

A Bottom-Up Derivation of Primary Wavefront Aberrations

Sukriti Mantri, 2026

Abstract

In this work, a bottom-up first principles derivation of the Wavefront Aberrations is done. Starting from a refracting lens surface, optical path length and the difference in optical path length for various rays is calculated and is then related to the wavefront aberration error. On-axis aberrations like Spherical Aberration are derived as well as off-axis aberrations like Coma, Astigmatism, Field Curvature, and Distortion are also derived. This work elaborates on the origin story of the wavefront aberration formulas that optical engineers are used to seeing in their day-to-day work.

Chief ray of a beam originating from a point intersects the aperture stop at the optical axis. Chief ray is also the ray around which the beam is centered. For an on-axis beam, the chief ray would strike the aperture stop as well as any optical element it meets at the same point ($y_0 = 0$). For an off-axis beam (oriented at an angle θ with the optical axis), the chief ray would intersect the aperture stop at $y_0 = 0$. However, it would generally cross other optical elements at non-zero y (unless the aperture stop is at the optical element). It is this understanding that helps in understanding the on-axis and off-axis aberrations, their dependence on field, pupil, and optical system, and their elimination methods.

Optical Path Length (OPL) is a property that is shared among the two methods of studying optics - geometrical/ray optics and physical/wave optics. We can use the OPL of a ray and then convert it into the wavefront and phase information – properties defining physical/wave optics. Variations in the optical path length of a beam originating from a single point can then lead to calculation of wavefront errors.

Optical Path Length is the medium-aware distance a ray travels. A ray takes more time in a material with higher refractive index, as light wave would travel at a lower speed than air (or any material with lower refractive index).

$$t = \frac{\text{distance}}{\text{speed}} = \frac{\text{distance}}{(\text{speed in air})/n} = n \times \frac{\text{distance}}{\text{speed in air}} = \text{time taken to travel } \{n * \text{distance}\}$$

Thus, every distance travelled is replaced by refractive index times the actual length to make it medium aware. In a more accurate sense, a light wave accumulates optical path length as it travels as shown below:

$$OPL = \int n \cdot ds \quad (1)$$

To understand wavefront aberrations, we, therefore, need to compare optical path travelled by rays as they cross various optical elements in their path.

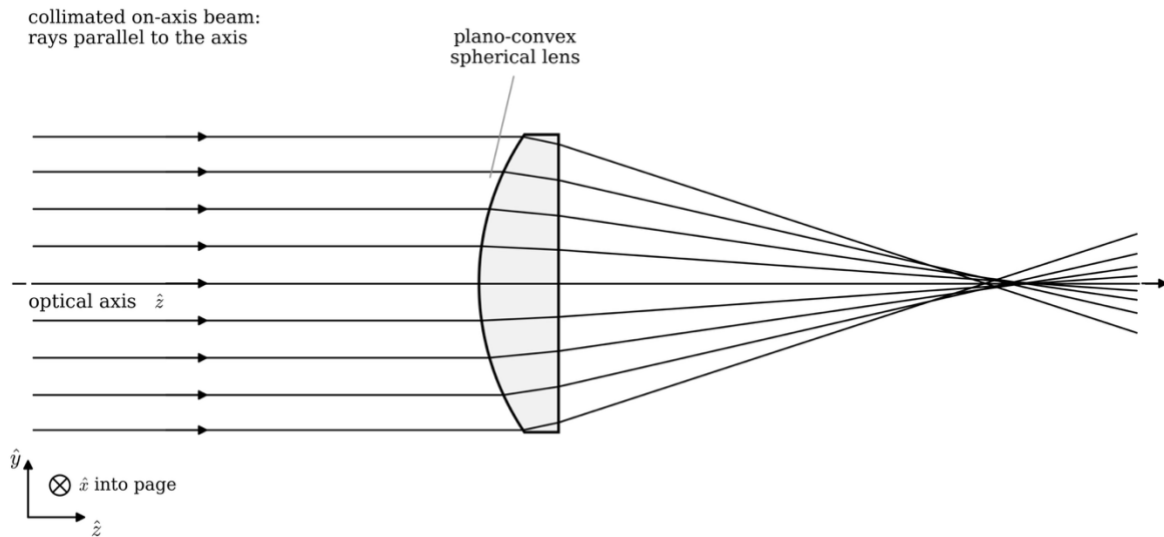


Figure 1: Plano Convex Spherical Lens with the optical axis along z .

