Generalized sine condition for image-forming systems with centering errors

S. A. Comastri and J. M. Simon

Laboratorio de Optica, Facultad de Ciencias Exactas y Naturales Universidad de Buenos Aires,
1428 Buenos Aires, Argentina

R. Blendowske

FH Aalen, Fachbereich Augenoptik, Gartenstrasse 135, D-73430 Aalen, Germany

The generalized sine condition for an image-forming system with centering errors but allowing for one symmetry plane is derived according to the Fourier optics approach. The variation of the wave-front-aberration function associated with a small displacement of field coordinates is given. The symmetry properties of aberrations are discussed. © 1999 Optical Society of America [S0740-3232(99)01803-7]

OCIS codes: 220.0220, 220.1010, 070.2580.

1. INTRODUCTION

As a map, scaled 1:1, is of little use to hikers, not all the aberration data that are accessible by the numerical analysis of optical systems can guide the process of optical design. On the contrary, this process depends heavily on data reduction and visualizing schemes. For centered image-forming systems, various performance plots are frequently used, such as longitudinal and transverse aberrations and optical path differences.[1,2]

An additional tool is the offense against the sine condition[1] (OSC). For a centered optical system (COS), it allows for off-axis information by the evaluation of on-axis data only. The generalized OSC was derived by Hopkins[3] for a general noncentered optical system (NCOS) by means of geometrical optics. We pick up this topic in the language of Fourier optics as developed in Refs. 4–6. Recently Mansuripur[7] reviewed the sine condition (SC). The link between considering the image of a diffraction grating and the Fourier optics approach was discussed in Ref. 8.

In 1873 Abbé applied the SC to microscope objectives for the first time. Because these microscope lenses are characterized by a small field of view and a large aperture, they are still the typical examples for the use of the OSC. Nevertheless, times changed, and microscope objectives and their applications did as well.

Some problems encountered today in the fabrication of microscope objectives provide a practical motivation for reconsidering the OSC. For metrology applications in the semiconductor industry, such as optical control of microchips, only high-end objectives are mandatory. They are nominally designed for a Strehl value greater than 99% for the field point on the optical axis. Additional aberrations, i.e., coma and astigmatism, are introduced at this point in production by decentrations, tilts, and surface errors. Nevertheless, a Strehl ratio greater than 97% for the produced objectives is sometimes required.

This value has been found to be reasonable for referring to these objectives as *production limited*.[9] The practical manner of achieving such a specification during assembly is discussed in Refs. 9 and 10. Moreover, for the so-called box-in-box overlay measurements (with a repeatability of some 5 nm) the performance of objectives, even if they are production limited, may depend crucially (besides other factors) on the variation of the remaining aberrations across the small imaged patch (typically some 60 μm in diameter). Asymmetric aberrations, i.e., coma, lead to systematic errors in the measured results. To study this problem the OSC (which is a property of the patch) is a higher-quality parameter than the Strehl ratio (evaluated for only one object point). This example features the typical conditions—high numerical aperture (up to 0.95 times the refractive index), small field, and NCOS—for applying the OSC and understanding the symmetry properties of the related aberrations.

In the present paper our approach is tailored, for simplicity, to a perturbated symmetrical optical system (PSOS). A PSOS (which is a special case of NCOS) is produced when an image-forming system is designed to be a COS but undergoes small manufacturing perturbations (decentrations and tilts) that transform it into a decentered system with one symmetry plane (this is explained in detail in Subsection 2.A). Throughout, we compare the results obtained for a PSOS with those expected for the reference COS. The surfaces of the object and the receiving screen of the PSOS are planes parallel to each other coinciding with those of the corresponding COS. To simplify the analysis, we confine our attention to object points in the neighborhood of the symmetry plane, although a generalization to other points is possible. To make the calculations easier we consider systems that are not telecentric. Moreover, we exclude from our analysis infinite conjugates, because these would obscure our line of reasoning. (If the PSOS is a microscope

objective designed for the image plane at infinity, we include a tube lens that renders a finite image distance.)

