



High-precision analysis of aberration contribution of Zernike freeform surface terms for non-zero field of view

SHUAI ZHANG,¹ XING ZHAO,^{1,2,*} DA LI,^{1,2}  HAO FENG,¹ SHANGNAN ZHAO,^{3,4,5,6} LINGJIE WANG,^{3,5,6} AND XIN ZHANG^{3,5,6}

¹Institute of Modern Optics, Eye Institute, Nankai University, Tianjin 300350, China

²Tianjin Key Laboratory of Micro-scale Optical Information Science and Technology, Tianjin 300350, China

³Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, Changchun 130033, China

⁴University of the Chinese Academy of Sciences, Beijing 100039, China

⁵State Key Laboratory of Applied Optics, Changchun 130033, China

⁶Key Laboratory of Optical System Advanced Manufacturing Technology, Chinese Academy of Sciences, Changchun 130033, China

*shaoxingtjnk@nankai.edu.cn

Abstract: Clarifying the aberrations arising from freeform surfaces is of great significance for maximizing the potential of freeform surfaces in the design of optical systems. However, the current precision in calculating aberration contribution of freeform surface terms for non-zero field of view is insufficient, impeding the development of freeform imaging systems with larger field of view. This paper proposes a high-precision analysis of aberration contribution of freeform surface terms based on nodal aberration theory, particularly for non-zero field points. Accurate calculation formulas of aberrations generated by Zernike terms on freeform surface are presented. Design examples illustrate that the calculation error of the provided formulas is 78% less than that of conventional theoretical values. Building upon high-precision analysis, we propose an optimization method for off-axis freeform surface systems and illustrate its effectiveness through the optimization of an off-axis three-mirror system. This research extends the applicability of nodal aberration theory in aberration analysis, offering valuable insights for the optimal design and alignment of optical freeform systems.

© 2024 Optica Publishing Group under the terms of the [Optica Open Access Publishing Agreement](#)

1. Introduction

Freeform surfaces offer a heightened degree of design flexibility and an enhanced capacity for controlling light, resulting in improved performance and compactness of optical systems [1]. The advancement of desensitization design methods [2] and cutting-edge manufacturing techniques [3] has greatly promoted the extensive incorporation of freeform surfaces in a range of optical systems, such as spectrometers [4], head-mounted displays [5–7], three-mirror imaging systems [8–10], and projection optics [11].

Freeform surface shapes are commonly represented by analytic functions and polynomials [12]. Among these, Zernike polynomials stand out for their orthogonality and completeness within the unit circle. The inherent quality renders Zernike polynomials highly valuable for characterizing surface shape and wavefront aberrations in optical design and testing. Additionally, the advancement of nodal aberration theory (NAT) has greatly facilitated the aberration analysis of optical imaging systems employing Zernike freeform surfaces. NAT was introduced by Shack and Thompson [13], building upon H.H. Hopkins' wavefront aberration theory and Buchroeder's concept of surface aberration field contributions [14]. NAT enables the analysis of aberrations in optical systems without rotational symmetry. Thompson extended NAT to third and fifth orders

[15–18], unveiling the multi-nodal aberration behavior in misaligned systems. Fuerschbach integrated NAT with Zernike freeform surfaces, presenting an analytical framework for aberration calculation [19], referred to here as the Aberration Calculation Framework (ACF). ACF is utilized to calculate aberration contributions of Zernike terms on freeform surfaces, whether located at the stop of the optical system or away from it. This framework not only quantifies aberrations but also uncovers their field-dependent characteristics. On the basis of ACF, Yang conducted comprehensive calculations of aberrations caused by Zernike terms on refractive and reflective surfaces in coaxial imaging systems [20]. Furthermore, Yang employed the scaled pupil vector to address the issue of elliptical footprints on tilted freeform surfaces. Experiment results show a significant improvement in predicting the aberration contribution of off-axis Zernike freeform surfaces terms at 0° field of view (FOV) [21]. Ju incorporated a scaled factor of the beam footprint into ACF and introduced modified explicit formulas, establishing the relationships between the magnitude of freeform surfaces and the resulting aberrations. Subsequently, Ju accurately calculated the figure error based on the measured wavefront aberrations at paraxial field points (0° and 0.07°) [22]. In addition, the quantitative evaluation of aberrations resulting from Zernike terms has significantly contributed to the advancement of research in optical design and alignment. It has supported the development of design methodologies for freeform surface systems [10], the calculation of aberrations due to mount-induced surface deformation [23], and the formulation of alignment strategies in the presence of figure errors [24,25].

The existing ACF has demonstrated its ability to accurately predict aberrations at paraxial field points generated by freeform surface terms. Nevertheless, optical imaging systems typically encompass a certain width of FOV, and it is equally essential to ensure satisfactory image quality at non-zero field points. Therefore, accurately calculating the aberration contributions of freeform surface terms for non-zero field points is imperative as well. However, due to the differences between the beams from zero field point and non-zero field points, such as the angle of incidence and the shape of the exit pupil, the existing ACF may exhibit large calculation errors when predicting the aberrations at non-zero field points, both in type and magnitude. Unfortunately, at present, there is a lack of precise analysis regarding the aberration contribution of freeform surface terms for general non-zero FOV, either in coaxial or off-axis optical systems. This limitation presents challenges in optical design by impeding the efficient identification of freeform terms that should be optimized based on the characteristics of non-zero-field aberrations. As a result, the optimization design process may heavily depend on design experience. Moreover, in optical alignment, this limitation hinders the accurate separation of surface figure errors from misalignments based on the aberrations of multiple field points, leading to a significant reduction in compensation efficiency. Hence, it is crucial to accurately calculate the aberration contribution of freeform terms for non-zero field points.

This paper presents a highly precise analysis of aberration contribution of Zernike freeform surface terms for non-zero field points, surpassing the accuracy of existing theories. We begin by an overview of the current analyses about the Zernike terms' aberration contribution for paraxial field of view, so as to establish the theoretical framework for subsequent sections. Next, a detailed discussion is conducted on the aberration contribution of Zernike terms on freeform surfaces located at the stop and away from the stop, within non-zero FOV. By introducing the chief ray's incident angle and the shape of the fitting exit pupil, we presented the precise aberration calculation formulas and substantiate their accuracy through design examples. When compared to traditional theoretical values, the proposed formulas exhibit a 78% reduction in calculation errors. Finally, we proposed a NAT-guided optimization method for off-axis freeform surface systems based on the high-precision analysis, and demonstrated its effectiveness by optimizing an off axis three-mirror system. This research is expected to aid in optical system design and alignment.

2. Aberration contribution of Zernike freeform surface terms for paraxial field of view

Zernike terms added to optical surface introduce optical path difference and generate wavefront aberration. The relationship between Zernike terms and the generated wavefront aberration W is represented as [21],

$$W = \frac{n' - n}{\lambda} \sum_i C_i Z_i(\vec{\rho}_f), \tag{1}$$

with

$$\vec{\rho}_f = (\rho_f, \varphi) = (\rho_{f,x}, \rho_{f,y}), \tag{2}$$

where λ is the wavelength, n and n' are the indices of refraction before and after freeform surface, especially for mirrors, $n = 1, n' = -1$ for odd times of reflection, and $n = -1, n' = 1$ for even times. $\vec{\rho}_f$ is the vector-form normalized aperture coordinate of the freeform surface that can be expanded in polar coordinates with radial component ρ_f and azimuthal component φ , or in Cartesian coordinates with horizontal component $\rho_{f,x}$ and vertical component $\rho_{f,y}$. C_i is the coefficient of the i th Zernike terms $Z_i(\vec{\rho}_f)$. $\sum_i C_i Z_i(\vec{\rho}_f)$ represents the surface sag generated by Zernike terms added to the surface, we refer to this as Zernike sag, denoted by $f(\vec{\rho}_f) = \sum_i C_i Z_i(\vec{\rho}_f)$. As the aberrations are generally fitted to the exit pupil plane, the normalized aperture vector of the freeform surface, $\vec{\rho}_f$, should be transformed to the normalized pupil vector, $\vec{\rho}_p$. The correspondence between $\vec{\rho}_f$ and $\vec{\rho}_p$ depends on the location of the freeform surface.