Imagine an optical system with a plano-convex spherical lens as shown in Fig 1. The optical axis is along z ($\hat{k}$). The vertical in-plane direction is y ($\hat{j}$) and the direction into the paper is x ($\hat{i}$).

Fig 1 also shows a bunch of rays parallel to the optical axis (on-axis). We will focus on two rays as shown in Fig 2 – one ray being the chief ray of that beam (being along the optical axis – and intersecting the lens

at its intersection with the optical axis - O) and the second one being a general ray hitting the lens at point A.

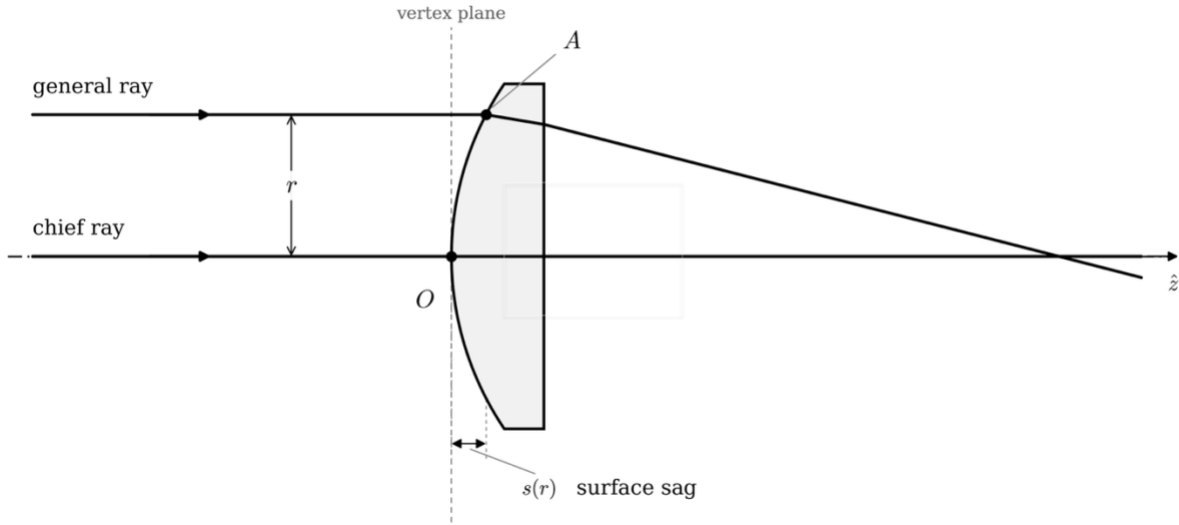


Figure 2: On axis field (zero-degree field angle for collimated beam). Two Rays are being shown – Chief ray and any general ray.

Taking the general ray, we now calculate its optical path length.

$$OPL_{ray A} = n_{medium} * (O'A) + n_{lens} * t(r) \quad (2)$$

Taking n_{medium} to be of air = 1.

$$OPL_{ray A} = (O'A) + n_{lens} * t(r)$$

$$OPL_{ray A} = (O'A) + n_{lens} * (t_0 - O'A)$$

$$OPL_{ray A} = n_{lens} * t_0 - O'A(n_{lens} - 1)$$

Chief ray is the special case where $O'A = 0$.

