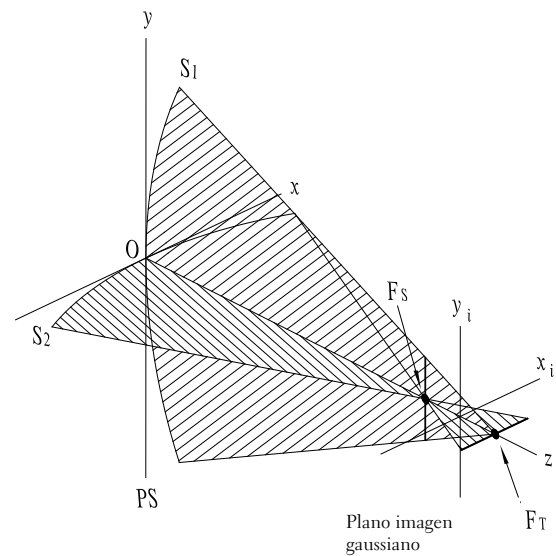


FIGURA 18. Astigmatismo

Nota. Los rayos tangenciales se enfocan en el punto F_T (centro de curvatura de S_1) y los rayos sagitales se enfocan en el punto F_S (centro de curvatura de S_2). La diferencia entre F_T y F_S se denomina astigmatismo longitudinal.

Para determinar en este caso la aberración de frente de onda, primero debemos fijar la esfera de referencia. Supongamos que el plano imagen gaussiano se encuentra localizado a la mitad de la distancia entre los puntos focales F_T y F_S , como se indica en la figura 18. Así, el radio de la esfera de referencia será $R = (R_1 + R_2)/2$. Con esto ya podemos hacer la resta entre el frente de onda astigmático y la esfera de referencia. Al igual que en la aberración de defoco, en lugar de tomar la forma exacta del frente de onda tórico para hacer la resta, tomaremos la aproximación de orden 2, de manera que el frente de onda con aberración de astigmatismo se puede escribir como:

$$S'(x, y) = \frac{x^2}{2R_2} + \frac{y^2}{2R_1} \quad (9)$$

esto es, la suma de dos paraboloides cilíndricos. Realizando la resta de la ecuación 9 con la esfera de referencia en la aproximación de orden 2 se llega a:

$$W(x, y) = -x^2 \left(\frac{R_2 - R}{2R_2 R} \right) - y^2 \left(\frac{R_1 - R}{2R_1 R} \right) \quad (10)$$

Identificando $a_{20} = -(R_2 - R)/2R_2R$ y $a_{02} = -(R_1 - R)/2R_1R$, y teniendo en cuenta que en la práctica R_2 y R_1 tienen valores muy cercanos a R , de modo que podemos aproximar $R_2R \approx R^2$ y $R_1R \approx R^2$, y ya que R es el valor promedio de los radios R_1 y R_2 , se llega a $a_{20} = -a_{02}$. Si ponemos $a_a = a_{20}$, la aberración de frente de onda para el astigmatismo a segundo orden se puede escribir simplemente por:

$$W(x, y) = a_a(x^2 - y^2) \quad (11)$$

De acuerdo con la figura 18, se tiene que $R < R_1$ y $R > R_2$, es decir $a_a > 0$, por lo que la forma de la superficie que representa la aberración de frente de onda para el astigmatismo es una silla de montar como la mostrada en la figura 19 (b). La diferencia de esta silla de montar con la mostrada en la figura 15 (c) es la orientación. Se dice que el astigmatismo descrito por la figura 19 (b) está orientado a 0° y el astigmatismo descrito por la figura 15 (c) está orientado a 45° . Una combinación de estas dos figuras permite describir la aberración de astigmatismo en cualquier orientación.

Considerando la figura 18 en la aproximación de orden 2, se puede mostrar que todos los rayos del frente de onda astigmático se enfocan en una línea centrada en el punto focal FS, luego el manojito de rayos se abre y en el plano imagen gaussiano estos están contenidos en un círculo, y nuevamente se enfocan en otra línea centrada en el punto focal FT (Mahajan, 1998). Estas líneas focales se denominan líneas focales astigmáticas y la distancia entre ellas es el astigmatismo longitudinal. La línea focal astigmática centrada en FT está contenida en el plano sagital y también se denomina línea focal tangencial. La línea focal astigmática centrada en FS está contenida en el plano tangencial y también se denomina línea focal sagital (las líneas focales astigmáticas son ortogonales al rayo principal del sistema óptico. En la figura 18 el eje z desempeña el papel del rayo principal). Note que el plano imagen gaussiano tal y como se ha definido es el plano de mejor enfoque, ya que allí la imagen de un punto objeto es un pequeño círculo denominado círculo de menor confusión.