2.1. Aberration generated by Zernike terms on freeform surface located at stop

When the freeform surface is located at stop, as shown in Fig. 1, the light beams from different field points share the same area of the freeform surface, the entire surface aperture is covered by the beam footprint.

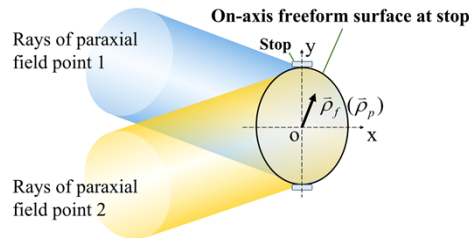


Fig. 1. The case when the freeform surface is located at stop, the rays from different field points share the same area.

Then we have,

$$\vec{\rho}_f = \vec{\rho}_p, \tag{3}$$

combine Eq. (1) and Eq. (3), the aberration generated by Zernike terms on freeform surface at stop can be calculated by,

$$W = \frac{n' - n}{\lambda} \sum_i C_i Z_i(\vec{\rho}_p). \tag{4}$$

In practice, the coordinate system of the surface located at stop is typically chosen as the reference coordinate system. This means that the surface located at stop is always considered on-axis. Therefore, the off-axis analysis exclusively focuses on the freeform surface located away from the stop.

2.2. Aberration generated by Zernike terms on freeform surface located away from stop

2.2.1. On-axis freeform surface

When the freeform surface is located away from stop, as shown in Fig. 2, the beam footprints of different field points cover different parts of the freeform surface, the aberration contribution of freeform to each field point differs and relies on sag of the specific irradiated surface area.

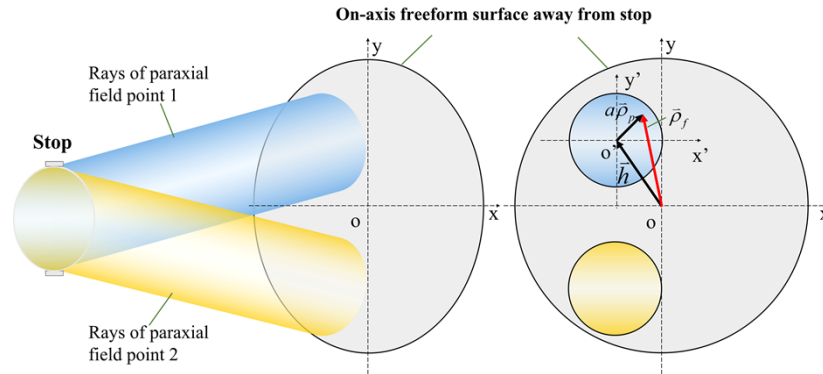


Fig. 2. The case when the freeform surface is located away from stop, the rays from different field points illuminate different areas.

In this case, a transformation from $\vec{\rho}_f$ to $\vec{\rho}_p$ is,

$$\vec{\rho}_f = a\vec{\rho}_p + \vec{h} = (a\rho_{f,x} + h_x, a\rho_{f,y} + h_y), \tag{5}$$

with scale factor $a = \frac{y}{R_n}$, where y denotes the marginal ray height on freeform surface, and R_n is the normalized radius of the freeform surface. \vec{h} is the pupil shift vector with horizontal component h_x and vertical component h_y , it defines the center of the shift beam footprint, that is,

$$\vec{h} = b\vec{H} = (h_x, h_y) = (bH_x, bH_y), \tag{6}$$

where $b = \frac{\bar{y}}{R_n}$, \bar{y} is chief ray height on freeform surface, \vec{H} represents the normalized field vector, consisting of horizontal component H_x and vertical component H_y . Combine Eq. (1) and Eq. (5), the aberration generated by Zernike terms on freeform surface away from stop can be calculated by,

$$W = \frac{n' - n}{\lambda} \sum_i C_i Z_i (a\vec{\rho}_p + \vec{h}). \tag{7}$$

2.2.2. Off-axis freeform surface

A surface is regarded as off-axis when decentered or tilted. In the case of a decentered surface, as illustrated in Fig. 3, the freeform surface is shifted from the dashed circle to the solid one, resulting in a shift of the surface origin, $\vec{\sigma}_f$.

In this case, the pupil shift vector \vec{h} in Eq. (6) is modified as $\vec{h} = b\vec{H} - \vec{\sigma}_f$, where $\vec{\sigma}_f = \frac{\vec{\delta}_f}{R_n}$, $\vec{\delta}_f$ is the physical surface displacement vector. By substituting the modified \vec{h} into Eq. (7), the aberration generated by Zernike terms on decentered freeform surface can be calculated, given

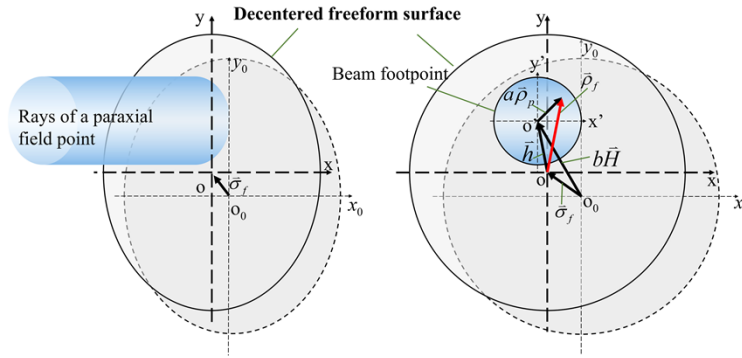


Fig. 3. Beam footprint on the decentered freeform surface that shift from the dashed boundary circle to the solid boundary circle

by Eq. (8).

$$W = \frac{n' - n}{\lambda} \sum_i C_i Z_i (a \vec{\rho}_p + b \vec{H} - \vec{\sigma}_f). \tag{8}$$

In practical cases, optical systems are typically designed plane-symmetric, so the surface is tilted by rotating only around the x or y axis. As illustrated in Fig. 4, when the beam is illuminated on a freeform surface tilted along the x-axis, as shown on the left side, the distribution of light beams on the surface is depicted on the right side. The effective beam footprint is not a circle but an approximate ellipse.

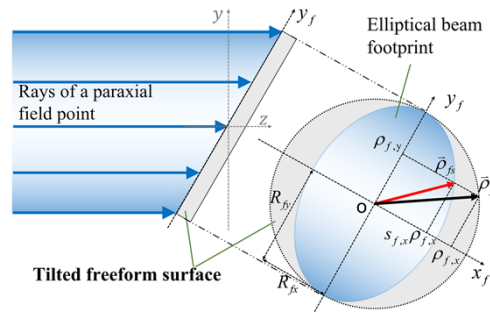


Fig. 4. The case when the beam is illuminated on a tilted freeform surface, the beam footprint is an approximate ellipse