Our approach is applicable to any image-forming system where the control of aberrations is critical (such as high-definition photography objectives) provided that object points in the neighborhood of the symmetry plane are considered. Nevertheless, it is especially useful for microscope objectives since, on the one hand, these often verify the conditions mentioned in the preceeding paragraph and, on the other, they require extremely strict control of residual aberrations because of their high numerical apertures.

We obtain the generalized sine condition by means of Fourier optics and relate its offense to the field derivatives of the wave-front-aberration function. In addition, we give the complete set of equations that can be used to study decentration tolerances for a PSOS by means of ray tracing. We also discuss the symmetry properties of aberrations. Tracing rays that are emanating from a fixed reference object point and applying the derived equations, the designer can gain insights into the aberrations that correspond to neighboring object points.

2. NOTATION

A. Perturbed Symmetrical Optical System
We consider as a reference optical system an undisturbed objective as it is usually designed, namely, as a COS. The object and image planes, S and S', respectively, are perpendicular to the optical axis z. The on-axis field point (the central object point) is termed $\tilde{O}_o$. The entrance and exit pupils[11] are centered at axial points $\tilde{E}_o$ and $\tilde{E}'_o$, respectively (see Fig. 1).

After being manufactured, the objective presents decentration and tilts of single lenses or lens groups and surface errors. It is rendered as a NCOS in the general sense. A first step in the approach to describe this system is the assumption of small perturbations. This allows for a linearized treatment of the individual contributions to the resulting overall aberrations. Each combination of decenterings can be compensated for by one shifting element, each combination of tilts can be leveled out by a tilt of the image plane, and each combination of surface errors can be neutralized by the rotation of a lens that has surface errors.
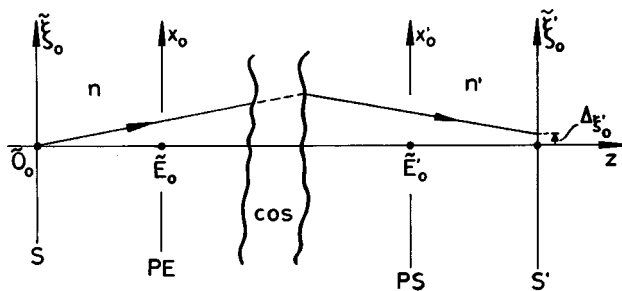

Fig. 1. Coordinate system for the reference COS: S and S', object and image plane surfaces; PE and PS, entrance and exit pupil planes; z, optical axis; $(\tilde{\xi}_o, \tilde{\eta}_o)$ $[(\tilde{\xi}'_o, \tilde{\eta}'_o)]$, coordinates at the object (image) with origin at $\tilde{O}_o$ (and at its image); (x_o, y_o) $[(x'_o, y'_o)]$, coordinates at the entrance (exit) pupil with origin at $\tilde{E}_o$ $(\tilde{E}'_o)$.

Through the work of Simon and Comastri,[6] the conceptual starting point for a general treatment, the optical invariant, has been derived for a general NCOS. Nevertheless, here we consider the constraint to an optical system with small decentrations and tilts (no surface error) such that there is a symmetry plane, that is, the constraint to a PSOS. The reason behind this restriction is to simplify the calculation and consider the case of highest practical interest, i.e., coma. As an elementary example, one may consider a PSOS with one decentration only, which is compensated for by the appropriate element. This configuration clearly features a symmetry plane. A better understanding of the OSC even for this restricted case gives us more insight into the behavior of aberrations. A simple approximation to the more general case of some decentrations, not necessarily oriented in one symmetry plane, is to combine all effects into a single aggregated one, and together with the shifting element they again yield a symmetry plane. Then we could apply our results to the more general case as well, as long as the approximation of the optical system to a PSOS is valid (this holds if Hopkins's basic theorem[3] is applicable).