ABERRACIONES DE ORDEN SUPERIOR

Los términos de orden 3 o mayores en la ecuación 5 permiten describir otro tipo de aberraciones de frente de onda denominadas de orden superior. Mediante un procedimiento análogo al mostrado en la sección de aberraciones de segundo orden se pueden deducir las expresiones matemáticas correspondientes. En esta sección solo haremos una breve descripción de dos de las aberraciones de orden superior más conocidas: la coma (orden 3) y la esférica (orden 4).

La aberración de coma tiene su origen cuando el aumento de un sistema óptico depende de la altura de los rayos en la pupila de salida. Lo anterior ocurre en un sistema con simetría rotacional cuando el punto objeto está fuera de eje, o en un punto objeto en eje cuando en el sistema óptico alguna de las lentes está descentrada o tiene una forma prismática. En ambos casos las distancias focales efectivas serán diferentes para los rayos que cruzan fuera de eje la pupila de salida (Hecht, 2000). El resultado final es que la imagen del punto objeto es una mancha con forma de cometa.

La aberración esférica ocurre para los rayos no paraxiales en superficies reflectoras o refractoras esféricas (de ahí su nombre). A medida que el rayo aumenta su altura en la pupila de salida, el rayo enfocará en un punto cada vez más alejado del punto imagen paraxial (gaussiano). La aberración esférica está presente para puntos objeto tanto en eje como fuera de eje y es la única aberración presente para puntos objeto en eje en un sistema óptico (de superficies esféricas) con simetría rotacional. La figura 13 ilustra muy bien esta aberración, de donde se puede deducir que la imagen de un punto objeto será una mancha circular.

La aberración de frente de onda de coma se puede describir mediante:

$$W(x, y) = a_c x(x^2 + y^2) \quad (12)$$

y la aberración de frente de onda esférica se puede describir mediante:

$$W(x, y) = a_s(x^2 + y^2)^2 \quad (13)$$

En la figura 19 (c) se muestra la aberración de coma descrita por la ecuación 12 y en la figura 19 (d) se muestra la aberración esférica descrita por la ecuación 13. La aberración de coma también se puede describir mediante $W(x, y) = a_c y(x^2 + y^2)$. La diferencia con la ecuación 12 está en que ahora la superficie estará inclinada en la dirección y . Al igual que en el astigmatismo, una combinación de las dos superficies comáticas permite representar la aberración de coma con cualquier orientación. Se suele asimilar la forma de la superficie correspondiente a la aberración de coma a una cuchara. Por otra parte, la aberración esférica tiene una forma de tazón: la base es más pronunciada que en el caso de la aberración de defoco, y luego aumenta muy rápido hasta la periferia. Así que para rayos cerca del eje la aberración esférica es despreciable, por lo que una manera de reducir esta aberración es disminuyendo el diámetro de la pupila de salida. Las aberraciones de astigmatismo, coma y esférica que hemos visto en esta sección, junto con la aberración de distorsión (que no hemos discutido) y curvatura de campo (relacionada con el defoco), se denominan aberraciones primarias o de Seidel (Hecht, 2000). La combinación de otros monomios de Taylor describen aberraciones que usualmente se clasifican como secundarias, terciarias, etc. (Mahajan, 1998).

POLINOMIOS DE ZERNIKE

Ya hemos visto que la aberración de frente de onda se puede representar matemáticamente como una suma de monomios de Taylor (coordenadas cartesianas). La combinación apropiada de monomios permite obtener las diferentes aberraciones ópticas. Otra posibilidad para representar las aberraciones ópticas es mediante polinomios en coordenadas polares (figura 2). Ya que las pupilas de los sistemas ópticos como el ojo humano son

circulares, resulta natural describir las aberraciones en la pupila de salida mediante coordenadas polares en lugar de coordenadas cartesianas. Las expresiones para las aberraciones en coordenadas polares resultan más simples de entender.

Para pasar del sistema coordenado cartesiano al sistema coordenado polar empleamos las siguientes relaciones:

$$x = r \cos \theta \quad (14a)$$

$$y = r \sin \theta \quad (14b)$$

donde $r = \sqrt{x^2 + y^2}$. Empleando estas relaciones, las ecuaciones para el defoco (ecuación 8), el astigmatismo (ecuación 11), la coma (ecuación 12) y la esférica (ecuación 13), tienen ahora la forma:

$$W(r, \theta) = a_d r^2 \quad (\text{defoco}) \quad (15)$$

$$W(r, \theta) = a_a r^2 \cos 2\theta \quad (\text{astigmatismo}) \quad (16)$$

$$W(r, \theta) = a_c r^3 \cos \theta \quad (\text{coma}) \quad (17)$$

$$W(r, \theta) = a_s r^4 \quad (\text{esférica}) \quad (18)$$

Las superficies que describen las ecuaciones anteriores son, por supuesto, las mostradas en la figura 19 para una pupila de radio 3 mm.