In this case, the normalized aperture vector $\vec{\rho}_f = (\rho_{f,x}, \rho_{f,y})$ should change to the effective scaled aperture vector $\vec{\rho}_{fs} = (s_{f,x} \rho_{f,x}, \rho_{f,y})$. where the tilt factor $s_{f,x}$ is given by,

$$s_{f,x} = \frac{R_{fx}}{R_{fy}}, \tag{9}$$

R_{fx} and R_{fy} represent the marginal ray heights in the x-direction and y-direction on the freeform surface, respectively. Similarly, when the surface is tilted about the y-axis, $\vec{\rho}_{fs} = (\rho_{f,x}, s_{f,y} \rho_{f,y})$. where the tilt factor $s_{f,y} = \frac{R_{fy}}{R_{fx}}$. So according to Eq. (5), the relationship between the effective

scaled aperture vector $\vec{\rho}_{fs}$ and the pupil vector $\vec{\rho}_p$ is,

$$\vec{\rho}_{fs} = (s_{f,x}\rho_{f,x}, s_{f,y}\rho_{f,y}) = (s_{f,x}a\rho_{p,x} + h_x, s_{f,y}a\rho_{p,y} + h_y), \quad (10)$$

where

$$s_{f,x} = \frac{R_{fx}}{R_{fs}}, s_{f,y} = 1, \text{ when tilted around x - axis}$$

$$s_{f,x} = 1, s_{f,y} = \frac{R_{fy}}{R_{fs}}, \text{ when tilted around y - axis}$$

Combine Eq. (7) and Eq. (10), the aberration generated by Zernike terms on off-axis freeform surface can be calculated by,

$$W = \frac{n' - n}{\lambda} \sum_i C_i Z_i [s_{f,x}a\rho_{p,x} + h_x, s_{f,y}a\rho_{p,y} + h_y]. \quad (11)$$

When the freeform surface is on-axis, $s_{f,x} = s_{f,y} = 1$, Eq. (11) reverts to Eq. (7). Therefore, Eq. (11) is applicable when the surface is located away from the aperture stop, whether it is on-axis or off-axis.

The aberration calculation formulas, specifically Eq. (4), Eq. (7), and Eq. (11), relies on an key approximation assuming that rays propagate along the local Z-axis [20]. As shown in Fig. 5, the zero-field incident ray is reflected at pupil position $\vec{\rho}$ of the freeform mirror. When approximating ray reflection along the Z-axis, the theoretical optical path difference due to Zernike sag, $f(\vec{\rho})$, is $2f(\vec{\rho})$, but the actual reflected ray deviates from the Z-axis. However, since the Zernike sag (on the order of wavelengths) is much smaller than the surface aperture (on the order of greater than millimeters), the deviation of the reflected ray from the Z-axis can be neglected. In the following section, this approximation will still be used for further analysis.

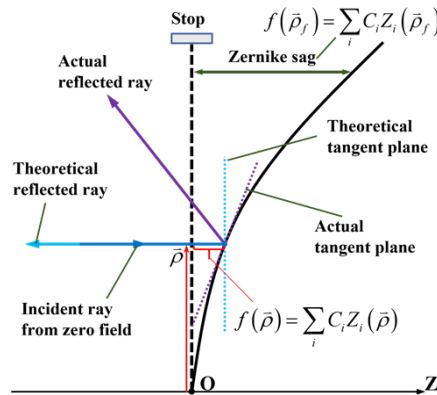


Fig. 5. The case when the zero-field incident ray is reflected at pupil position $\vec{\rho}$ of the freeform mirror.

3. Aberration contribution of Zernike freeform terms for general non-zero field of view

3.1. Freeform surface located at stop

In Fig. 5, the premise that reflected rays are transmitted along the Z-axis implies that the tangent plane at each pupil position of the freeform surface is perpendicular to the Z-axis. With this approximation, we analyzed the aberration contribution of at-stop Zernike freeform surface terms

for general non-zero FOV. As an example, a simple one-mirror system is shown in Fig. 6(a), the freeform surface is located at stop. For enhanced visualization, the freeform departure of the surface is presented, representing surface maps after removing the basic sphere [10]. The rays from non-zero field point are obliquely incident to the surface, with the incident angle of chief ray α_{in} . Figure 6(b) illustrates the behavior of a certain ray within the system as it irradiates at the pupil position $\vec{\rho}$ of the freeform mirror, both before and after the addition of Zernike terms. Under the approximation that the tangent plane of the surface is perpendicular to the Z-axis, the two reflected rays are parallel. For the sake of presentation, we have magnified the separation between the two reflected rays, BC , even though it is, in reality, negligibly small compared to the surface aperture, and similarly, $\Delta\vec{\rho} \ll \vec{\rho}$.

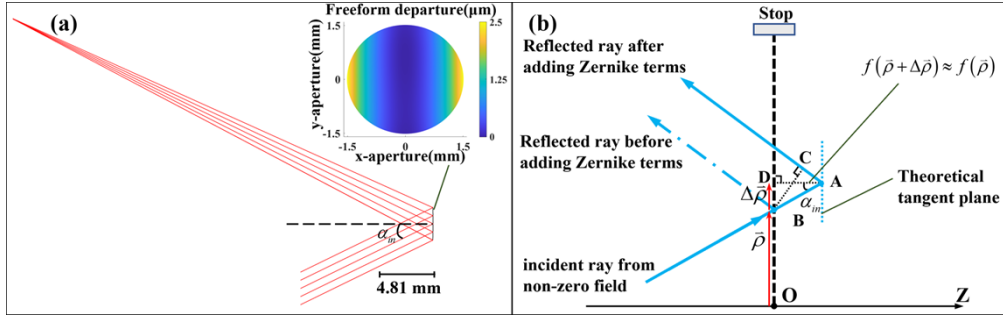


Fig. 6. (a) Schematic of one-mirror system where the freeform surface is located at stop and the rays from non-zero field point are obliquely incident to the surface, (b) The case when a certain ray within the one-mirror system irradiates at the pupil position $\vec{\rho}$ of the freeform mirror, under the approximation that the tangent plane of the surface is perpendicular to the Z-axis.

In this case, the optical path difference introduced by the Zernike sag $f(\vec{\rho})$ is

$$\begin{aligned}
 \Delta W &= BA + AC \\
 &= BA(1 + \cos(2\alpha_{in})) \\
 &= 2BA\cos^2(\alpha_{in}) \quad , \\
 &= 2AD \cos(\alpha_{in}) \\
 &= 2f(\vec{\rho} + \Delta\vec{\rho}) \cos(\alpha_{in})
 \end{aligned}
 \tag{12}$$

as $\Delta\vec{\rho} \ll \vec{\rho}$, we have,

$$\Delta W \approx 2f(\vec{\rho}) \cos(\alpha_{in}).
 \tag{13}$$

This implies that the generated non-zero-field aberrations result from the incidence-direction orthogonal component of the actual Zernike sag, and it can be written as $s_g \sum_i C_i Z_i(\vec{\rho}_p)$, where $s_g = \cos(\alpha_{in})$. Then according to Eq. (4), for general non-zeros field points, the aberration contribution of at-stop Zernike freeform terms can be calculated by,

$$W = s_g \frac{n' - n}{\lambda} \sum_i C_i Z_i(\vec{\rho}_p) = s_g \sum_i M_i Z_i(\vec{\rho}_p),
 \tag{14}$$

where $M_i = \frac{n' - n}{\lambda} C_i$.

Additionally, wave aberrations are typically fitted to a plane located at the exit pupil and perpendicular to the chief ray [26], we refer to this plane as the fitting exit pupil (FEP). In a

system with a round exit pupil, the aperture of the zero-field-point beam is circular at the exit pupil surface. Correspondingly, for non-zero-field points, the apertures of beams which are incident obliquely on the exit pupil plane, is approximately an ellipse at the FEP, as illustrated in Fig. 7.

