Zernike polynomials

Frederik Zernike, 1953 Nobel prize
Reminder

\[ W = W(y', \rho, \varphi) \]
\[ = W_{000} + W_{200} y'^2 + W_{400} y'^4 \]
\[ + W_{020} \rho^2 \]
\[ + W_{111} y' \rho \cos \varphi \]
\[ + W_{040} \rho^4 + W_{131} y' \rho^3 \cos \varphi + W_{222} y'^2 \rho^2 \cos^2 \varphi + W_{2205} y'^2 \rho^2 + W_{311} y'^3 \rho \cos \varphi \]

...
Advantages

• Zernike polynomials are of great interest in many fields:
  – optical design
  – optical metrology
  – adaptive optics
  – ophtalmology (corneal topography, ocular aberrometry)
  – freeform optics...

• For a circular pupil, Zernike polynomials form an **orthonormal basis**. Hence formalism is easier

• Set of basis shapes or topographies of a surface

• “Real” surface (wavefront, mirror…) is constructed from linear combination of basis shapes or modes

Thierry Lépine - Optical design
Disadvantages

• There are some optical systems with non-circular pupils:
  – Telescopes: annular and hexagonal pupils...
  – Lasers: squared pupils

• In these cases, there exist specific orthonormal polynomials other than the Zernike ones
Definitions

Warning: VERY CONFUSING
- what axis is the reference axis? x, y
- [fringe, standard, Noll, Zygo, Wyant, Born & Wolf... ] ordering
- normalization: zero-to-peak (0-P), RMS
- units: waves, micrometers

\[ Z_n^m(\rho, \phi) = \begin{cases} 
\sqrt{\frac{2(n+1)}{1+\delta_{m0}}} R_n^{|m|}(\rho) \sin(m\phi) & \text{if } m \geq 0 \\
-\sqrt{2(n+1)} R_n^{|m|}(\rho) \cos(m\phi) & \text{if } m < 0 
\end{cases} \]

The polynomials \( R_n^m(\rho), n \in N, m \in Z \), are defined such that:

\[ R_n^m(\rho) = \sum_{k=0}^{n-m} \frac{(-1)^k (n-k)!}{k! \left[ \frac{n+m}{2} - k \right]! \left[ \frac{n-m}{2} - k \right]!} \rho^{n-2k}, \text{ with } n \geq |m| \text{ et } n - |m| \text{ even} \]
Consequences

\( R_n^m (\rho) \) is a polynomial in \( \rho \) of degree \( n \) and is naught if \(|m| > n \) or \((n - |m|)\) uneven.

\[ R_n^m (\rho) = R_n^{-m} (\rho) \]

\[ |R_n^m (\rho)| \leq 1 \]

\[ R_n^m (1) = 1 \]

Zernike polynomials are orthonormal if and only if:

\[ \frac{1}{\pi} \int_0^1 \int_0^{2\pi} Z_n^m (\rho, \varphi) Z_n^{m'} (\rho, \varphi) \rho d\rho d\varphi = \delta_{nn'} \delta_{mm'} \]

All wavefronts errors \( W \) can then be expressed as a sum of the \( Z_n^m \):

\[ W(y', \rho, \varphi) = \sum_{n,m} c_{nm} (y') Z_n^m (\rho, \varphi) \]

More about field dependence of \( c_{nm} \), see for instance:
This leads to very interesting properties:

– The coefficient associated with the piston is equal to the mean value of the aberration function:

\[ \overline{W} = \frac{1}{\pi} \int \int W \rho d\rho d\varphi = c_{00} \]

– The quadratic sum of the coefficients, except that of the piston, is equal to the variance of the aberration function:

\[ \overline{W}^2 = \frac{1}{\pi} \int \int W^2 \rho d\rho d\varphi \Rightarrow \sigma_W^2 = (\overline{W} - \overline{W})^2 = \sum_{(n,m)\neq(0,0)} c_{nm}^2 \]

– Each coefficient, except that of the piston, is equal to the RMS value of the associated aberration

– The value of a coefficient does not depend on the number of polynomials used in the development:

\[ c_{nm} = \frac{1}{\pi} \int \int W Z_n^m \rho d\rho d\varphi \]
The most important! Each term contains the appropriate amount of lower order terms in order to minimize the RMS wavefront error of that term. So Zernike polynomials represent balanced classical aberrations.