Para sistemas ópticos con pupilas circulares de radio r_0 resulta conveniente normalizar la coordenada radial en la pupila de salida con respecto al radio r_0 . Así la nueva coordenada radial será:

$$\rho = \frac{r}{r_0} \quad (19)$$

y las ecuaciones 15, 16, 17 y 18 se pueden reescribir como:

$$W(r, \theta) = b_d \rho^2 \quad (\text{defoco}) \quad (20)$$

$$W(r, \theta) = b_a \rho^2 \cos 2\theta \quad (\text{astigmatismo}) \quad (21)$$

$$W(r, \theta) = b_c \rho^3 \cos \theta \quad (\text{coma}) \quad (22)$$

$$W(r, \theta) = b_s \rho^4 \quad (\text{esférica}) \quad (23)$$

donde los nuevos coeficientes son $b_a = a_a r_0^2$, $b_b = a_b r_0^2$, $b_c = a_c r_0^3$ y $b_s = a_s r_0^4$. Note que ahora la coordenada radial ρ no tiene unidades y su valor va desde 0 hasta 1. Las unidades (en milímetros, micrómetros, longitudes de onda, etc.) están contenidas en los coeficientes; entonces, cuando $\rho = 1$ el valor real en la pupila de salida será r_0 .

Conocidas las expresiones matemáticas para cada uno de los términos de aberración, ahora podemos preguntarnos de qué manera podemos reducir el efecto de las aberraciones. Por ejemplo, fijemos nuestra atención en la aberración esférica mostrada en la figura 13. Se observa que si el plano imagen lo desplazamos un poco hacia la izquierda, podemos obtener como imagen una mancha circular de menor tamaño (denominada también círculo de menor confusión) en comparación con el tamaño de la mancha que se tiene en el plano imagen gaussiano. Esto es equivalente a mantener el plano imagen en su posición original e introducir un término de defoco en el frente de onda

en la pupila de salida, de modo que el círculo de menor confusión quede en el plano imagen. Lo anterior implica sumar (o restar) un término de defoco a la ecuación 23, así la nueva expresión para la aberración esférica optimizada será $W(r, \theta) = b_s \rho^4 + b_d \rho^2$. Entonces, el resultado de sumar un término de defoco apropiado a la aberración esférica permite reducir el efecto de la aberración sobre la imagen. Este tipo de procedimiento se denomina compensación de las aberraciones (Mahajan, 1998).

El tipo de compensación que hemos descrito anteriormente se basa en el criterio de minimizar el tamaño de la PSF. Sin embargo, la compensación también se puede realizar si tomamos como criterio disminuir la distorsión del frente de onda con respecto a la esfera (o plano) de referencia. Los dos tipos de compensación no necesariamente dan el mismo resultado.

Antes de ver el resultado con el segundo criterio de compensación debemos hacer algunos comentarios sobre la superficie de referencia. Hemos