Thus,

$$OPL_{ray O} = n_{lens} * t_0$$

Thus, the **optical path difference (OPD)** is the difference between the optical path accumulated by the chief ray and the general ray.

$$OPD = OPL_{ray A} - OPL_{ray O}$$

$$OPD = n_{lens} * t_0 - O'A(n_{lens} - 1) - n_{lens} * t_0$$

$$OPD = -O'A(n_{lens} - 1) \quad (3)$$

Thus, the optical path difference is the function of the distance along optical axis of how much the surface dips (or sags) in comparison to its location at the vertex. This is called as the surface sag. To calculate OPD, we need to calculate the surface sag. Please note that that here the way we have defined OPD, negative optical path difference implies a phase lag from chief ray and positive optical path difference implies a phase gain from chief ray.

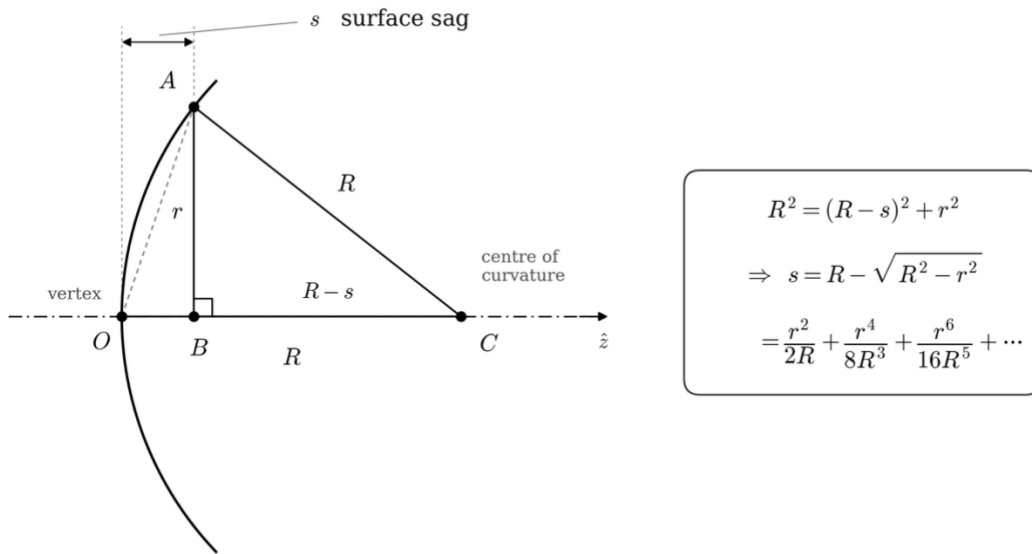


Figure 3: Figure showing the surface sag and the triangle ABC which can then be used to calculate the surface sag and the OPL travelled by any ray

Fig 3 can be used to understand the following set of equations. Triangle A-B-C forms a right triangle

Therefore,

$$\begin{aligned}
 BC^2 + AB^2 &= AC^2 \\
 (R - s)^2 + (r)^2 &= (R)^2 \\
 s^2 - 2 * R * s + r^2 &= 0 \\
 s &= R - \sqrt{(R^2 - r^2)} \\
 s &= R - R(1 - \frac{r^2}{R^2})^{1/2}
 \end{aligned}$$

Using binomial series expansion, the surface sag (s) comes out to be:

$$\begin{aligned}
 s &= R - R \left(1 - \frac{1}{2} \left(\frac{r^2}{R^2} \right) - \frac{1}{8} \left(\frac{r^2}{R^2} \right)^2 - \frac{1}{16} \left(\frac{r^2}{R^2} \right)^3 + \dots \right) \\
 s &= R \left(\frac{1}{2} \left(\frac{r^2}{R^2} \right) + \frac{1}{8} \left(\frac{r^2}{R^2} \right)^2 + \frac{1}{16} \left(\frac{r^2}{R^2} \right)^3 + \dots \right) \\
 s &= \frac{r^2}{2R} + \frac{r^4}{8R^3} + \frac{r^6}{16R^5} + \dots
 \end{aligned} \tag{4}$$