Using Zernike polynomials, the evaluation of the Strehl ratio is simpler (see further in this course).

Many examples will be studied in classwork.
# The first Zernike polynomials

<table>
<thead>
<tr>
<th>n</th>
<th>m</th>
<th>polynomial</th>
<th>name</th>
<th>Notation from Zygo, REOSC*...</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td>piston</td>
<td>$Z^0_0$</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>$2 \rho \sin \varphi$</td>
<td>tilt X</td>
<td>$Z^1_1$</td>
</tr>
<tr>
<td></td>
<td>-1</td>
<td>$2 \rho \cos \varphi$</td>
<td>tilt Y</td>
<td>$Z^{-1}_1$</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>$\sqrt{3} \left(2 \rho^2 - 1\right)$</td>
<td>defocus</td>
<td>$Z^0_1$</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>$\sqrt{6} \rho^2 \sin 2 \varphi$</td>
<td>astigmatism 0°</td>
<td>$Z^2_2$</td>
</tr>
<tr>
<td></td>
<td>-2</td>
<td>$\sqrt{6} \rho^2 \cos 2 \varphi$</td>
<td>astigmatism 45°</td>
<td>$Z^{-2}_2$</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>$\sqrt{8} \left(3 \rho^3 - 2 \rho\right) \sin \varphi$</td>
<td>coma X</td>
<td>$Z^1_2$</td>
</tr>
<tr>
<td></td>
<td>-1</td>
<td>$\sqrt{8} \left(3 \rho^3 - 2 \rho\right) \cos \varphi$</td>
<td>coma Y</td>
<td>$Z^{-1}_2$</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>$\sqrt{5} \left(6 \rho^4 - 6 \rho^2 + 1\right)$</td>
<td>spherical aberration</td>
<td>$Z^0_2$</td>
</tr>
</tbody>
</table>

*$\left(R^m_{2n-m}\right)_{This\ lecture} = \left(R^m_n\right)_{Zygo,\ REOSC...}$

Thierry Lépine - Optical design
The first Zernike polynomials

The first Zernike polynomials

\[ Z_1^1 \text{ (tilt X, } c_{11} = 1 \times \lambda) \]

\[ Z_2^0 \text{ (defocus, } c_{20} = 1 \times \lambda) \]

\[ Z_2^2 \text{ (astigmatism } 0^\circ, c_{22} = 1 \times \lambda) \]

\[ Z_3^{-1} \text{ (coma Y, } c_{3-1} = 1 \times \lambda) \]

\[ Z_4^0 \text{ (spherical aberration, } c_{40} = 1 \times \lambda) \]
3rd order aberrations: best focus et tolerances

- We saw that: \( \sigma_w^2 = \left( W - \bar{W} \right)^2 = \sum_{(n,m) \neq (0,0)} c_{nm}^2 \)
- And we know that:

\[
S \approx 1 - \frac{4\pi^2}{\lambda^2} \sigma_w^2
\]

- So:

\[
S \approx 1 - \frac{4\pi^2}{\lambda^2} \sum_{(n,m) \neq (0,0)} c_{nm}^2 \\
\approx 1 - \frac{4\pi^2}{\lambda^2} \left( c_{11}^2 + c_{1-1}^2 + c_{20}^2 + \ldots \right)
\]
3rd order aberrations: best focus et tolerances

• At the beginning, the reference sphere is centered on the paraxial image. The Strehl ratio depends on the tilts et defocus coefficients: \( c_{11}, c_{1-1}, c_{20} \)

• We can change the reference sphere, applying the opposite previous tilts et defocus to the center of the reference sphere