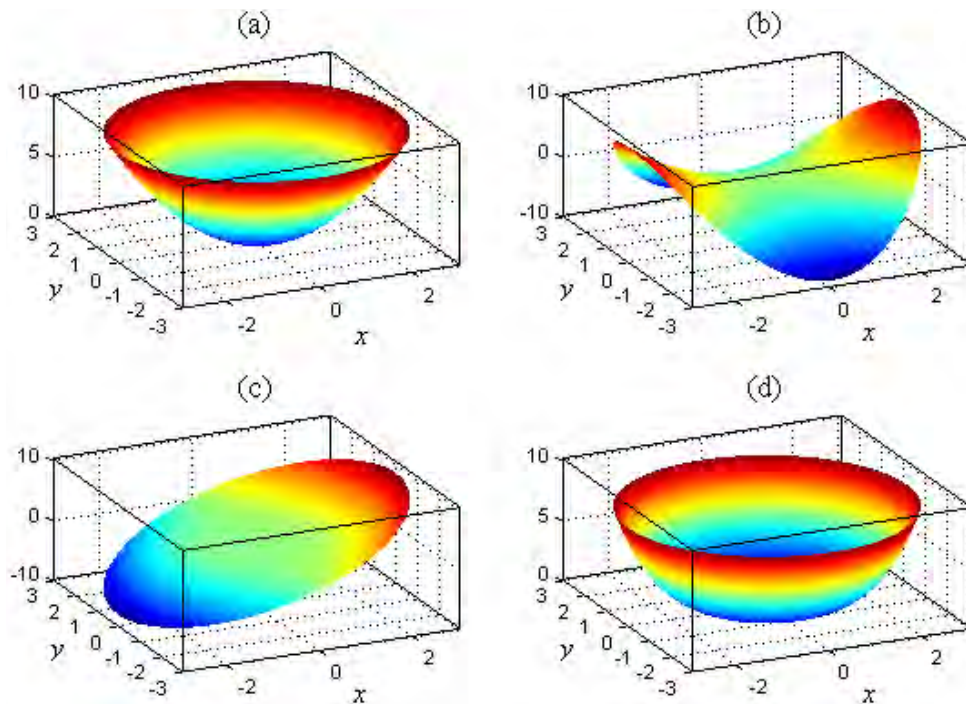


FIGURA 19. Aberración de frente de onda en coordenadas cartesianas en una pupila circular de radio 3 mm

Nota. (a) de foco; (b) astigmatismo; (c) coma; (d) esférica

visto que para sistemas formadores de imagen la superficie de referencia es una esfera cuyo radio es la distancia entre la pupila de salida y el punto imagen gaussiano, pero en un sistema como el aberrómetro ocular la superficie de referencia es un plano. La superficie de referencia ha sido introducida con el fin de comparar las distorsiones del frente de onda real con un frente de onda ideal; si nuestro interés ahora es minimizar las distorsiones del frente de onda, conviene definir una nueva superficie de referencia. Esta nueva superficie para los sistemas formadores de imagen será la esfera media. El radio de esta esfera es tal que el promedio de las desviaciones del frente de onda con respecto a la esfera será igual a cero, por lo tanto, con respecto a la esfera media, el promedio de la aberración de frente de onda es cero (sin incluir el término de pistón). En el caso del aberrómetro ocular el frente de onda de referencia sigue siendo el plano, ya que en estos sistemas se incluye un par de lentes o un conjunto de espejos o prismas (sistema Badal) que introducen un defoco al frente de onda incidente en el ojo para compensar la miopía (o hipermetropía) y así tener a la salida un frente de onda casi plano (Liang et ál., 1994).

Una vez definida la esfera media, se define ahora la varianza de la aberración de frente de onda con respecto a su valor promedio. Ya que el promedio es cero, entonces la varianza será:

$$\sigma_w = \langle W^2(\rho, \theta) \rangle \quad (24)$$

es decir, la desviación cuadrática media de la aberración de frente de onda, la cual está medida ahora con respecto a la esfera media. Si dos superficies de aberración tienen la misma esfera media, la superficie que tiene la menor distorsión será la que tenga la menor varianza.

Fijemos nuevamente la atención en la aberración esférica dada por la ecuación 23. Para que esta aberración de frente de onda tenga un valor medio de cero y una varianza mínima, debemos sumar un término de pistón y un término de defoco de acuerdo con $6\rho^4 - 6\rho^2 + 1$ (Malacara

y Malacara, 1994). Un análisis similar se realiza sobre los otros términos de aberración, por ejemplo: para minimizar la coma se puede ver que si cambiamos la inclinación de la figura 19 (c), de modo que quede horizontal, entonces tendremos la mínima varianza y un promedio igual a cero. Lo anterior implica sumar (o restar) un término de inclinación de acuerdo con $(3\rho^3 - 2\rho)\cos\theta$. Por otra parte, la aberración de frente de onda para el astigmatismo (figura 19 [c]) ya está compensada, así que en este caso no se requiere sumar ningún término adicional. Este resultado viene de la manera como se definió el astigmatismo en la figura 18. En la figura 20 se muestran los frentes de onda compensados para la aberración esférica y de coma: la forma de la aberración de frente de onda de coma (figura 20 [b]) genera una PSF de menor tamaño y con mayor simetría que la que genera la forma de la aberración de frente de onda de coma dada en la figura 19 (c). Igual ocurre para la aberración esférica (figura 20 [a]), siendo la nueva PSF de menor tamaño que la PSF generada por la aberración de frente de onda de la figura 19 (d).

Por tanto, las expresiones para las aberraciones esférica, coma y astigmatismo, de acuerdo con el segundo criterio de compensación y siguiendo la notación recomendada por la Sociedad Americana de Óptica (OSA) (Thibos et ál., 2000), tienen la forma:

$$W(\rho, \theta) = C_4^0 \sqrt{5}(6\rho^4 - 6\rho^2 + 1) \quad (25)$$

$$W(\rho, \theta) = C_3^1 \sqrt{8}(3\rho^3 - 2\rho)\cos\theta \quad (26)$$

$$W(\rho, \theta) = C_2^2 2\rho^2 \cos 2\theta \quad (27)$$

En estas expresiones se tienen nuevos coeficientes identificados con dos subíndices y un factor adicional, cuyo significado será aclarado más adelante.