Now that we have surface sag (s) as a function of r and R, we can place 's' from equation 4 into equation 3.

$$\begin{aligned}
 OPD &= -s * (n_{lens} - 1) \\
 OPD &= - \left(\frac{r^2}{2R} + \frac{r^4}{8R^3} + \frac{r^6}{16R^5} + \dots \right) * (n_{lens} - 1)
 \end{aligned} \tag{5}$$

This OPD can also be written as :

$$OPD = - \left(\frac{x^2 + y^2}{2R} + \frac{(x^2 + y^2)^2}{8R^3} + \frac{(x^2 + y^2)^3}{16R^5} + \dots \right) * (n_{lens} - 1) \tag{6}$$

Optical Path Difference is directly translatable to wavefront error by dividing OPD by wavelength – thus giving the difference in terms of number of waves.

This OPD has many terms, the contribution of which keeps on increasing as r increases (Higher r implies larger pupil, also implies larger numerical aperture, thus faster optical system, and smaller $F/\#$ -> all of these are thus synonymous to increasing wavefront aberrations.

The general ray shown travels less distance in the higher refractive index medium (slower medium) than the chief ray. Thus, at any instant, this ray is ahead of the chief ray. This implies - to compensate for this, if it ever intends to reach the focus at the same time, the general ray must travel greater distance. The first term of OPD – quadratic term – is a perfect imaging condition. If a general ray accumulates only this much difference in optical path (in comparison to the central ray), it compensates for it in its path, thereby coming a to focus. **Fig 4** shows the OPL accumulated by two rays – one along the optical axis and other along the ray hitting lens at point r distance from the vertex. The two rays accumulate the following optical length after the refraction. The net OPL accumulated by ray OA is also the distance between the vertex of the lens (lens intersection with the optical axis) and the paraxial focal point i.e., paraxial focal length.

$$OPL_{OA} = f$$

$$OPL_{O'A} = \sqrt{(r^2 + f^2)} = f \sqrt{\left(1 + \frac{r^2}{f^2}\right)} = f + \frac{r^2}{2f} + \dots$$

So for the two rays to come to perfect focus, the difference in OPL should be negligible.

$$\text{Net Optical Path Difference} = \text{OPD} + OPL_{O'A} - OPL_{OA} = 0$$

$$-\left(\frac{r^2}{2R}\right) * (n_{lens} - 1) + f + \frac{r^2}{2f} - f = 0$$

$$-\left(\frac{r^2}{2R}\right) * (n_{lens} - 1) + \frac{r^2}{2f} = 0$$

$$-\left(\frac{r^2}{2R}\right) * (n_{lens} - 1) = -\frac{r^2}{2f}$$

$$f(\lambda) = \frac{R}{(n_{lens}(\lambda) - 1)}$$

This is indeed the thin lens formula for the focal length of a lens with one powered surface.

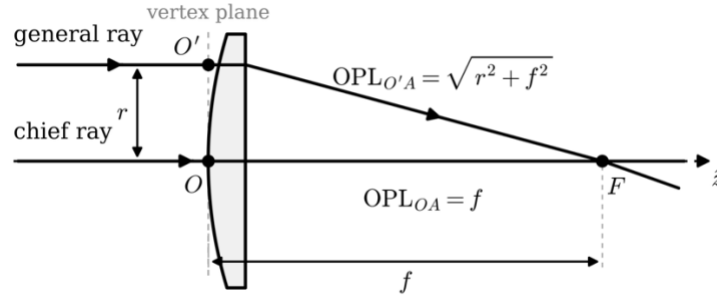


Figure 4: Image to show the extra path length covered by the marginal ray in comparison to the chief ray. The path length difference due to glass is not shown here but present in equations.