• With respect to this new reference sphere, tilts and defocus have been canceled: the Strehl ratio is maximized, and the center of this new sphere is the best focus!
3\textsuperscript{rd} order aberrations: best focus et tolerances

- Hence for the best focus:

\[
S \approx 1 - \frac{4\pi^2}{\lambda^2} \sum_{(n,m)\neq(0,0),(1,1),(1,-1),(2,0)} c_{nm}^2
\]
Case of the 3\textsuperscript{rd} order spherical aberration

• Reminder:
  – With the Seidel formalism, we could determine the position of the best focus, using rather tedious calculations...
Reminder : Seidel formalism

At the best focus, the Strehl ratio is maximized, i.e. the variance of $\sigma_W^2$ is minimized.

We need to introduce a defocus: $W = W_{040} \rho^4 + \varepsilon_z \frac{R_p^2}{2R^2} \rho^2 = a \rho^4 + b \rho^2$

$$\sigma_w^2 = \overline{W^2} - \overline{W}^2$$

$$\overline{W} = \frac{1}{\pi} \int_0^{2\pi} \int_0^1 W \rho d\rho d\theta = \left(\frac{a}{3} + \frac{b}{2}\right)$$

$$\overline{W^2} = \frac{1}{\pi} \int_0^{2\pi} \int_0^1 W^2 \rho d\rho d\theta = \left(\frac{a^2}{5} + \frac{ab}{2} + \frac{b^2}{3}\right)$$

$$\Rightarrow \sigma_w^2 = \left(\frac{4a^2}{45} + \frac{ab}{6} + \frac{b^2}{12}\right)$$

We are looking for the defocus (b) for which $\sigma_w^2$ is minimized, i.e. $\frac{\partial \sigma_w^2}{\partial b} = 0$,

Hence: $b = -a$, so: $\varepsilon_z \frac{R_p^2}{2R^2} = -W_{040} \Rightarrow \varepsilon_z = -\frac{2R^2}{R_p^2} W_{040} = \frac{LA}{2}$

The best focus is halfway between the paraxial et marginal focii!
Reminder : Seidel formalism

Let us try to evaluate the maximum aberration function values for each case:

\[ W_{\text{paraxial}} = W_{040} \rho^4 \]

\[ \Rightarrow (W_{\text{paraxial}})_{\max} = W_{040} \]

\[ W_{BF} = W_{040} \left( \rho^4 - \rho^2 \right) \]

When \( W_{BF} \) is maximized:

\[ \frac{\partial W_{MF}}{\partial \rho} = 0 \Rightarrow \rho = \frac{1}{\sqrt{2}} \Rightarrow (W_{BF})_{\max} = -\frac{W_{040}}{4} = -\frac{(W_{\text{paraxial}})_{\max}}{4} \]

For the best focus, the maximum wavefront error is 4 times smaller than it is for the paraxial focus, and the wavefront is retarded, with respect to the reference sphere.

We note also that the aberration function for the best focus is naught on the axis (obvious) and at the edge of the pupil:
Same case with the Zernikes

If the reference sphere is centered on the paraxial image, we know that: \( W = W_{040} \rho^4 \)

We write \( W \) as a sum of the \( Z_n^m \): \( W = c_{40} Z_4^0 + c_{20} Z_2^0 (+c_{00} Z_0^0) \)

We easily get: \( c_{20} = \frac{W_{040}}{2\sqrt{3}}, c_{40} = \frac{W_{040}}{6\sqrt{5}} \)

At the best focus, we "cancel" the coefficient \( c_{20} \), hence: \( W = c_{40} Z_4^0 \). So, \( Z_4^0 \) is the compensated spherical aberration (we introduced the appropriate amount of defocus to minimize the RMS wavefront error).