B. Coordinate Systems
The surfaces of the object and receiving screen are planes parallel to each other that coincide with the surfaces S and S', respectively, of the reference COS. These surfaces are maintained throughout the paper even when tilts are present; that is, the compensation of these tilts by a tilt of the image plane is not contemplated. We consider arbitrary aperture reference planes at the entrance and the exit Π and Π', respectively (which may coincide but are not necessarily the entrance and the exit pupil planes of the reference COS).

To describe the PSOS there are three types of coordinate systems[3]: base, local, and canonical (see Appendix A). The coordinates used in the calculations are the canonical ones, so in what follows, only these are considered (see Fig. 2). We take into account a reference object point Q_ξ in the symmetry plane (the subscript ξ is used to stress this fact). The base ray (central ray of the beam) $Q_\xi E$ intersects the plane Π at point E, the plane Π' at point E', and the plane S' at point Q'_ξ (considered as the geometrical image of Q_ξ[3]).

The canonical coordinates in the object are (ξ_T, η_S, ζ) with origin in point O, defined as the foot of the perpendicular from E to the object plane S. These coordinates are chosen as follows: ξ_T is the intersection of the symmetry plane with the plane S, ζ is along OE, and η_S is perpendicular to the symmetry plane. Similarly, in the image we have coordinates $(\xi'_T, \eta'_S, \zeta')$ (all primed variables belong to image space).

The canonical coordinates in the entrance and exit reference spheres are (X_T, Y_S, Z) and (X'_T, Y'_S, Z') with origin at E and E', respectively and $Z(Z')$ on the prolongation of $\zeta(\zeta')$. The systems (X_T, Y_S, Z) and (X'_T, Y'_S, Z') are parallel to (ξ_T, η_S, ζ) and $(\xi'_T, \eta'_S, \zeta')$, respectively. With use of vector notation and two components only, the canonical coordinates at the entrance and exit reference spheres of a general ray from Q_ξ are
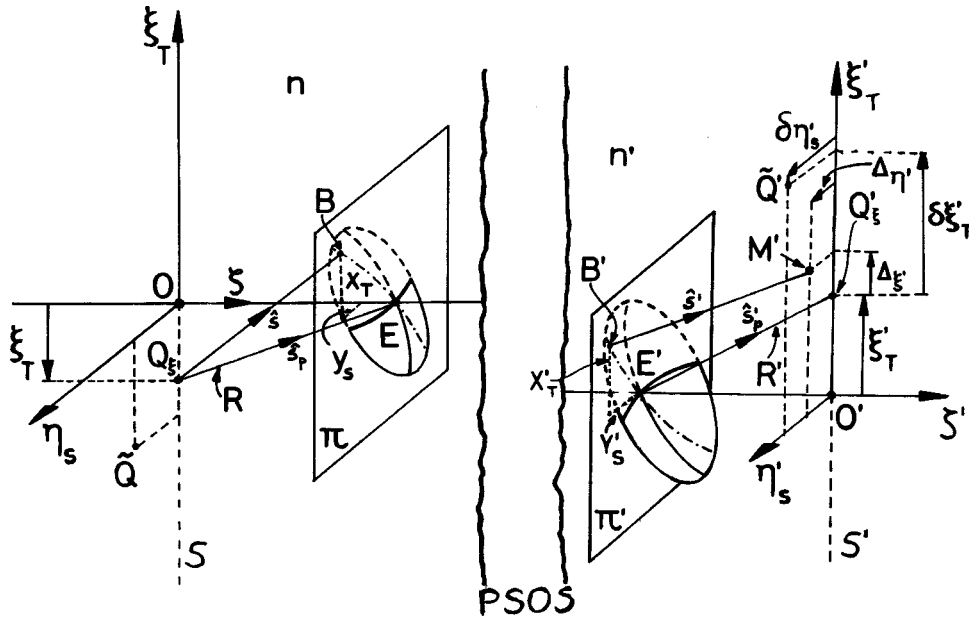
Fig. 2.   Canonical coordinate systems for the object point Q_ξ of the PSOS: S and S', object and image plane surfaces; Π and Π', entrance and exit aperture planes; Q_ξ, object point; $Q_\xi E$, base ray; O (O'), foot of the perpendicular from E (E') to S (S'); B and B', points where the general ray from Q_ξ intersects the reference spheres; (X_T, Y_S) $[(X_T', Y_S')]$, canonical axes in the entrance (and exit) reference spheres; (ξ_T, η_S) $[(\xi_T', \eta_S')]$, canonical coordinates in the object (image); $\tilde{Q}$, object point in the neighborhood of Q_ξ with image at $\tilde{Q}'$; $(\delta\xi_T', \delta\eta_S')$, relative canonical coordinates in the image for $\tilde{Q}'$; $\Delta_{\xi'}$, $\Delta_{\eta'}$, transverse aberration for the general ray from Q_ξ.