Refractive index of a material is a function of wavelength – this is also known as dispersion. So, for different wavelengths, the focal point would be different (due to different refractive index). This is what leads to **axial chromatic aberration**. This equation also tells that focal point is inversely proportional to delta in refractive index (which makes sense as bigger difference in refractive index of the two mediums imply more ray bending implying converging faster/closer). This also shows when blue wavelength travelling in air hits a glass lens (lower wavelength, thus higher refractive index, thus higher delta in refractive index of air and glass) focusses earlier than red ray.

$$\text{Chromatic Axial Aberration} = f_{\lambda_2} - f_{\lambda_1} = \frac{R}{(n_{\text{lens}}(\lambda_2) - 1)} - \frac{R}{(n_{\text{lens}}(\lambda_1) - 1)}$$

$$\text{Chromatic Axial Aberration} = \frac{R(n_{\text{lens}}(\lambda_1) - n_{\text{lens}}(\lambda_2))}{(n_{\text{lens}}(\lambda_2) - 1) * (n_{\text{lens}}(\lambda_1) - 1)}$$

$$\text{Chromatic Axial Aberration} = \frac{(n_{\text{lens}}(\lambda_1) - n_{\text{lens}}(\lambda_2))}{RC_1C_2}$$

To recap - the extra length accumulated is compensated by the OPD between the two rays due to the lens and its sag. Thus, if only first term in OPD equation was application, perfect focus was possible. But that would be possible only if $r \ll R$ such that higher powers of r/R (and r/f) can be ignored.

Now moving back to the full description of the optical path difference as shown in equation 5.

In the full description, spherical surface creates a sag that lead to optical path difference leading to aberration – called as spherical aberration. The second term $\frac{r^4}{8R^3} * (n - 1)$ is the primary spherical aberration.

$$W_{hkl} = -\frac{r^4}{8R^3} * (n - 1)$$

Here, {hkl} are used to represent the wavefront aberration type and its dependence. 'h' represents the dependence on the field height (remember till now we are only talking about the on-axis field with the chief ray intersecting the lens surface and the aperture at the same $y = 0$), 'k' represents the dependence on the pupil radius (r), 'l' represents the dependence on the angle $\cos\varphi$.

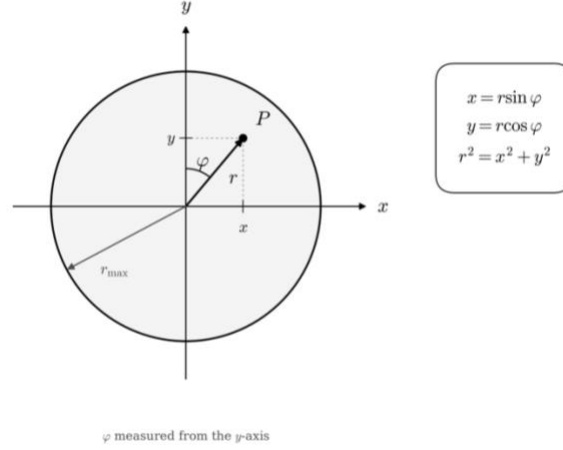


Figure 5: Relating the cartesian coordinates to polar coordinates

The radius 'r' represents a point in the x-y plane given by (x,y) as shown in in Fig 5. Here x and y can also be represented in the polar coordinates {r, φ }.

$$x = r \sin \varphi$$

$$y = r \cos \varphi$$

So for the primary spherical aberration $h = 0$, $k = 4$, $l = 0$. And thus,

$$W_{040} = -\frac{r^4}{8R^3} * (n - 1)$$

In this aberration, one can see that there is a dependence on 'r', but there is also a dependence on system parameters like R and n.

