

# Using ray matrices to derive analytical expressions of optical aberrations

Ignacio Moreno, Pascuala García-Martínez and Carlos Ferreira

1. Departamento de Ciencia de Materiales, Óptica y Tecnología Electrónica, Universidad Miguel Hernández de Elche, 03202 Elche (Alicante), Spain.
2. Departament d'Òptica, Universitat de València, 45100 Burjassot (Valencia), Spain.

## ABSTRACT

In this work we show an easy way to derive approximated analytical expressions for spherical aberration. It is based on the use of  $3 \times 3$  ray matrices, where additional elements are included to account for the errors in the ray height and angle propagation. The technique, which is commonly used to analyze misaligned systems in optical resonators, is here applied to account for deviations from the paraxial regime caused by spherical aberration.

## 1. INTRODUCTION

Ray matrix formalism is a useful theory applied to paraxial geometrical optics, which employs  $2 \times 2$  matrices to describe optical systems [1,2]. This formulation allows complex, multi-element, optical systems to be analyzed and simplified using easy calculations. This matrix theory is also very useful for teaching Optics, since it is well-valued by students because the related mathematics are simple. This formalism is included in many textbooks devoted to Geometrical Optics, although usually included in a final stage, as a practical tool for numerical calculations. However, it has been demonstrated to be very useful also to develop analytical derivations. For instance, in Refs. [3,4] we presented a full derivation of Fourier Optics systems based on this matrix formalism. It is also commonly employed to analyze stability conditions in optical resonators [5], as well as its connection with the fractional Fourier transform [6].

From the practical point of view, the paraxial approximation is not fulfilled and “perfect” imaging are not obtained. Aberration theory is required to complete the description. This subject is included in advanced Optics courses, usually defined in terms of the distortion of the ideal spherical wavefront, named the wavefront aberration function,  $W$ . The aberration function is well defined and all primary aberrations are expressed in terms of it. However this aberration function theory can be quite annoying for students since the polynomial expressions are large, and analytical expressions are difficult to derive. Following our previous works [3,6], we believe that the application of a common tool, as the ray matrix theory, to teach different modules of an Optics course permits the students to achieve a deeper understanding of the practical aspects of its application, and provides a uniform and elegant framework. It is our purpose here to add an additional step in this sequence, and use the ray matrix formalism to explain aberrations concepts.

Spherical aberration is the most common aberration in optical imaging. Thus, we have first attempted to derive related analytical expressions using the ray matrix theory. We analyze non-paraxial rays, where an axial object point is not imaged onto a single point. Instead rays at different heights focus at different axial locations. Generalized ray transfer matrices that account for aberration were introduced in Ref. [7]. There, two  $2 \times 2$  matrices were required, one for the radial component, and another for the azimuthal component. Here, however, we use the extended  $3 \times 3$  ray matrix method introduced in the field of laser resonators [5]. This extended method is based on the introduction of two parameters,  $\Delta x$  and  $\Delta \sigma$ , which account respectively for possible errors in the height and angle coordinates of a ray traversing an optical system. While this method has been used to analyze errors caused by misaligned optical elements, we show next that it can also be applied to account for spherical aberration in optical systems. For that purpose, we consider these error parameters produced by the deviation of the ray with respect to the paraxial solution. Using this mathematical treatment we are able to derive, in a simple manner, analytical equations describing the longitudinal (LSA) or axial (ASA) spherical aberration.

The paper is organized as follows: in section 2 we briefly review the usual ray matrix method, as well as the extended 3×3 matrix. Its application to account for spherical aberration is introduced in Section 3. The conclusions are presented in Section 4.