$$\chi = (X_T, Y_S), \qquad \chi' = (X_T', Y_S'), \qquad (1)$$

where boldface indicates that the mathematical character is a vector.

The radius of the local entrance reference sphere is termed $R = EQ_\xi$ and has curvature center at Q_ξ.   Correspondingly, $R' = E'Q_\xi'$ is the radius of the local exit reference sphere (in Fig. 2, R is negative and R' is positive).   A general ray from Q_ξ intersects the entrance and exit reference spheres at points B and B', respectively, and the image plane S' at point M', since there might be transverse aberrations $\Delta = (\Delta_{\xi'}, \Delta_{\eta'})$ (measured from Q_ξ').

Because we are interested in variations of the optical path difference related to small shifts of the field point, we consider a general third object point, $\tilde{Q}$ (not necessarily in the symmetry plane) in the neighborhood of Q_ξ. For this point, $\tilde{Q}E...E'\tilde{Q}'$ is the image-space associated base ray, defined as that which intersects the same point E' of the exit reference sphere as does the base ray from Q_ξ.[3]   Again the intersection of the base ray with S', point $\tilde{Q}'$, is considered the geometrical image point of $\tilde{Q}$.[3] For $\tilde{Q}$ we have the same pupil locations as for Q_ξ but accordingly adjusted radii $\tilde{R} = \tilde{Q}E$ and $\tilde{R}' = E'\tilde{Q}'$.   The canonical coordinates at the object (and image) plane for $\tilde{Q}$ are $(\tilde{\xi}_T, \tilde{\eta}_S)$ $[(\tilde{\xi}_T', \tilde{\eta}_S')]$ with origin at point O (O'). We define relative canonical coordinates in the object $(\tilde{\xi}_T - \xi_T, \tilde{\eta}_S - \eta_S)$ and in the image $(\tilde{\xi}_T' - \xi_T', \tilde{\eta}_S' - \eta_S')$[3] ($\eta_S = \eta_S' = 0$ for Q_ξ).   With vector notation, the relative canonical coordinates in the object and the image are the row vectors

$$\delta\Psi = (\delta\xi_T, \delta\eta_S) = (\tilde{\xi}_T - \xi_T, \tilde{\eta}_S - \eta_S),$$

$$\delta\Psi' = (\delta\xi_T', \delta\eta_S') = (\tilde{\xi}_T' - \xi_T', \tilde{\eta}_S' - \eta_S'). \qquad (2)$$

The subscripts T and S for the canonical coordinates are used following Hopkins,[3] but for a PSOS these coordinates coincide with the local ones (see Appendix A), so to avoid tedious notation in what follows, we suppress them.