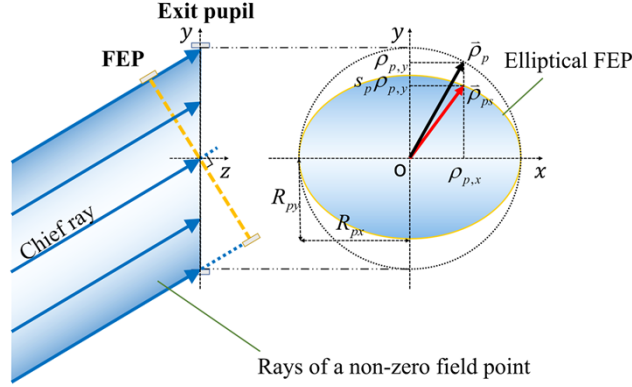


Fig. 7. The elliptical FEP of the non-zero field point

In this case, the normalized pupil vector $\vec{\rho}_p = (\rho_{p,x}, \rho_{p,y})$ should change to the effective scaled FEP vector $\vec{\rho}_{ps} = (\rho_{p,x}, s_p \rho_{p,y})$, where the s_p is given by,

$$s_p = \frac{R_{py}}{R_{px}}, \quad (15)$$

where R_{px} and R_{py} are the marginal ray heights in the x-direction and y-direction on FEP, respectively, and can be obtained via ray trace.

In a more general situation, according to Eq. (3), the relationship between the normalized aperture vector $\vec{\rho}_f$ and the effective scaled FEP vector $\vec{\rho}_{ps}$ is,

$$\begin{aligned} \vec{\rho}_f &= \vec{\rho}_p = (\rho_{p,x}, \rho_{p,y}) \\ &= (\rho_{ps,x}/s_{p,x}, \rho_{ps,y}/s_{p,y}) \end{aligned} \quad (16)$$

where

$$\begin{aligned} s_{p,x} &= \frac{R_{px}}{R_{py}}, s_{p,y} = 1, \text{ when } R_{px} \leq R_{py} \\ s_{p,x} &= 1, s_{p,y} = \frac{R_{py}}{R_{px}}, \text{ when } R_{px} > R_{py} \end{aligned}$$

Combine Eq. (14) and Eq. (16), for general non-zero field points, the aberration contribution of at-stop Zernike freeform terms can be accurately calculated by,

$$\begin{aligned} W &= s_g \sum_i M_i Z_i(\vec{\rho}_p) \\ &= s_g \sum_i M_i Z_i(\rho_{p,x}, \rho_{p,y}) \\ &= s_g \sum_i M_i Z_i(\rho_{ps,x}/s_{p,x}, \rho_{ps,y}/s_{p,y}) \end{aligned} \quad (17)$$

When the rays from zero field point pass the freeform surface located at stop, $s_g = 1$, $s_{p,x} = s_{p,y} = 1$, Eq. (17) transforms back to Eq. (4), so Eq. (17) is actually available for any field

point. For further analysis, we express the generated aberration W calculated by Eq. (17) as the general form of fringe Zernike polynomials. According to Eq. (17), the aberration generated by the i th Zernike term on the freeform surface is,

$$W_i = s_g M_i Z_i(\rho_{ps,x}/s_{p,x}, \rho_{ps,y}/s_{p,y}). \quad (18)$$

Based on the study [21], Eq. (18) can be expand into XY polynomials and then rearranged into fringe Zernike polynomials, they can be written as the form of vector multiplication, given by,

$$W_i = \mathbf{C}_i^T \mathbf{G} = \mathbf{M}_i^{rearrT} \mathbf{Z}^{adj} \quad (19)$$

where \mathbf{G} is the vector containing XY polynomial terms. \mathbf{C}_i is the corresponding coefficient vector. \mathbf{Z}^{adj} is the vector containing fringe Zernike polynomial terms which has been adjusted to remove constant terms and tilt terms without affecting the results. \mathbf{M}_i^{rearr} is the corresponding coefficient vector to be calculated. \mathbf{G} and \mathbf{Z}^{adj} are related by a transformation matrix \mathbf{T} , that is, $\mathbf{Z}^{adj} = \mathbf{T}\mathbf{G}$. The specific \mathbf{G} , \mathbf{Z}^{adj} and \mathbf{T} employed here can be found in study [21]. Therefore, according to Eq. (19), the vector of the fringe Zernike coefficient \mathbf{M}_i^{rearr} can be obtained by,

$$\mathbf{M}_i^{rearr} = (\mathbf{M}_i^{rearrT})^T = (\mathbf{C}_i^T \mathbf{G} (\mathbf{Z}^{adj})^{-1})^T = (\mathbf{C}_i^T (\mathbf{T})^{-1})^T = (\mathbf{T}^T)^{-1} \mathbf{C}_i, \quad (20)$$

based on Eq. (20), the aberration generated by the Zernike term on the freeform surface is allowed to be expressed using fringe Zernike polynomials.

Conventionally, the designed optical system is symmetric about yoz plane, and as a result, the Zernike terms considered here also maintain symmetry about the yoz plane. Figure 8 provides calculation formulas of fringe Zernike aberrations generated by Zernike terms, typically up to the 4th order, for the freeform surface at the aperture stop.

We validate the above calculation equations using the one-mirror system shown in Fig. 6(a), the wavelength is 500 nm, the test field points ranges from $(0^\circ, 0^\circ)$ to $(0^\circ, 55^\circ)$ in 5° intervals along the y-direction, we build the system in optical software Code V. Taking Zernike term Z9 as an example, we add Z9 with coefficient of 1λ to the freeform surface, the generated aberrations in each test field point are calculated by formulas derived from Eq. (17) (in Fig. 8), representing the proposed theoretical values. The comparison between the proposed theoretical value, traditional theoretical value (generated aberrations calculated by formulas derived from Eq. (4)) and the actual value (generated aberrations read from Code V) is shown in Fig. 9. Traditional theory suggests that a Zernike term on the freeform surface at stop generates a single type of aberration independent of the field. This theory holds accurate for small fields but progressively deviates from the actual values as the field increases. In contrast, the proposed theoretical value aligns more consistently with the actual values across a range of test field points. In addition, we obtained the maximum calculation error of the proposed theory for the total aberrations generated by each Zernike term in Fig. 8. The comparison of the calculation error between the proposed theoretical value and traditional theoretical value is illustrated in Fig. 10. The calculation error of traditional theoretical value increases as the field increases, whereas the error in the proposed theoretical value remains consistently small at all field points, with a maximum calculation error of 0.015λ .

3.2. Freeform surface located away from stop

In addition to the at-stop freeform surfaces, we further obtain the aberration contribution of non-stop freeform surface terms. Considering the angle of incidence and the shape of FEP, combine Eq. (11) and Eq. (17), the aberration generated by Zernike terms on freeform surface

Ci: Coefficient of the ith Zernike term added to freeform surface		Kj: Amount of the jth Zernike aberration generated by Ci	
C4	$K4 = M_4 \left(\frac{s_g}{2s_{p,x}^2} + \frac{s_g}{2s_{p,y}^2} \right)$	C5	$K5 = M_5 \left(\frac{s_g}{4s_{p,x}^2} - \frac{s_g}{4s_{p,y}^2} \right)$
C8	$K8 = M_8 \frac{s_g(3s_{p,x}^2 + s_{p,y}^2)}{4s_{p,x}^2 s_{p,y}^3}$	C11	$K11 = M_{11} \left(-\frac{s_g(s_{p,x}^2 - s_{p,y}^2)}{4s_{p,x}^2 s_{p,y}^3} \right)$
C9	$K9 = M_9 \frac{s_g(3s_{p,x}^4 + 2s_{p,x}^2 s_{p,y}^2 + 3s_{p,y}^4)}{8s_{p,x}^4 s_{p,y}^4}$	C12	$K12 = M_{12} \left(\frac{s_g}{2s_{p,x}^4} + \frac{s_g}{2s_{p,y}^4} \right)$
C17	$K17 = M_{17} \frac{s_g(s_{p,x}^2 - s_{p,y}^2)^2}{16s_{p,x}^4 s_{p,y}^4}$	C17	$K17 = M_{17} \left(\frac{s_g(s_{p,x}^2 - s_{p,y}^2)^2}{16s_{p,x}^4 s_{p,y}^4} \right)$