If we want to eliminate the primary spherical aberration, one way is to reduce the r -> This is known as the stopping down approach or reducing the numerical aperture or entrance beam diameter or increasing the F/# or making the lens slower. Or we can also change the curvature of the lens given by $c = \frac{1}{R}$. This is

done by changing the surface of the lens using a conic (k). When a conic is added, the primary wavefront aberration changes in the following way.

$$W_{040} = -\frac{(1+k)r^4}{8R^3} * (n-1) = -\frac{(1+k) * c^3 * r^4}{8} * (n-1)$$

The intention behind adding conic is to reduce the primary spherical aberration. The net OPD becomes –

$$OPD = -\left(\frac{r^2}{2R} + \frac{(1+k) * r^4}{8R^3} + \frac{r^6}{16R^5} + \dots\right) * (n_{lens} - 1)$$

There are more terms visible in this OPD which are higher order spherical aberration terms, active as we go far from the optical axis.

$$W_{060} = -\frac{r^6}{16R^5} * (n-1)$$

Stopping down would also reduce these higher-order spherical aberrations. Just like conic talked about earlier, aspherizing lens surface helps in eliminating these aberrations.

$$W_{060} = -\frac{(1+A_6)r^6}{16R^5} * (n-1)$$

The conic parameter ‘k’ is a special case of aspherizing and can also be represented by A_4 . The net OPD for on-axis beam thus becomes:

$$OPD = -\left(\frac{r^2}{2R} + \frac{(1+k) * r^4}{8R^3} + \frac{(1+A_6) * r^6}{16R^5} + \dots\right) * (n_{lens} - 1)$$

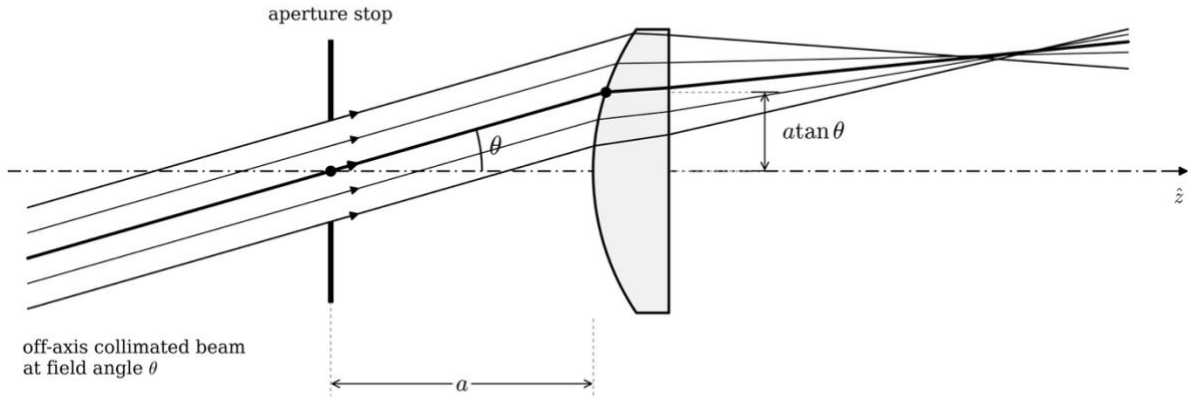


Figure 6: A plano-convex lens with an off-axis object field. Here $a \tan \theta$ represents the chief ray height at the lens interface – due to the stop being placed at some distance from the lens surface.

Now let's move on to off-axis fields (Refer Fig 6). Whether the beam is on axis or off axis, the sag equation of an optical surface remains same. As mentioned earlier, the chief ray of an off axis beam much like on axis beam intersects the aperture stop at the optical axis. Because of this, the chief ray of an off-axis beam would strike the lens surface distance away from the optical axis (this part is different from the on-axis beam). If the beam is at an angle θ from the optical axis,

$$y_0 = a * \tan \theta$$

Where a is the distance between the aperture stop and the optical element / lens surface along the optical axis.