Los polinomios de aberración compensados con base en la minimización de la varianza de la aberración de frente de onda dados por las ecuaciones 25, 26 y 27 son ejemplos de los polinomios de Zernike.

Los polinomios de Zernike están definidos en coordenadas polares (ρ, θ) , donde ρ es la coordenada

radial normalizada (ecuación 19) y θ es el ángulo que varía de 0° a 360° . Cada polinomio de Zernike consiste en tres términos: un factor de normalización, un polinomio radial y un término angular (seno o coseno). Por ejemplo, en las ecuaciones 25, 26 y 27 los términos de normalización son $\sqrt{5}$, $\sqrt{8}$ y 2; los polinomios radiales son $6\rho^4 - 6\rho^2 + 1$, $3\rho^3 + 2\rho$ y ρ^2 y los términos angulares $\cos\theta$ para la coma, $\cos 2\theta$ para el astigmatismo y para la esférica no se tiene ningún término angular. La forma general de los polinomios de Zernike es:

$$Z_n^m(\rho, \theta) = \begin{cases} N_n^m R_n^m(\rho) \cos(m\theta) & \text{para } m \geq 0 \\ -N_n^m R_n^m(\rho) \sin(m\theta) & \text{para } m < 0 \end{cases} \quad (28)$$

donde el índice inferior (n) denota el grado del polinomio radial y el índice superior (m) denota la frecuencia angular. La forma explícita del factor de normalización N_n^m y del polinomio radial $R_n^m(\rho)$, así como las propiedades matemáticas de los polinomios de Zernike, se puede consultar en Mahajan (1998), Thibos et ál. (2000), Malacara y DeVore (1992).

Empleando los polinomios de Zernike, cualquiera de las aberraciones se puede escribir como $C_n^m Z_n^m(\rho, \theta)$. Finalmente, el frente de onda real puede tener cualquier tipo de distorsión, de modo que la aberración total de frente de onda será la suma de los diferentes términos de aberración. Entonces, en general, la aberración de frente de onda tiene la forma:

$$W(\rho, \theta) = \sum_{n=0}^k \sum_{\substack{m=-n \\ n-|m|=\text{par}}}^n C_n^m Z_n^m(\rho, \theta) \quad (29)$$

incluyendo $n - |m| = 0$. Desde el punto de vista matemático $k \rightarrow \infty$, pero en la práctica se toman los primeros términos. ¿Cuántos términos de aberración se deben tomar? La respuesta varía: para investigadores como Atchison (2004), con $k = 6$ (lo que da 28 términos) es suficiente para modelar adecuadamente el frente de onda, mientras que la OSA (Thibos et ál., 2000) sugiere $k = 7$ (lo que da 36 términos). En cualquier caso, el número de términos es considerable, por lo que no resulta práctico escribir la suma explícita dada por la ecuación 29 para representar el frente de onda. En lugar de ello, se emplea un histograma de los índices C_n^m , tal y como se indicó en la figura 16.

En la figura 21 se muestra la reconstrucción experimental de un frente de onda aberrado generado por un modelo físico del ojo humano empleando un aberrómetro experimental tipo Hartmann-Shack (Galeano, 2009). El modelo de ojo consiste en una lente de contacto hecha en PMMA (córnea) con cara externa tórica y cara interna esférica, un diafragma de apertura de 8 mm de diámetro (iris), una lente positiva en vidrio óptico de distancia focal 20 mm y diámetro 12 mm (cristalino), y una superficie plana reflectora (retina). La luz empleada para iluminar el modelo del ojo es un láser de He-Ne (longitud de onda $\lambda = 632,8 \text{ nm}$).

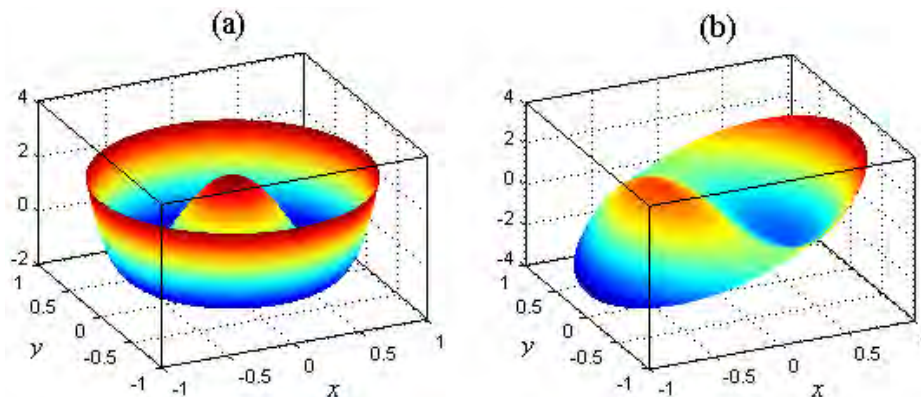


FIGURA 20. Aberraciones compensadas: (a) esférica; (b) coma

Nota. Estas superficies corresponden a los polinomios de Zernike Z_4^0 y Z_3^1 .