Fig. 8. Formulas of fringe Zernike aberrations generated by Zernike terms on the freeform surface located at stop

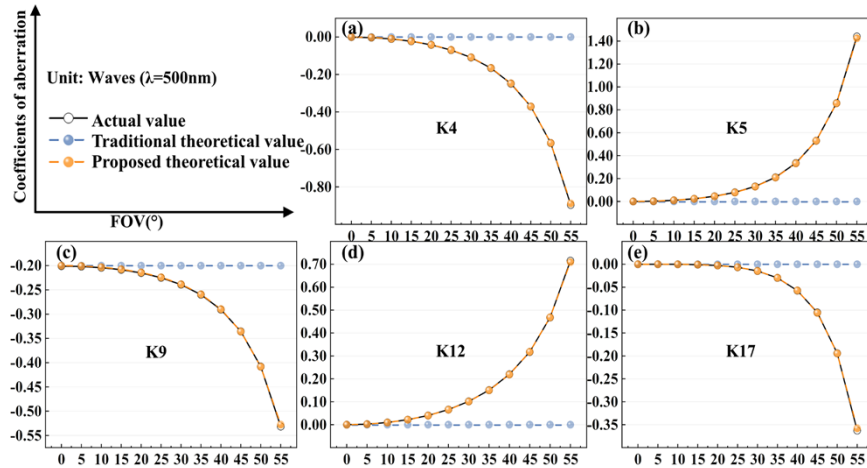


Fig. 9. Comparison between the proposed theoretical value, traditional theoretical value and the actual value when adding Z9 to the freeform surface located at stop, the generated aberrations are (a)K4, (b) K5, (c) K9, (d) K12 and (e) K17. The line between the data points is for the purpose of showing the trend of the generated aberrations.

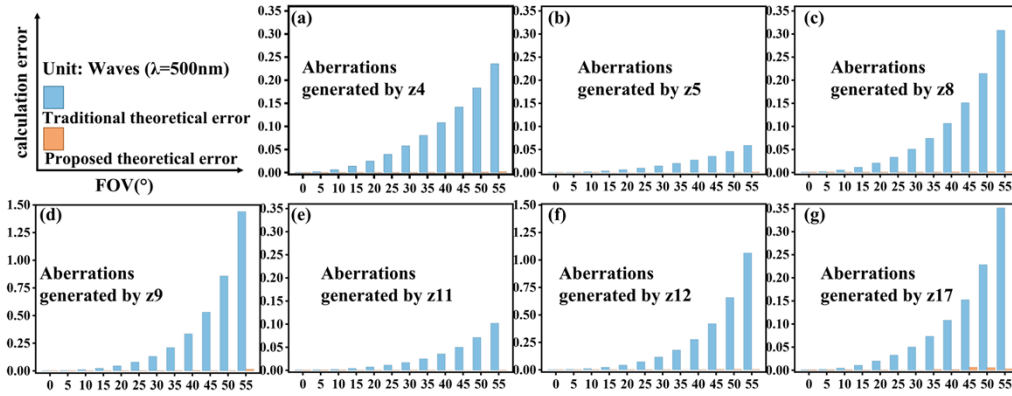


Fig. 10. Comparison of the calculation error of total aberrations between the proposed theoretical value and traditional theoretical value when adding (a) z4, (b) z5, (c) z8, (d) z9, (e) z11, (f) z12 and (g) z17 to the freeform surface located at stop

away from stop can be calculated by,

$$W = s_g \sum_i M_i Z_i(Q_x \rho_{ps,x} + h_x, Q_y \rho_{ps,y} + h_y), \quad (21)$$

where,

$$Q_x = \frac{as_{fx}}{s_{px}},$$

$$Q_y = \frac{as_{fy}}{s_{py}}.$$

In the same way, the coefficients of aberrations in forms of fringe Zernike polynomials derived from Eq. (21) are listed in Fig. 11.

We built an off-axis two-mirror system shown in Fig. 12 to validate the calculation formulas in Fig. 11. The entrance pupil is 10 mm. The primary mirror is a plane mirror tilted by 45° , the freeform surface is located at secondary mirror which is tilted by 45° and decentered by $5\sqrt{2}/2$ mm. The system has a $10^\circ \times 4^\circ$ FOV and an effective focal length of 189.2 mm. The wavelength is 500 nm. One set of test field points is symmetric about yoz plane, which ranges from $(0^\circ, 0^\circ)$ to $(5^\circ, 0^\circ)$ in 1° intervals along x direction, add each Zernike term in Fig. 11 with coefficient of 1λ to the freeform surface, the comparison of the calculation error of the total aberrations between the proposed theoretical value, traditional theoretical value (generated aberrations calculated by formulas derived from Eq. (11)) is illustrated in Fig. 13.

Another set of test field points is asymmetric and ranges from $(0^\circ, -2^\circ)$ to $(0^\circ, 2^\circ)$ in 1° intervals along y direction, in the same way, the comparison of the calculation error of the total aberrations between the proposed theoretical value, traditional theoretical value is illustrated in Fig. 14.

For both sets of test field points, the proposed theoretical value is more accurate than the traditional theoretical value, with a maximum error of 0.066λ . However, compared to the situation of at-stop freeform surface, the error of proposed theoretical value is larger, this is due to the use of the ellipse as approximation to describe the shape of beam footprint and the FEP, the approximation error is hard to avoid. Additionally, as mentioned before, the aberration formulas are approximate assuming that rays are transmitted along the Z direction or the chief-ray direction, avoiding the complex tracing for every single ray. Hence the calculation error introduced by the approximate formulas can be regarded as an inevitable tradeoff. Nevertheless, when compared to the conventional theoretical values (with an error range of 0.048λ to 0.302λ) at both sets of test