The ' r ' in the sag equation is the r from the optical axis while the r used in the OPD equation is from the chief ray of the beam. These two quantities are same for the on-axis beam but not for the off axis beam. We know the chief ray of the beam is at $y_0 \hat{j}$ from the optical axis. So the new r that goes into the sag equation is

$$\mathbf{r} = \mathbf{r} + y_0 \hat{j}$$

$$\mathbf{r} = (x\hat{i} + (y + y_0)\hat{j})$$

So now the OPD (equation 5) becomes,

$$OPD = -\left(\frac{r^2}{2R} + \frac{r^4}{8R^3} + \frac{r^6}{16R^5} + \dots\right) * (n_{lens} - 1)$$

$$OPD = -\left(\frac{(x^2 + (y + y_0)^2)}{2R} + \frac{((x^2 + (y + y_0)^2)^2)}{8R^3} + \frac{(x^2 + (y + y_0)^2)^3}{16R^5} + \dots\right) * (n_{lens} - 1)$$

Let's concentrate on the fourth order in r term:

$$OPD = -\left(\frac{((x^2 + (y + y_0)^2)^2)}{8R^3}\right) * (n_{lens} - 1)$$

$$OPD = -\left(\frac{(x^4 + (y + y_0)^4 + 2 * x^2 * (y + y_0)^2)}{8R^3}\right) * (n_{lens} - 1)$$

$$OPD = -\left(\frac{(x^4 + (y + y_0)^4 + 2 * x^2 * (y + y_0)^2)}{8R^3}\right) * (n_{lens} - 1)$$

$$OPD = -\left(\frac{(x^4 + y^4 + 4y^3y_0 + 6y^2y_0^2 + 4yy_0^3 + y_0^4 + 2x^2y^2 + 2x^2y_0^2 + 4x^2yy_0)}{8R^3}\right) * (n_{lens} - 1)$$

$$OPD = -\left(\frac{((x^2 + y^2)^2 + 4y^3y_0 + 6y^2y_0^2 + 4yy_0^3 + y_0^4 + 2x^2y_0^2 + 4x^2yy_0)}{8R^3}\right) * (n_{lens} - 1)$$

$$OPD = -\left(\frac{((x^2 + y^2)^2 + y_0(4y^3 + 4x^2y) + 2y_0^2 * (3y^2 + x^2) + 4yy_0^3 + y_0^4)}{8R^3}\right) * (n_{lens} - 1)$$

$$OPD = -\left(\frac{((x^2 + y^2)^2 + 4y_0(x^2 + y^2)y + 2y_0^2 * (y^2 + x^2 + 2y^2) + 4yy_0^3 + y_0^4)}{8R^3}\right) * (n_{lens} - 1)$$

$$OPD = - \left(\frac{(x^2 + y^2)^2}{8R^3} - \frac{y_0(x^2 + y^2)y}{2R^3} - \frac{y_0^2 * (y^2 + x^2)}{4R^3} - \frac{y_0^2 * (y^2)}{2R^3} - \frac{4yy_0^3}{2R^3} - \frac{y_0^4}{8R^3} \right) * (n_{lens} - 1)$$

$$OPD = \left(-\frac{(r^2)^2}{8R^3} - \frac{y_0(r^2)r \cos \varphi}{2R^3} - \frac{y_0^2 * (r^2)}{4R^3} - \frac{y_0^2 * (r^2 \cos \varphi^2)}{2R^3} - \frac{(r \cos \varphi)y_0^3}{2R^3} - \frac{y_0^4}{8R^3} \right) * (n_{lens} - 1) \quad (7)$$

Separating the terms in equation 7:

1) Primary Spherical Aberration

$$W_{040} = -\frac{(r^2)^2}{8R^3} * (n_{lens} - 1)$$

2) Coma

$$W_{131} = -\frac{y_0 r^3 \cos \varphi}{2R^3} * (n_{lens} - 1)$$

3) Astigmatism

$$W_{222} = -\frac{y_0^2 * (r^2 \cos \varphi^2)}{2R^3} * (n_{lens} - 1)$$

4) Field Curvature

$$W_{220} = -\frac{y_0^2 * (r^2)}{4R^3} * (n_{lens} - 1)$$

5) Distortion

$$W_{311} = -\frac{(r \cos \varphi y_0^3)}{2R^3} * (n_{lens} - 1)$$

Therefore, the same order term that was giving spherical aberration for on-axis gives all the fourth order wavefront aberrations for off-axis fields.