¿Cómo se obtienen, en la práctica, los coeficientes de Zernike? La primera información que nos entrega un aberrómetro tipo Hartmann-Shack es el patrón de Hartmann distorsionado que produce el ojo bajo prueba (imagen superior izquierda en la figura 21). Las posiciones de los puntos del patrón de Hartmann se comparan con las correspondientes posiciones de los puntos en el patrón de Hartmann de referencia, el cual se obtiene cuando sobre el arreglo de microlentes Hartmann-Shack incide un frente de onda plano. Las diferencias en las posiciones están representadas en las ecuaciones 4a y 4b) como Δx_i y Δy_i (siendo i el número del punto en el patrón). Con base en las ecuaciones 4a y 4b se obtienen las pendientes de la aberración de frente de onda, pero en forma discreta, ya que con estas ecuaciones solo podemos tener las pendientes en las posiciones de los puntos del patrón de Hartmann. Por lo tanto, se requiere de un proceso de interpolación, de manera que tengamos un mapa continuo de las pendientes en

cada dirección; una vez hecho esto, se integran las pendientes para obtener la superficie que describe la aberración de frente de onda. Hasta aquí, lo que tenemos es un conjunto de datos que graficados muestran una superficie (imagen superior derecha en la figura 21). Para obtener una ecuación que represente esta superficie se debe realizar un proceso de ajuste de los datos al conjunto de polinomios de Zernike escogido (para $k = 7$, por ejemplo). El resultado del ajuste son los coeficientes de Zernike (estas son las variables desconocidas en el proceso) que se grafican en el histograma (imagen inferior derecha en la figura 21). Los métodos de ajuste son diversos (Malacara y DeVore, 1992); el más directo es el que emplea las propiedades de ortogonalidad de los polinomios de Zernike, pero también se emplea con frecuencia el método de mínimos cuadrados (la descripción detallada de los métodos de ajuste de polinomios se sale del alcance de este artículo. Estos procedimientos requieren de conocimientos básicos en cálculo y

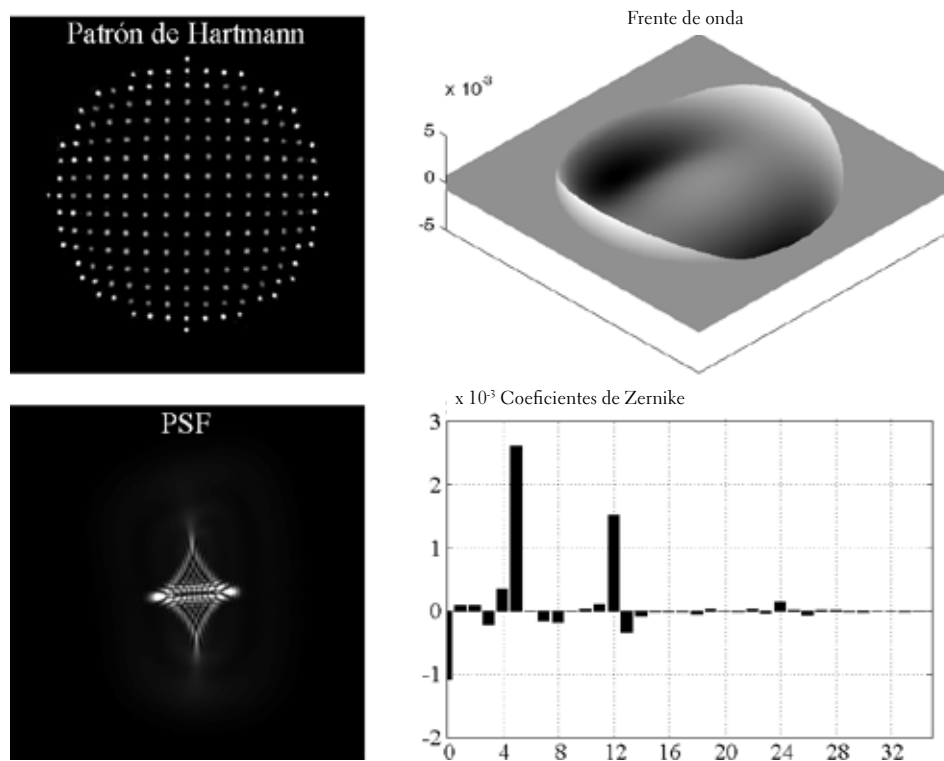


FIGURA 21. Medición del frente de onda generado por un modelo de ojo humano empleando un aberrómetro experimental tipo Hartmann-Shack

Nota. A partir del patrón de Hartmann se reconstruye el frente de onda y luego se analiza en términos de los coeficientes de Zernike. A partir del frente de onda también se puede evaluar la PSF en la retina.