Ci: Coefficient of the ith Zernike term added to freeform surface		Kj: Amount of the jth Zernike aberration generated by Ci	
C4 $K4 = M_4 \frac{s_g(Q_x^2 + Q_y^2)}{2}$ $K5 = M_4 s_g(Q_x^2 - Q_y^2)$		C5 $K4 = M_5 \frac{s_g(Q_x^2 - Q_y^2)}{4}$ $K5 = M_5 \frac{s_g(Q_x^2 + Q_y^2)}{2}$	
C8 $K4 = M_8 \frac{3h_y s_g(Q_x^2 + 3Q_y^2)}{4}$ $K5 = M_8 \frac{3h_y s_g(Q_x^2 - 3Q_y^2)}{2}$ $K6 = M_8 3Q_x Q_y h_x s_g$ $K8 = M_8 \frac{Q_y s_g(Q_x^2 + 3Q_y^2)}{4}$ $K11 = M_8 \frac{3Q_y s_g(Q_x^2 - Q_y^2)}{4}$			
C9 $K4 = M_9 \frac{3s_g(3Q_x^4 + 2Q_x^2 Q_y^2 + 24Q_x^2 h_x^2 + 8Q_x^2 h_y^2 - 4Q_x^2 + 3Q_y^4 + 8Q_y^2 h_x^2 + 24Q_y^2 h_y^2 - 4Q_y^2)}{8}$ $K5 = M_9 \frac{-3s_g(-3Q_x^4 - 24Q_x^2 h_x^2 - 8Q_x^2 h_y^2 + 4Q_x^2 + 3Q_y^4 + 8Q_y^2 h_x^2 + 24Q_y^2 h_y^2 - 4Q_y^2)}{8}$ $K6 = M_9 24Q_x Q_y h_x h_y s_g$ $K7 = M_9 2Q_x h_x s_g(3Q_x^2 + Q_y^2)$ $K8 = M_9 2Q_y h_y s_g(Q_x^2 + 3Q_y^2)$ $K9 = M_9 \frac{s_g(3Q_x^4 + 2Q_x^2 Q_y^2 + 3Q_y^4)}{4}$ $K10 = M_9 6Q_x h_x s_g(Q_x^2 - Q_y^2)$ $K11 = M_9 6Q_y h_y s_g(Q_x^2 - Q_y^2)$ $K12 = M_9 \frac{3s_g(Q_x^4 - Q_y^4)}{4}$ $K17 = M_9 \frac{3s_g(Q_x^2 - Q_y^2)^2}{4}$			
C11 $K4 = M_{11} \frac{3h_y s_g(Q_x^2 - Q_y^2)}{4}$ $K5 = M_{11} \frac{3h_y s_g(Q_x^2 + Q_y^2)}{2}$ $K6 = M_{11} 3Q_x Q_y h_x s_g$ $K8 = M_{11} \frac{Q_y s_g(Q_x^2 - Q_y^2)}{4}$ $K11 = M_{11} \frac{Q_y s_g(3Q_x^2 + Q_y^2)}{4}$			
C12 $K4 = M_{12} \frac{-3s_g(-Q_x^4 - 8Q_x^2 h_x^2 + Q_x^2 + Q_y^4 + 8Q_y^2 h_y^2 - Q_y^2)}{2}$ $K7 = M_{12} 4Q_x^3 h_x s_g$ $K8 = M_{12} (-4Q_y^3 h_y s_g)$ $K5 = M_{12} \frac{3s_g(Q_x^4 + 8Q_x^2 h_x^2 - Q_x^2 + Q_y^4 + 8Q_y^2 h_y^2 - Q_y^2)}{2}$ $K9 = M_{12} \frac{s_g(Q_x^4 - Q_y^4)}{2}$ $K10 = M_{12} 4Q_x^3 h_x s_g$ $K11 = M_{12} 4Q_y^3 h_y s_g$ $K12 = M_{12} \frac{s_g(Q_x^4 + Q_y^4)}{2}$ $K17 = M_{12} \frac{s_g(Q_x^4 - Q_y^4)}{2}$			
C17 $K4 = M_{17} \frac{3s_g(Q_x^2 - Q_y^2)(Q_x^2 - Q_y^2 + 8h_x^2 - 8h_y^2)}{16}$ $K6 = M_{17} (-12Q_x Q_y h_x h_y s_g)$ $K5 = M_{17} \frac{3s_g(Q_x^2 + Q_y^2)(Q_x^2 - Q_y^2 + 8h_x^2 - 8h_y^2)}{16}$ $K7 = M_{17} Q_x h_x s_g(Q_x^2 - Q_y^2)$ $K8 = M_{17} (-Q_y h_y s_g(Q_x^2 - Q_y^2))$ $K9 = M_{17} \frac{s_g(Q_x^2 - Q_y^2)^2}{16}$ $K10 = M_{17} Q_x h_x s_g(Q_x^2 + 3Q_y^2)$ $K11 = M_{17} (-Q_y h_y s_g(3Q_x^2 + Q_y^2))$ $K12 = M_{17} \frac{s_g(Q_x^4 - Q_y^4)}{8}$ $K17 = M_{17} \frac{s_g(Q_x^4 + 6Q_x^2 Q_y^2 + Q_y^4)}{8}$			

Fig. 11. Formulas of aberrations generated by Zernike terms on the off-axis freeform surface

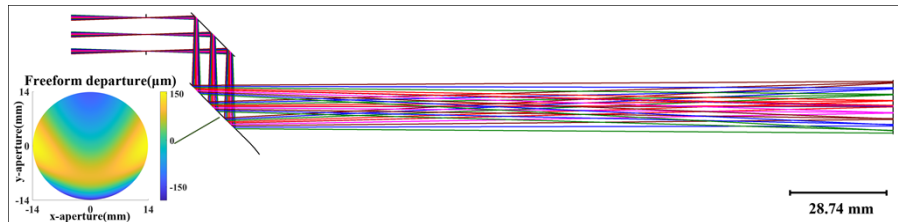


Fig. 12. Schematic of the off-axis two-mirror system

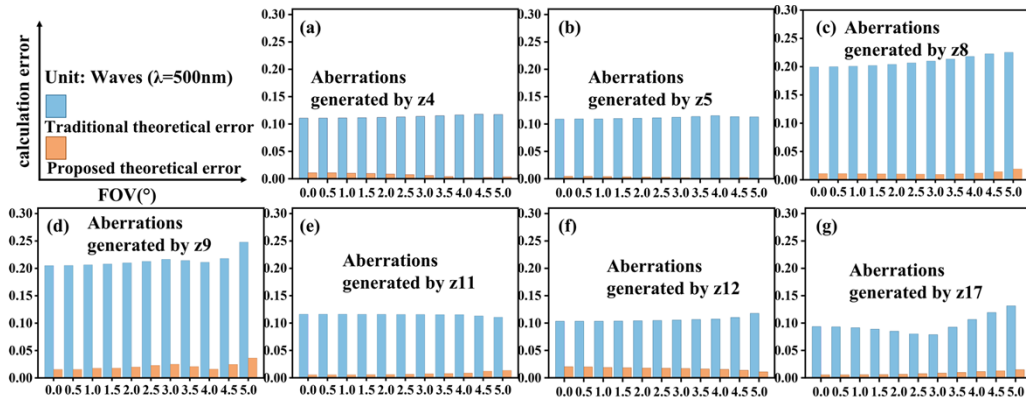


Fig. 13. Comparison of the calculation error of total aberrations between the proposed theoretical value and traditional theoretical value when adding (a) z4, (b) z5, (c) z8, (d) z9, (e) z11, (f) z12 and (g) z17 to the off-axis freeform surface, the test field points are along the x direction that is symmetric about yoz plane

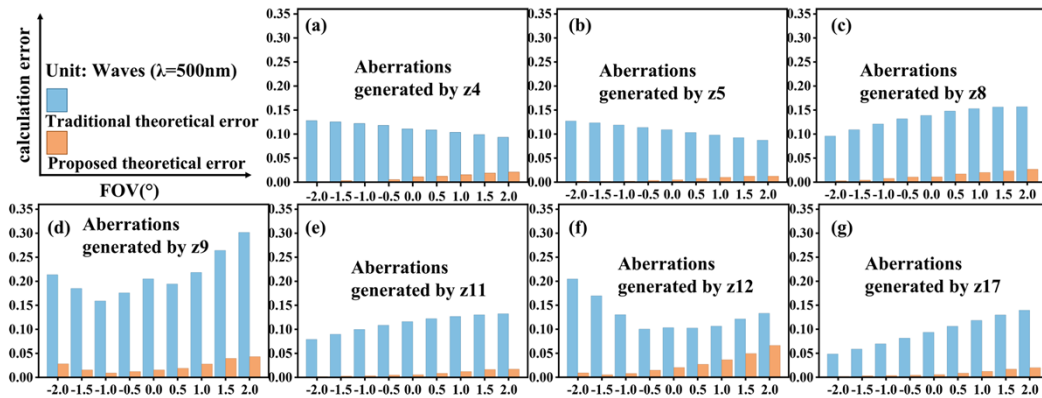


Fig. 14. Comparison of the calculation error of total aberrations between the proposed theoretical value and traditional theoretical value when adding (a) z4, (b) z5, (c) z8, (d) z9, (e) z11, (f) z12 and (g) z17 to the off-axis freeform surface, the test field points are along the y direction that is not symmetric about yoz plane

field points, the proposed theoretical values exhibit a much narrower error range of 0.001λ to 0.066λ . The proposed formulas reduce the calculation error by more than 78% and significantly improve computational accuracy. Based on the proposed formulas, the Zernike terms on each surface of the system have the individual aberration contribution corresponding to their respective surface, and they can all be accurately calculated without being confused.