Application 1:

Let us suppose that \( W \) is equal to \( \frac{\lambda}{4} \) at the edge of the pupil (\( \rho = 1 \)), ie. \( W_{040} = \frac{\lambda}{4} \). Hence:

\[
S_{\text{paraxial focus}} \approx 1 - \frac{4\pi^2}{\lambda^2} \left( c_{20}^2 + c_{40}^2 \right) \approx 0.78 \quad \text{et} \quad S_{\text{best focus}} \approx 1 - \frac{4\pi^2}{\lambda^2} \left( c_{40}^2 \right) \approx 0.99
\]
Application 2:
We evaluate the tolerance at the 2 focii, assuming that $S \geq 0.8$. We get:

- paraxial focus: $1 - \frac{4\pi^2}{\lambda^2} \left(c_{20}^2 + c_{40}^2\right) \geq 0.8 \Rightarrow W_{040} \leq 0.24 \lambda$

- best focus: $1 - \frac{4\pi^2}{\lambda^2} \left(c_{40}^2\right) \geq 0.8 \Rightarrow W_{040} \leq 0.95 \lambda$

Application 3:
We evaluate the position of the best focus.

For this, we need to introduce a defocus $\varepsilon_z \frac{R_p^2}{2R^2} \rho^2$ to compensate the term $c_{20}Z_2^0$.

Hence: $\varepsilon_z = -\frac{2R^2}{R_p^2} W_{040} = \frac{LA}{2}$
If the reference sphere is centered on the paraxial image, we have: \( W = W_{131} y' \rho^3 \cos \varphi \)

We write \( W \) as a sum of the \( Z_n^m : W = c_{3-1} Z_3^{-1} + c_{1-1} Z_1^{-1} (+c_{00} Z_0^0) \)

We get: \( c_{1-1} = \frac{W_{131} y'}{3}, c_{3-1} = \frac{W_{131} y'}{3\sqrt{8}} \)

At the best focus, we "cancel" the coefficient \( c_{11} \), hence: \( W = c_{31} Z_3^{-1} \). Then, \( Z_3^{-1} \) is the compensated coma (we introduced the appropriate amount of tilt to minimize the RMS wavefront error)

Application 1:

We evaluate the tolerance at the 2 focii, assuming that: \( S \geq 0,8 \). We get:

- paraxial focus: \( 1 - \frac{4\pi^2}{\lambda^2} \left( c_{3-1}^2 + c_{1-1}^2 \right) \geq 0,8 \Rightarrow W_{131} y' \leq 0,2 \lambda \)

- best focus: \( 1 - \frac{4\pi^2}{\lambda^2} \left( c_{3-1}^2 \right) \geq 0,8 \Rightarrow W_{131} y' \leq 0,60 \lambda \)

Application 2:

We evaluate the position of the best focus.

For this, we need to introduce a tilt \( \varepsilon_y \frac{R_p}{R} \rho \sin \varphi \) to compensate the term \( c_{1-1} Z_1^{-1} \).

So: \( \varepsilon_y = -\frac{2}{3} W_{131} y' \frac{R}{R_p} \)

This tilt is equal to the third of the distance between the paraxial image and the center of the greatest circle of coma (\( \rho = 1 \))

Reminder: \( y' \) et \( \rho \) are dimensionless!
Astigmatism

If the reference sphere is centered on the paraxial image, we have:
\[ W = W_{222} y'^2 \rho^2 \cos^2 \varphi \]

We write \( W \) as a sum of the \( Z_n^m \):
\[ W = c_{-2} Z_2^{-2} + c_{20} Z_0^0 (+c_{00} Z_0^0) \]

Then:
\[ c_{20} = \frac{W_{222} y'^2}{4\sqrt{3}}, \quad c_{-2} = \frac{W_{222} y'^2}{2\sqrt{6}} \]

At the best focus, we "cancel" the term \( c_{20} \), then:
\[ W = c_{-2} Z_2^{-2} \]
Hence, \( Z_2^{-2} \) is the compensated astigmatism (we introduced the appropriate amount of defocus to minimize the RMS wavefront error)

Application 1:

We evaluate the tolerance at the 2 focii, assuming that: \( S \geq 0,8 \). We get:

- paraxial focus: \( 1 - \frac{4\pi^2}{\lambda^2} (c_{-2}^2 + c_{20}^2) \geq 0,8 \Rightarrow W_{222} y'^2 \leq 0,28\lambda \)

- best focus: \( 1 - \frac{4\pi^2}{\lambda^2} (c_{-2}^2) \geq 0,8 \Rightarrow W_{222} y'^2 \leq 0,34\lambda \)

Application 2:

We evaluate the position of the best focus.