## 2. PARAXIAL AND EXTENDED RAY MATRICES

The ray matrix method considers rotationally symmetric optical systems under the paraxial approximation [1,2]. The optical system is regarded as a set of optical components placed between two traverse planes, located at  $z=z_1$  and  $z=z_2$  (Fig 1). The paraxial approximation implies that angular coordinates ( $\sigma$ ) follow the small angle approximation and it can be considered as the ray slope,  $\sigma=dx/dz$ . The optical system changes the position and the angle of the input ray. An input ray at the incidence plane, with coordinates  $(x_1, \sigma_1)$ , is transformed into an output ray with coordinates  $(x_2, \sigma_2)$ . In the paraxial approximation, the relations among these coordinates are linear, and they can be written in the form:

$$\begin{pmatrix} x_2 \\ \sigma_2 \end{pmatrix} = \mathbf{M} \begin{pmatrix} x_1 \\ \sigma_1 \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} x_1 \\ \sigma_1 \end{pmatrix}. \quad (1)$$

$\mathbf{M}$  is the ray matrix, with ABCD components. Ray matrices corresponding to the some basic blocks for building optical systems, like a free propagation (FP), a spherical refractive surface (R), and a thin lens (TL), are respectively:

$$\mathbf{M}_{FP}(d) = \begin{pmatrix} 1 & d \\ 0 & 1 \end{pmatrix}, \quad \mathbf{M}_R(n, n', R) = \begin{pmatrix} 1 & 0 \\ -\frac{n'-n}{n'R} & 1 \end{pmatrix}, \quad \mathbf{M}_{TL}(f') = \begin{pmatrix} 1 & 0 \\ -\frac{1}{f'} & 1 \end{pmatrix}, \quad (2)$$

being  $d$  the propagation distance,  $n$ ,  $n'$  and  $R$  the refractive indices of input and output media, and the radius of curvature of the refractive surface, respectively and  $f'$  the lens focal length.

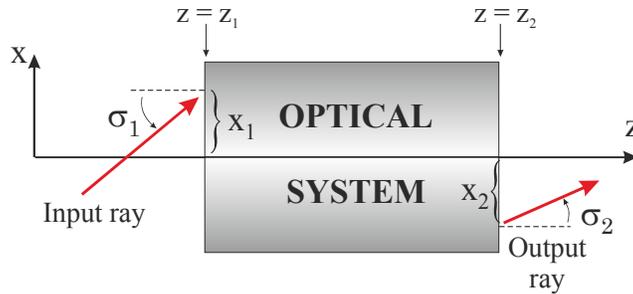


Fig. 1. Representation of a first order optical system by means of a ray transfer matrix.

The extended ray matrix method was proposed to account for misalignments in laser resonators [5]. Possible decentring and/or tilt of the optical elements were introduced as two error parameters,  $\Delta x$  and  $\Delta \sigma$ , in the height and the angle coordinates displacements of the emerging ray. Therefore, Eq. (1) must be transformed to:

$$\begin{pmatrix} x_2 \\ \sigma_2 \\ 1 \end{pmatrix} = \begin{pmatrix} A & B & \Delta x \\ C & D & \Delta \sigma \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ \sigma_1 \\ 1 \end{pmatrix}. \quad (3)$$

which introduces the extended 3×3 ray matrix, which includes the standard ray matrix as the top-left 2×2 submatrix. The advantage of such method becomes clear when cascading multiple misaligned elements in optical systems. So, a whole

optical system can be treated as a group of elements, as it is usual in this formalism, by multiplying the corresponding optical element matrices. The final 3×3 matrix provides the information of the propagated error in the output ray coordinates as the elements  $\Delta x$  and  $\Delta \sigma$  in its third column.

### 3. APPLICATION TO SPHERICAL ABERRATION

Here we apply the above 3×3 ray matrix method to analyze spherical aberration. While this aberration can not be considered as a misalignment (perfectly centered elements present spherical aberration), we consider that that rays are affected by coordinate errors from paraxial regime due to the aberration. Let us consider Fig. 2(a), which represents a spherical surface between two media with refractive indices  $n$  and  $n'$ , respectively and radius of curvature,  $R$ .  $S$  denotes the axial point on the spherical surface, and  $C$  its center of curvature.  $F'_0$  is the paraxial focus, located at the focal distance  $f'_0$  from  $S$ . This distance is given by the classical geometrical optics formula derived from the first order (paraxial) approximation as:

$$f'_0 = \frac{n'-n}{n'R}. \quad (4)$$

However, due to the spherical aberration, a parallel ray impinging the surface at height  $x$  will not focus on  $F'_0$  but on another axial point  $F'_x$  (the true and paraxial ray trajectories are the blue and red lines in Fig. 2(a), respectively). The equation giving this focus point is probably the simplest derivation in third order approximation, and the result is given by [8]:

$$\frac{1}{f'_x} = \frac{1}{f'_0} \left( 1 + \frac{n^2 x^2}{2n'^2 R^2} \right). \quad (5)$$

where  $f'_x$  denotes the distance from  $S$  to  $F'_x$ . We can regard the deviation of the real ray trajectory from the paraxial ray in terms of the error parameters in the 3×3 ray matrix formalism. For simplicity, we can approximate  $\Delta x \approx 0$  and consider that all errors are produced in the angular coordinate  $\Delta \sigma$ . Considering Eqs. (4) and (5), this angular deviation is given by:

$$\Delta \sigma = \sigma_x - \sigma_0 = -x \left( \frac{1}{f'_x} - \frac{1}{f'_0} \right) = -\frac{n^2 (n'-n)}{2n'^3 R^3} x^3. \quad (6)$$

Using this simple result, we can derive some properties about the spherical aberration of a lens. Let us consider a TL in air, composed by two spherical refractions of curvatures  $R_1$  and  $R_2$ , and a medium of index  $N$  (Fig. 2(b)). We have considered the extended angular error matrix and we have assumed a thin lens as a successive optical elements arranged in cascade having different degree of angular errors in Eq. (6). Then, the matrix describing the TL is given by:

$$\mathbf{M}_{TL} = \begin{pmatrix} 1 & 0 & 0 \\ -\frac{1}{f'_2} & N & \Delta \sigma_2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ -\frac{1}{f'_1} & \frac{1}{N} & \Delta \sigma_1 \\ 0 & 0 & 1 \end{pmatrix}. \quad (7)$$

where

$$f'_1 = \frac{N-1}{NR_1}, \quad f'_2 = \frac{1-N}{R_2}, \quad \Delta \sigma_1 = \frac{(1-N)}{2N^3 R_1^3} x^3, \quad \Delta \sigma_2 = \frac{N^2(N-1)}{2R_2^3} x^3. \quad (8)$$

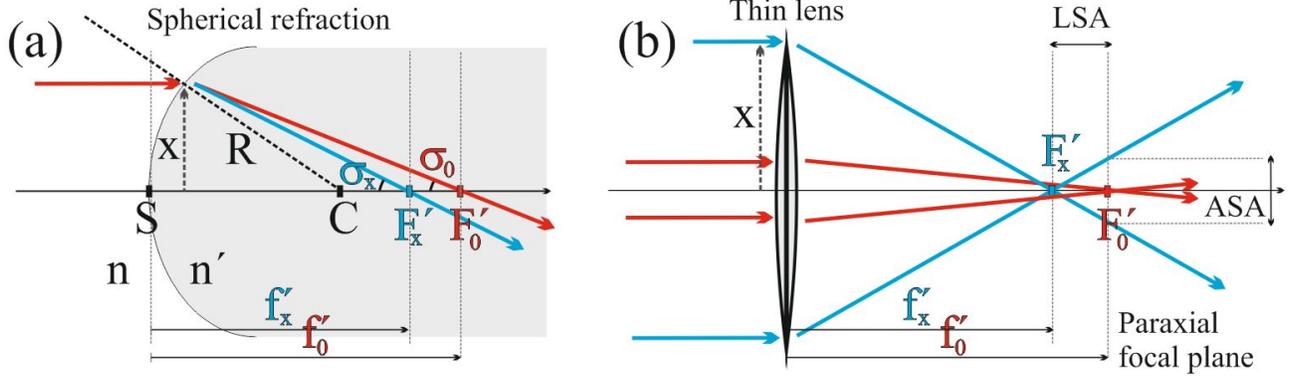


Fig. 2. Representation of the focalization of a parallel ray in (a) a spherical refractive surface, (b) a thin lens.