4. NAT-guided optimization method for off-axis freeform system

In this section, we present an application scenario of the proposed high-precision aberration analysis. The optimization strategy based on NAT utilizes the predictable aberration field generated by freeform surface terms to compensate the aberration field of the system, generating a reasonable optimization starting point for ray-tracing-based optimization. This strategy offers flexibility in identifying the freeform terms that require optimization based on the system's aberration characteristics. It aids in determining the optimization direction and mitigates the inefficiency caused by relying solely on ray-tracing-based optimization algorithms [10]. However,

due to the insufficient accuracy of the traditional theory in calculating the aberration contribution for general non-zero FOV, the accurate quantitative optimization guidance based on NAT is only suitable for coaxial systems with paraxial FOV [27].

Here, with the proposed high-precision aberration calculation formulas, we present a NAT-guided optimization method for off-axis freeform surface system with general non-zero FOV. This method supports the optimization for systems containing multiple freeform surfaces, reducing the complexity of the optimization process. The total optimization flowchart is as Fig. 15, two steps are included, Step (1) is the optimization for initial system. In this step, we first construct the NAT-guided error function $W_{sys}(C_{s,i})$, where the coefficient of the Zernike freeform surface term is designated as the optimization variable. Then, we employ the gradient descent algorithm to optimize the error function, extract the optimal coefficients of Zernike terms and complete the initial construction of the freeform surfaces. The error function $W_{sys}(C_{s,i})$ is denoted as,

$$W_{field}(C_{s,i}) = \sum_j \left(W_{0(j)} + \sum_s W_{z(s,j)}(C_{s,i}) \right)^2, \quad (22)$$

$$W_{sys}(C_{s,i}) = \sum_{field} W_{field}(C_{s,i})$$

where $W_{field}(C_{s,i})$ is the comprehensive wave aberration of a specific field point, $C_{s,i}$ is the i th Zernike coefficient of the s th freeform surface. For this field point, $W_{0(j)}$ is the amount of the j th fringe Zernike aberration. $W_{z(s,j)}$ is the amount of the j th fringe Zernike aberration generated by $C_{s,i}$, it is calculated by the proposed formulas in Fig. 8 and Fig. 11. The NAT-guided error function $W_{sys}(C_{s,i})$ is the sum of the $W_{field}(C_{s,i})$ of all selected field points. This error function denotes the superposition of the aberrations of the initial system and the aberration contribution of the freeform Zernike terms, reflecting the comprehensive level of aberrations when Zernike terms are added to optical surfaces. The optimized system after Step (1) is taken as the optimization starting point of Step (2). Step (2) involves further optimizing the surface shape parameters, thickness, decenter and tilt. This optimization is performed using the damped least squares method integrated into the optical software, employing default error functions. Iterate the Step (1) and (2) if the system performance doesn't meet the target.

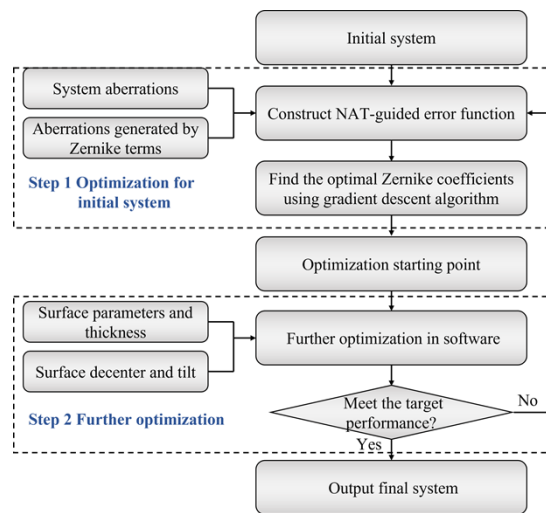


Fig. 15. Flowchart of the NAT-guided optimization method

Using the proposed method, we optimized an off-axis three-mirror system, the specifications and target performance are shown in Table 1. The initial system is shown in Fig. 16(a) with

the primary mirror located at stop. All three surfaces are spherical, with radius of 1521.3 mm for PM, 468.1 mm for SM, and 667.3 mm for TM. In Step (1), we optimized the NAT-guided error function in MATLAB, the variables are the Zernike coefficients of three surfaces. In each iteration of the gradient descent process, the error function continues to decrease, obtaining the efficient optimized Zernike terms. After about 20 iterations, the decrease is not obvious, we stop iteration and extract the Zernike terms as the optimal solution. The obtained system with constructed freeform surfaces is shown in Fig. 16(b), the RMS wavefront error (RMS WFE) is reduced from 68.639λ of initial system to 5.679λ . Then we set the Zernike coefficients, conics, tilts of the three surfaces and the image plane distance as the variables to take a further optimization in Code V, after 6 iterations of damped least squares, the target performance was met. The final system with three optimized freeform surfaces is shown in Fig. 16(c).

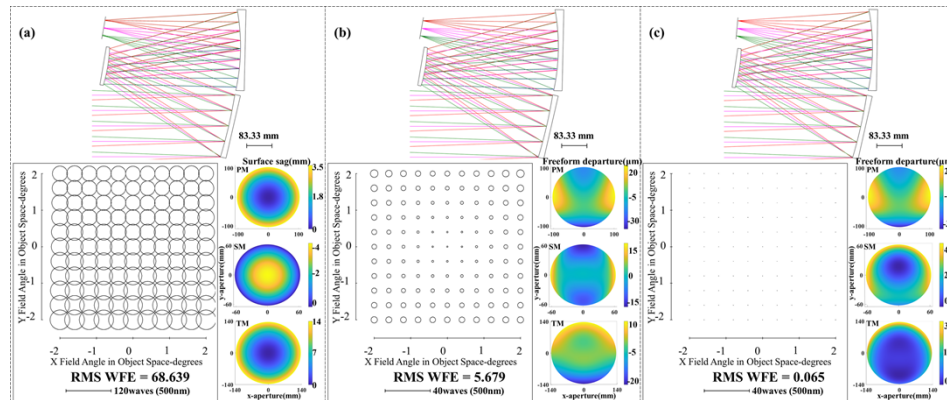


Fig. 16. Schematic of the three-mirror systems and their RMS WFE. (a) initial system, (b) system after the NAT-guided surface construction(Step (1)), and (c) the final system optimized by the proposed method

Table 1. Specifications and target performance of the three-mirror system

Entrance pupil(mm)	200	Wavelength(nm)	500
Focal length(mm)	600	RMS wavefront error(λ)	≤ 0.07
Full field of view($^{\circ}$)	4×4	Distortion	$\leq 1\%$

As a further proof of the proposed optimization method, we replaced the NAT-guided operation for solving the Zernike coefficients in Step (1) with a direct software optimization, and obtained a system optimized without NAT guidance. Then perform Step (2) on this system under the same constraints, and with the same variables as the NAT-guided system. The optimization process terminates when it either reaches the target specification or converges. The comparison between the final system optimized with and without NAT guidance in terms of the RMS WFE and distortion are shown in Table 2. It suggested that the system optimized with NAT guidance achieved a better optical performance. We attribute this to the fact that the proposed method incorporates NAT to account for the field-dependent property of aberrations, thereby compensating for the entire aberration field of the initial system, without relying exclusively on ray-tracing-based algorithm. Therefore, Step (1) with NAT-guidance generates a more reasonable optimization starting point which enables Step (2) to output the ideal final system.