For this, we need to introduce a defocus \( \varepsilon_z \frac{R_p^2}{2R^2} \rho^2 \) to compensate the term \( c_{20} Z_0^0 \). Hence:

\[ \varepsilon_z = - \frac{R^2}{R_p^2} W_{222} y'^2 \]

This defocus is equal to half the distance between the 2 focii.
Field curvature

For the reference sphere centered on the paraxial image, we have: \( W = W_{2205} y'^2 \rho^2 \)

We write \( W \) as a sum of the \( Z_n^m \): \( W = c_{20} Z_2^0 (+c_{00} Z_0^0) \)

We get: \( c_{20} = \frac{W_{2205} y'^2}{2\sqrt{3}} \)

The field curvature IS a defocus, depending on the field!
At the best focus, there is no aberration!

Application 1:
We evaluate the tolerance for the paraxial focus, assuming that: \( S \geq 0,8 \). We get:

\[
1 - \frac{4\pi^2}{\lambda^2} \left( c_{20}^2 \right) \geq 0,8 \Rightarrow W_{2205} y'^2 \leq 0,25 \lambda
\]

Application 2:
Where is the best focus?

For this, we introduce a defocus \( \mathcal{E}_z \frac{R_p^2}{2R^2} \rho^2 \) to compensate the term \( c_{20} Z_2^0 \). Hence:

\[
\mathcal{E}_z = -\frac{2R^2}{R_p} W_{2205} y'^2
\]

already demonstrated previously!
Distorsion

For the reference sphere centered on the paraxial image, we know that: \( W = W_{311} y^3 \rho \cos \varphi \)

We write \( W \) as a sum of the \( Z_n^m \): \( W = c_{1-1} Z_1^{-1} \)

We easily get: \( c_{1-1} = \frac{W_{331} y^3}{2} \)

The distorsion IS a tilt, depending on the field!

At the best focus, there is no aberration!

Application 1:

We evaluate the tolerance at the paraxial focus, assuming that: \( S \geq 0.8 \). We get:

\[
1 - \frac{4\pi^2}{\lambda^2} \left( c_{1-1} \right)^2 \geq 0.8 \Rightarrow W_{331} y^3 \leq 0.14 \lambda
\]

Application 2:

What is the position of the best focus?

To do this, we introduce a tilt \( \varepsilon_y \frac{R_p}{R} \rho \sin \varphi \) to compensate the term \( c_{1-1} Z_1^{-1} \). Hence:

\[
\varepsilon_y = -\frac{R}{R_p} W_{331} y^3
\]

already demonstrated previously!
Tolerances: summary

<table>
<thead>
<tr>
<th>3rd order aberration</th>
<th>Paraxial focus</th>
<th>Best focus</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spherical aberration</td>
<td>0,24(\lambda)</td>
<td>0,95(\lambda)</td>
</tr>
<tr>
<td>Coma</td>
<td>0,2(\lambda)</td>
<td>0,6(\lambda)</td>
</tr>
<tr>
<td>Astigmatism</td>
<td>0,28(\lambda)</td>
<td>0,34(\lambda)</td>
</tr>
<tr>
<td>Field curvature</td>
<td>0,25(\lambda)</td>
<td></td>
</tr>
<tr>
<td>Distorsion</td>
<td>0,14(\lambda)</td>
<td></td>
</tr>
</tbody>
</table>
Zernike polynomials for 3\textsuperscript{rd} and 5\textsuperscript{th} orders