Now let's move on to how knowing the dependence of these wavefront aberrations can help us in understanding how to eliminate them.

Spherical aberration, for instance, depends on R (refracting surface radius of curvature). This makes **bending the lens** (changing R_1, R_2 for two surfaces of the lens while keeping the focal length constant) a way to reduce the spherical aberration. Spherical Aberration also depends on r (pupil radius) – this makes stopping down the lens (reducing the aperture size) a way to reduce spherical aberration. **Stopping down the lens** can be used to reduce all the above mentioned aberrations due to their dependence on ' r '. Note, however, their dependence is different as a function of power. In addition, as already seen before, adding a conic (k) or aspherizing (A_6, A_8 etc.) can also reduce the spherical aberration term to zero.

Coma, Astigmatism, FC, Distortion are all dependent on y_0 . Recall that y_0 came into being due to the chief rays coming from an off-axis point intersecting the lens surface at a distance from optical axis. This happened because the aperture stop was at a distance from the lens (and the chief ray intersects the stop at the optical axis). This makes **stop shift** - moving the aperture stop such that y_0 changes another way to reduce all the off-axis aberrations. Dependence on y_0 also means that on-axis fields would not have Coma, Astigmatism, FC, and Distortion.

Also note that Coma and Distortion are odd powers of y_0 . This shows an interesting point. If multiple elements are added that are **symmetrically** placed around the aperture stop such that the y_0 is positive and negative for two same lenses sitting on either side of the stop. Then the net coma contribution would become zero for this system. Similarly, the net distortion of this system would also become zero.

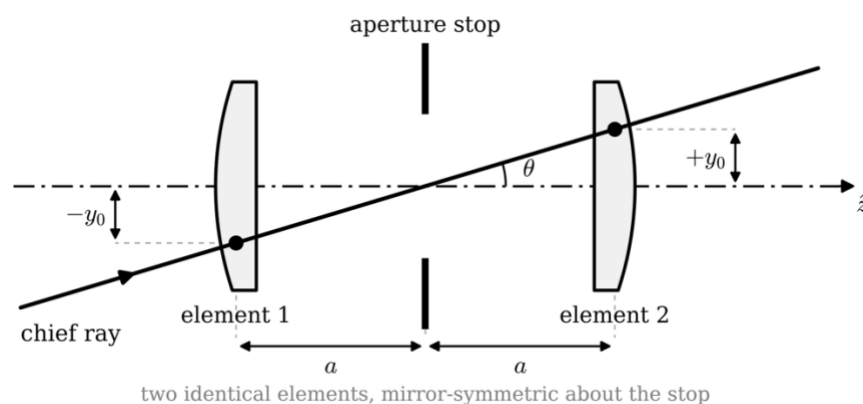


Figure 7: The effect of symmetry on Odd powered aberrations via effect on y_0 . Note that the chief ray is shown as not refracting for easy visualization of y_0 . However, the concept is valid even when ray bending happens.

Conclusion

In conclusion, this paper derives the wavefront aberration from the simplest starting point of a refracting lens surface. Starting from the surface of the lens and the sag equation, optical path lengths travelled by various rays can be compared and their difference can be calculated leading to wavefront aberrations of all types. Knowledge and understanding of the origin of various wavefront aberrations can aid in understanding the well-established ways used to eliminate aberrations as well as in predicting new ways to eliminate the aberrations.

Acknowledgements

All the figures were generated using Claude Code.