Table 2. RMS WFE and distortion of the system optimized with and without NAT guidance

	RMS WFE(λ)	distortion
System optimized with NAT guidance	0.065	0.89%
System optimized without NAT guidance	0.197	2.11%

5. Conclusion and discussion

In this paper, a high-precise analysis of aberration contribution of Zernike freeform surface terms for non-zero field points is provided. As a theoretical basis, we first reviewed the existing analyses of the aberrations generated by Zernike freeform surface terms. Subsequently, we conduct a comprehensive analysis of the aberration contribution of Zernike terms on both at-stop and non-stop surfaces for non-zero FOV. The chief ray's angle of incidence and the shape of the fitting exit pupil are considered to derive the precise aberration calculation formulas. Design examples have demonstrated the enhanced calculation accuracy of the proposed formulas, resulting in a remarkable 78% reduction in calculation error compared to traditional theoretical values. Finally, we proposed a NAT-guided optimization method for off-axis freeform surface systems based on the high-precision analysis, and validated its effectiveness through the optimization of a three-mirror system. This research contributes to the broader application of NAT in aberration analysis and offers valuable insights for the optimal design and alignment of optical freeform systems.

This paper explores the aberration contribution of Zernike freeform surfaces. However, for some surface shapes exhibiting special characteristics, Zernike polynomials may have limitations in fitting accuracy. In such cases, alternative types of freeform basis functions are necessary for accurate representation [12]. For example, a surface shape with the information concentrated around the center of the analysis domain is better suited for representation using vortex Jinc functions [28]. For surfaces with abrupt variations, such as the anterior eye surface at the limbus region, Bessel functions are more capable of representation [29]. And Chebyshev polynomials are suitable to represent the complete ophthalmic surfaces [30]. Clarifying the aberration contribution of these alternative basis functions aids in understanding the aberration behavior of optical systems containing surfaces with special characteristics, providing theoretical guidance for improving system performance. Therefore, we hope that our study can offer insights for conducting these potential researches in the future.

Funding. National Natural Science Foundation of China (62075106); Nankai University Eye Institute (NKYKD202204, NKYKK202209).

Disclosures. The authors declare no conflicts of interest.

Data availability. Data underlying the results presented in this paper are not publicly available at this time but may be obtained from the authors upon reasonable request.

References

1. J. P. Rolland, M. A. Davies, T. J. Suleski, *et al.*, "Freeform optics for imaging," *Optica* **8**(2), 161 (2021).
2. Z. Qin, Q. Meng, and X. Wang, "Desensitization design method of a freeform optical system based on local curve control," *Opt. Lett.* **48**(1), 179–182 (2023).
3. L. Zhu, Z. Li, F. Fang, *et al.*, "Review on fast tool servo machining of optical freeform surfaces," *Int J Adv Manuf Technol* **95**(5-8), 2071–2092 (2018).
4. J. Reimers, A. Bauer, K. P. Thompson, *et al.*, "Freeform spectrometer enabling increased compactness," *Light: Sci. Appl.* **6**(7), e17026 (2017).
5. D. Cheng, H. Chen, C. Yao, *et al.*, "Design, stray light analysis, and fabrication of a compact head-mounted display using freeform prisms," *Opt. Express* **30**(20), 36931–36948 (2022).
6. H. Huang and H. Hua, "High-performance integral-imaging-based light field augmented reality display using freeform optics," *Opt. Express* **26**(13), 17578–17590 (2018).

7. A. Wilson and H. Hua, "Design and demonstration of a vari-focal optical see-through head-mounted display using freeform Alvarez lenses," *Opt. Express* **27**(11), 15627–15637 (2019).
8. W. Wu, H. Wang, G. Jin, *et al.*, "Fast automatic design method for freeform imaging systems through system construction and correction," *Opt. Lett.* **45**(18), 5140–5143 (2020).
9. F. Duerr and H. Thienpont, "Freeform imaging systems: Fermat's principle unlocks "first time right" design," *Light: Sci. Appl.* **10**(1), 95 (2021).
10. A. Bauer, E. M. Schiesser, and J. P. Rolland, "Starting geometry creation and design method for freeform optics," *Nat. Commun.* **9**(1), 1756 (2018).
11. Z. Zhuang, Y. Chen, F. Yu, *et al.*, "Field curvature correction method for ultrashort throw ratio projection optics design using an odd polynomial mirror surface," *Appl. Opt.* **53**(22), E69–76 (2014).
12. Z. Gao, Q. Yuan, X. Li, *et al.*, "Review of optical freeform surface representation technique and its application," *Opt. Eng.* **56**(11), 1 (2017).
13. R.V. Shack and K. Thompson, "Influences of alignment errors of a telescope system on its aberation field," *Proc. SPIE* **251**, 146–153 (1980).
14. R.A. Buchroeder, "Tilted component optical systems," Ph.D dissertation (University of Arizona, 1976).
15. K. Thompson, "Description of the third-order optical aberrations of near-circular pupil optical systems without symmetry," *J. Opt. Soc. Am. A* **22**(7), 1389–1401 (2005).
16. K. P. Thompson, "Multinodal fifth-order optical aberrations of optical systems without rotational symmetry: spherical aberration," *J. Opt. Soc. Am. A* **26**(5), 1090–1100 (2009).
17. K. P. Thompson, "Multinodal fifth-order optical aberrations of optical systems without rotational symmetry: the comatic aberrations," *J. Opt. Soc. Am. A* **27**(6), 1490–1504 (2010).
18. K. P. Thompson, "Multinodal fifth-order optical aberrations of optical systems without rotational symmetry: the astigmatic aberrations," *J. Opt. Soc. Am. A* **28**(5), 821–836 (2011).
19. K. Fuerschbach, J. P. Rolland, and K. P. Thompson, "Theory of aberration fields for general optical systems with freeform surfaces," *Opt. Express* **22**(22), 26585–26606 (2014).
20. T. Yang, J. Zhu, and G. Jin, "Nodal aberration properties of coaxial imaging systems using Zernike polynomial surfaces," *J. Opt. Soc. Am. A* **32**(5), 822–836 (2015).
21. T. Yang, D. Cheng, and Y. Wang, "Aberration analysis for freeform surface terms overlay on general decentered and tilted optical surfaces," *Opt. Express* **26**(6), 7751–7770 (2018).
22. G. Ju, C. Yan, Z. Gu, *et al.*, "Computation of astigmatic and trefoil figure errors and misalignments for two-mirror telescopes using nodal-aberration theory," *Appl. Opt.* **55**(13), 3373–3386 (2016).
23. K. Fuerschbach, J. P. Rolland, and K. P. Thompson, "Extending nodal aberration theory to include mount-induced aberrations with application to freeform surfaces," *Opt. Express* **20**(18), 20139–20155 (2012).
24. X. Zhang, D. Zhang, S. Xu, *et al.*, "Active optical alignment of off-axis telescopes based on nodal aberration theory," *Opt. Express* **24**(23), 26392–26413 (2016).
25. J. Wang, X. He, J. Luo, *et al.*, "Alignment algorithm of nonsymmetric off-axis reflective astronomical telescopes based on the modified third-order nodal aberration theory," *Opt. Express* **30**(8), 13159–13183 (2022).
26. Code V Reference Manual (Synopsys Inc., 2012)
27. S. Zhang, L. Xiao, X. Zhao, *et al.*, "Optimization method using nodal aberration theory for coaxial imaging systems with radial basis functions based on surface slope," *Appl. Opt.* **60**(1), 1 (2021).
28. J. E. Gomez-Correa, M. A. Platas-Garza, J. P. Trevino, *et al.*, "Topographic synthesis of arbitrary surfaces with vortex Jinc functions," *Appl. Opt.* **59**(13), D95–D103 (2020).
29. J. P. Trevino, J. E. Gomez-Correa, D. R. Iskander, *et al.*, "Zernike vs. Bessel circular functions in visual optics," *Ophthalmic Physiol Opt* **33**(4), 394–402 (2013).
30. M. M. Rodrigues, A. Rosa, N. Vieira, *et al.*, "Modeling ophthalmic surfaces using Zernike, Bessel and Chebyshev type functions," *J. Phys.: Conf. Ser.* **1194**, 012093 (2019).