The matrix product in Eq. (7) results in:

$$\mathbf{M}_{TL} = \begin{pmatrix} 1 & 0 & 0 \\ -(N-1)\left(\frac{1}{R_1} - \frac{1}{R_2}\right) & 1 & \Delta\sigma \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ -\frac{1}{f'_0} & 1 & \Delta\sigma \\ 0 & 0 & 1 \end{pmatrix}. \quad (9)$$

which recovers the TL ray matrix in Eq. (2), including the lens maker formula that provides the lens paraxial focal length,  $1/f'_0 = (N-1)(1/R_1 - 1/R_2)$ . However, an TL angular deviation,  $\Delta\sigma$ , is also present in Eq. (9), given by:

$$\Delta\sigma = N\Delta\sigma_1 + \Delta\sigma_2 = \frac{N-1}{2} \left[ \frac{N^2}{R_2^3} - \frac{1}{N^2 R_1^3} \right] x^3 \equiv -\Omega x^3. \quad (10)$$

$\Delta\sigma$  depends on the height,  $x$ , to the third, with a proportional term  $\Omega$  defined in the equation. Now it becomes a simple ray matrix exercise (free propagation + TL (Eq.(6) matrix) to find the distance  $f'_x$  where a parallel ray ( $\sigma_1=0$ ) impinging the thin lens at height,  $x$ , focalize in the axis. The obtained result is:

$$f'_x = \frac{x}{\frac{x}{f'_0} - \Delta\sigma} = \frac{f'_0}{1 + \Omega f'_0 x^2} \cong f'_0 - \frac{1}{2} \Omega f'_0 x^2, \quad (11)$$

where we have approximated the denominator as  $(1 + ay^2)^{-1} \cong 1 - \frac{1}{2} ay^2$ , being  $y = x^2 \ll 1$  and  $a = \Omega f'_0$ . Then Eq. (11) leads to the following relation of the longitudinal spherical aberration (Fig. 2(b)):

$$LSA = f'_0 - f'_x = \frac{1}{2} \Omega f'_0 x^2. \quad (12)$$

This shows the LSA second order dependence with the height  $x$ .

Analogously, we can propagate the ray emerging from the lens until the paraxial focal plane located at distance  $f'_0$ , and calculate its height, providing a measure of the axial spherical aberration (ASA). The result is:

$$\text{ASA} = f'_0 \Delta\sigma = -\Omega f'_0 x^3, \quad (13)$$

which shows its third order dependence with  $x$ .

#### 4. CONCLUSIONS

In summary, we have introduced a new way to treat spherical aberration in terms of the extended  $3 \times 3$  ray matrix method. We believe this is a powerful tool to derive related analytical expressions which otherwise require very extensive derivations. Therefore it is a useful tool to teach aberrations to students who are familiar with the standard ray matrix method. For example, the strict derivation of the last two results (dependence of LSA and ASA with  $x^2$  and  $x^3$ , respectively) is rather annoying, and it is usually avoided in texts devoted to Optics. Even texts [8] which include chapters specifically devoted to this issue, present the result without a complete derivation. The ray matrix method proposed here constitutes a formal framework familiar to students in Optics, and it provides a relatively short derivation of these relations. We believe the method can be extended to obtain other related results.

#### ACKNOWLEDGEMENTS

We acknowledge financial support from Spanish *Ministerio de Economía y Competitividad*, through projects FIS2012-39158-C02-02 and FIS2010-16646, and by *Fondo Europeo de Desarrollo Regional* (FEDER).

#### REFERENCES

1. A. Gerrard and J. M. Burch, *Introduction to matrix methods in Optics*. Dover Publications, New York (1975).
2. G. Kloos, *Matrix Methods for Optical Layout*, SPIE Press, Bellingham, WA (2007).
3. I. Moreno, M. M. Sánchez-López, C. Ferreira, J. A. Davis, F. Mateos, "Teaching Fourier optics through ray matrices", *European Journal of Physics* 26, 261-271 (2005).
4. I. Moreno, C. Ferreira, "Fractional Fourier transform and Geometrical Optics", Chapter 3 in *Advances in Imaging and Electron Physics*, Vol. 161, 89-146 (2010).
5. A. E. Siegman, *Lasers*, University Science Books (1986).
6. I. Moreno, P. García-Martínez, C. Ferreira, "Teaching stable two-mirror resonators through the fractional Fourier transform", *European Journal of Physics* 31, 273-284 (2010).
7. T. M. Jeong, D.-K. Ko, J. Lee, "Generalized ray-transfer matrix for an optical element having an arbitrary wavefront aberration", *Optics Letters* 30, 3009-2011 (2005).
8. F. A. Jenkins, H. E. White, *Fundamentals of Optics*, McGraw-Hill (1981).