álgebra lineal. La referencia Malacara y DeVore [1992] es una buena fuente para entender estos métodos). Finalmente, en la figura 21, en la parte inferior izquierda, se ha incluido el cálculo de la PSF a partir de la superficie ajustada con los coeficientes de Zernike (ecuación 29); esta PSF simula la imagen que el ojo bajo prueba vería de un punto objeto. Por otra parte, es común emplear la PSF para simular la imagen que un sujeto vería de un objeto extendido, usualmente algún optotipo. Lo anterior da una idea objetiva de cómo está la visión del sujeto.

En el histograma de la figura 21 los coeficientes se enumeran con un solo índice, de acuerdo con la siguiente relación (Thibos et ál., 2000):

$$j = \frac{n(n+2)+m}{2} \quad (30)$$

Entonces, para identificar el tipo de aberración, usamos los coeficientes C_n^m y para graficar en un histograma estos coeficientes usamos el índice j , dado por la ecuación 30. De este histograma podemos deducir lo siguiente: de los 36 términos sobresalen C_2^2 ($j = 5$) y C_4^0 ($j = 12$), es decir, astigmatismo primario a 0° y esférica primaria; luego con un menor peso sobresalen C_2^{-2} ($j = 3$), C_2^0 ($j = 4$), C_3^{-1} ($j = 7$), C_3^1 ($j = 8$), C_4^{-2} ($j = 13$) y C_6^0 ($j = 24$), es decir, astigmatismo primario a 45° , defoco, coma primaria a 90° , coma primaria a 0° , astigmatismo secundario a 0° y esférica secundaria, respectivamente. Los demás términos se omiten, ya que su contribución relativa es muy pequeña. Aunque el coeficiente de defoco es pequeño, debemos recordar que en la aberración esférica compensada también hay un término de defoco, de manera que el defoco real (ecuación 20) resulta de sumar estos términos de segundo orden. Los resultados obtenidos con el experimento concuerdan muy bien con el sistema físico bajo análisis. La lente de contacto tórica empleada para simular la córnea es la responsable de los términos de astigmatismo C_2^{-2} , C_2^2 y C_4^{-2} . Por otra parte, el $F/\#$ de la lente positiva que simula el cristalino es 2,5 ($= f/D = 20/8$), lo que introduce una aberración esférica notable, de ahí los términos C_4^0 y C_6^0 . En el experimento, el espejo que simula la

retina se desplazó hasta la posición media entre las líneas focales astigmáticas, de ahí el término de defoco. Nótese que la forma de la PSF no es el patrón de difracción correspondiente al círculo de menor confusión del astigmatismo, debido a la presencia de la aberración esférica. Por último, los términos de coma C_3^{-1} y C_3^1 sugieren que hay un ligero descentramiento entre la lente de contacto y la lente positiva, o que entre las caras de la lente de contacto existe una forma ligeramente prismática. En conclusión, en el modelo de ojo las aberraciones dominantes son el astigmatismo y la esférica.

Un análisis similar se puede hacer en ojos humanos reales. Investigaciones sobre el tipo de aberraciones más comunes en ojos reales muestran que además de las aberraciones defoco, astigmatismo, coma y esférica también son comunes las aberraciones trefoil (Z_3^{-3} , Z_3^3), tetrafoil (Z_4^{-4} , Z_4^4) y pentafoil (Z_5^{-5} , Z_5^5) (Porter, Guirao y Williams, 2000). En el apéndice 1 se muestran las ecuaciones explícitas y las gráficas de los polinomios de Zernike hasta orden 5.

CONCLUSIONES

Los polinomios de Zernike constituyen una base matemática adecuada para representar las aberraciones oculares. Cada término de Zernike representa la aberración de frente de onda compensada de una de las aberraciones ópticas, de modo que su varianza con respecto a una superficie de referencia adecuada es la mínima y, además, el promedio de las distorsiones del frente de onda con respecto a la superficie de referencia es cero. En un sistema óptico formador de imagen la superficie de referencia es una esfera, mientras que en el aberrómetro tipo Hartmann-Shack la superficie de referencia es un plano. Es más, en el aberrómetro tipo Hartmann-Shack la aberración de frente de onda corresponde al frente de onda real. Como cada polinomio de Zernike está compensado, lo anterior también nos permite saber cuánto de las aberraciones que compensan el término de más alto

orden se requiere para obtener la PSF óptima en la imagen final. El histograma de coeficientes de Zernike es una forma conveniente de representar el frente de onda total, ya que este nos da una idea rápida de cuáles son los términos de aberración que dominan la forma global de la aberración de frente de onda.