<table>
<thead>
<tr>
<th>( n )</th>
<th>( m )</th>
<th>polynomial</th>
<th>name</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td>piston</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>( 2 \rho \sin \varphi )</td>
<td>tilt X</td>
</tr>
<tr>
<td>-1</td>
<td></td>
<td>( 2 \rho \cos \varphi )</td>
<td>tilt Y</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>( \sqrt{3}(2 \rho^2 - 1) )</td>
<td>defocus</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>( \sqrt{6} \rho^2 \sin 2\varphi )</td>
<td>3\textsuperscript{rd} order astigmatism 0°</td>
</tr>
<tr>
<td>-2</td>
<td></td>
<td>( \sqrt{6} \rho^2 \cos 2\varphi )</td>
<td>3\textsuperscript{rd} order astigmatism 45°</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>( \sqrt{8}(3 \rho^3 - 2 \rho) \sin \varphi )</td>
<td>3\textsuperscript{rd} order coma X</td>
</tr>
<tr>
<td>-1</td>
<td></td>
<td>( \sqrt{8}(3 \rho^3 - 2 \rho) \cos \varphi )</td>
<td>3\textsuperscript{rd} order coma Y</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>( \sqrt{5}(6 \rho^4 - 6 \rho^2 + 1) )</td>
<td>3\textsuperscript{rd} order spherical aberration</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>( \sqrt{10}(4 \rho^4 - 3 \rho^2) \sin 2\varphi )</td>
<td>5\textsuperscript{th} order astigmatism 0°</td>
</tr>
<tr>
<td>4</td>
<td>-2</td>
<td>( \sqrt{10}(4 \rho^4 - 3 \rho^2) \cos 2\varphi )</td>
<td>5\textsuperscript{th} order astigmatism 45°</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>( \sqrt{12}(10 \rho^5 - 12 \rho^3 + 3 \rho) \sin \varphi )</td>
<td>5\textsuperscript{th} order coma X</td>
</tr>
<tr>
<td>5</td>
<td>-1</td>
<td>( \sqrt{12}(10 \rho^5 - 12 \rho^3 + 3 \rho) \cos \varphi )</td>
<td>5\textsuperscript{th} order coma Y</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>( \sqrt{7}(20 \rho^6 - 30 \rho^4 + 12 \rho^2 - 1) )</td>
<td>5\textsuperscript{th} order spherical aberration</td>
</tr>
</tbody>
</table>
To go further!

Orthonormal polynomials in wavefront analysis: error analysis

Guang-ming Dai¹,* and Virendra N. Mahajan²

¹Laser Vision Correction Group, Advanced Medical Optics, 510 Cottonwood Drive, Milpitas, California 95035, USA
²The Aerospace Corporation, 2350 East El Segundo Boulevard, El Segundo, California 90245, USA

*Corresponding author: george.dai@amo-inc.com

Received 4 April 2008; accepted 23 May 2008;
posted 2 June 2008 (Doc. ID 94594); published 23 June 2008

1 July 2008 / Vol. 47, No. 19 / APPLIED OPTICS

Zernike circle polynomials are in widespread use for wavefront analysis because of their orthogonality over a circular pupil and their representation of balanced classical aberrations. However, they are not appropriate for noncircular pupils, such as annular, hexagonal, elliptical, rectangular, and square pupils, due to their lack of orthogonality over such pupils. We emphasize the use of orthonormal polynomials for such pupils, but we show how to obtain the Zernike coefficients correctly. We illustrate that the wavefront fitting with a set of orthonormal polynomials is identical to the fitting with a corresponding set of Zernike polynomials. This is a consequence of the fact that each orthonormal polynomial is a linear combination of the Zernike polynomials. However, since the Zernike polynomials do not represent balanced aberrations for a noncircular pupil, the Zernike coefficients lack the physical significance that the orthonormal coefficients provide. We also analyze the error that arises if Zernike polynomials are used for noncircular pupils by treating them as circular pupils and illustrate it with numerical examples. © 2008 Optical Society of America

OCIS codes: 010.1090, 010.7350, 220.1010, 220.3180, 220.0220, 330.4460.

Thierry Lépine - Optical design