REFERENCIAS

- Atchison, D. A. (2004). Recent Advances in Representation of Monochromatic Aberrations of Human Eyes. *Clinical and Experimental Optometry*, 87(3), 138-148.
- Atchison, D. A. & Smith, G. (2000). *Optics of the Human Eye*. Oxford: Butterworth-Heinemann.
- Galeano, J. C. (2009). *Medición de aberraciones ópticas empleando un arreglo de microlentes Hartmann-Shack con aplicación en oftalmología*. Tesis de Maestría en Ciencias-Física, Facultad de Ciencias, Universidad Nacional de Colombia.
- Goodman, J. W. (1996). *Introduction to Fourier Optics*. New York: McGraw-Hill.
- Hecht, E. (2000). *Óptica*. Madrid: Addison Wesley Iberoamericana.
- Liang, J. et al. (1994). Objective Measurement of Wave Aberrations of the Human Eye with the Use of a Hartmann-Shack Wave-front Sensor. *Journal of the Optical Society of America*, 11(7), 1949-1957.
- Mahajan, V. N. (1998). *Optical Imaging and Aberrations*, Part I. Washington: SPIE Optical Engineering Press, Bellingham.
- Malacara, D., & DeVore, S. L. (1992). Interferogram Evaluation and Wavefront Fitting. In *Optical Shop Testing*. (2nd ed.). New York: John Wiley and Sons.
- Malacara, D., & Malacara, Z. (1994). *Handbook of Lens Design*. New York: Marcel Dekker.
- Porter, J., Guirao, A. & Williams, D. R. (2000). A Compact Description of the Eye's Monochromatic Aberrations in a Large Population. *Vision Science and its Applications, Trends in Optics and Photonics - Optical Society of America*, 35, 199-204.
- Rayces, J. L. (1964). Exact Relation between Wave Aberration and Ray Aberration, *Optica Acta*, 11(2), 85-88.
- Thibos, L. N. (2000). Principles of Hartmann-Shack Aberrometry. *Vision Science and its Applications, Trends in Optics and Photonics - Optical Society of America*, 35, 163-169.
- Thibos, L. N. et al. (2000). Standards for Reporting the Optical Aberrations of Eyes. *Vision Science and its Applications, Trends in Optics and Photonics - Optical Society of America*, 35, 232-244.

Recibido: 17 de marzo del 2011

Aceptado: 8 de septiembre del 2011

CORRESPONDENCIA

Yobani Mejía Barbosa

ymejia@unal.edu.co

Apéndice 1. Polinomios de Zernike

j	n	m	$Z_n^m(\rho, \theta)$	j	n	m	$Z_n^m(\rho, \theta)$
0	0	0	1	11	4	-2	$\sqrt{10}(4\rho^4-3\rho^2)\text{sen}2\theta$
1	1	-1	$2\rho\text{sen}\theta$	12	4	0	$\sqrt{5}(6\rho^4-6\rho^2+1)$
2	1	1	$2\rho\cos\theta$	13	4	2	$\sqrt{10}(4\rho^4-3\rho^2)\cos 2\theta$
3	2	-2	$\sqrt{6}\rho^2\text{sen}2\theta$	14	4	4	$\sqrt{10}\rho^4\cos 4\theta$
4	2	0	$\sqrt{3}(2\rho^2-1)$	15	5	-5	$\sqrt{12}\rho^5\text{sen}5\theta$
5	2	2	$\sqrt{6}\rho^2\cos 2\theta$	16	5	-3	$\sqrt{12}(5\rho^5-4\rho^3)\text{sen}3\theta$
6	3	-3	$\sqrt{8}\rho^3\text{sen}3\theta$	17	5	-1	$\sqrt{12}(10\rho^5-12\rho^3+3\rho)\text{sen}\theta$
7	3	-1	$\sqrt{8}(3\rho^3-2\rho)\text{sen}\theta$	18	5	1	$\sqrt{12}(10\rho^5-12\rho^3+3\rho)\cos\theta$
8	3	1	$\sqrt{8}(3\rho^3-2\rho)\cos\theta$	19	5	3	$\sqrt{12}(5\rho^5-4\rho^3)\cos 3\theta$
9	3	3	$\sqrt{8}\rho^3\cos 3\theta$	20	5	5	$\sqrt{12}\rho^5\cos 5\theta$
10	4	-4	$\sqrt{10}\rho^4\text{sen}4\theta$				

Apéndice 2. Representación en tres dimensiones de los polinomios de Zernike hasta orden 5