3.   SINE CONDITION BY USE OF FOURIER OPTICS AND ITS OFFENSE

A.   Fourier Optics Approach

For a patch on the object plane, the complex field can be Fourier decomposed, and each component can be identified as a plane wave.[12]   The shape of each wave front is distorted as it propagates through the system, and its evaluation is as difficult as the computation of aberrations.   Nevertheless, we can link a ray to each wave front and expand the wave-front-aberration function around this ray.   The zeroth- and first-order terms are determined by the optical path and inclination of the ray and correspond to a plane wave front.   The superposition of all the plane wave fronts emanating from the object and captured by the optical system is an adequate description for the field in a small patch if the origins of the expansions corresponding to every wave front coincide.   This holds precisely at the image of the reference object point where, up to first order, all the rays intersect.   Then in both the object and the image we consider small patches surrounding the reference point and associate with the Fourier expansion a superposition of plane waves.   Each

plane wave at the image corresponds to a plane wave at the object. In what follows we give the corresponding formulas.

The complex field at the object, $U(\delta\xi, \delta\eta)$, is

$$U(\delta\xi, \delta\eta) = \int_{-\infty}^{\infty} \hat{U}(\nu_x, \nu_y)$$
$$\times \exp\{-i[2\pi(\nu_x\delta\xi + \nu_y\delta\eta)]\}d\nu_x d\nu_y, \quad (3)$$

with generalized spatial frequencies $\boldsymbol{\nu} = (\nu_x, \nu_y)$ and amplitude or angular spectrum $\hat{U}(\nu_x, \nu_y)$.[4,12,13] We specify each ray by a normalized or unit ray vector (s_x, s_y, s_z) (with $s_x^2 + s_y^2 + s_z^2 = 1$ so that each component is a direction cosine). It is usually convenient[4,6] to refer the ray directions to a fixed base ray (often the principal when a COS is considered) with a unit ray vector (s_{px}, s_{py}, s_{pz}). In the following we use two independent components only, written as $\mathbf{s} = (s_x, s_y)$ and $\mathbf{s}_p = (s_{px}, s_{py})$. Given a general ray with the direction (s_x, s_y, s_z), we link to it the vector of generalized spatial frequency,

$$\boldsymbol{\nu} = \frac{n}{\lambda}(\mathbf{s} - \mathbf{s}_p), \quad (4)$$

where λ is the vacuum wavelength and n is the refractive index in object space.

Each plane wave front present at the object plane and captured by the optical system is focalized in the second focal plane[14] and gives rise to a quasi-spherical wave front at the image plane. The complex field at the image, $U'(\delta\xi', \delta\eta')$, is

$$U'(\delta\xi', \delta\eta') = \int\int_{-\infty}^{\infty} \hat{U}'(\nu_x', \nu_y')\exp\{-i[2\pi(\nu_x'\delta\xi' + \nu_y'\delta\eta') + b((\delta\xi')^2 + (\delta\eta')^2) + o(\delta\xi', \delta\eta')]\}d\nu_x'd\nu_y', \quad (5)$$

where $2\pi(\nu_x'\delta\xi' + \nu_y'\delta\eta')$ corresponds to the expansion of each wave front up to first order, b takes into account its curvature, and $o(\delta\xi', \delta\eta')$ is a correction to the spherical wave front that appears as a result of aberrations. For small aberrations, b and $o(\delta\xi', \delta\eta')$ can be assumed independent of spatial frequency, and the phase factor $\exp\{ib[(\delta\xi')^2 + (\delta\eta')^2] + o(\delta\xi', \delta\eta')\}$ can be written outside the integral and disappears when the intensity is computed. Then $U'(\delta\xi', \delta\eta')$ is equivalent to a superposition of plane waves with generalized spatial frequencies

$$\boldsymbol{\nu}' = \frac{n'}{\lambda}(\mathbf{s}' - \mathbf{s}_p') \quad (6)$$

such that each spatial frequency $\boldsymbol{\nu}'$ corresponds to a frequency $\boldsymbol{\nu}$.

It is worth pointing out that often intensities, instead of fields, are associated with the picture of geometrical optical rays. Furthermore, the modulation transfer function (MTF) is related to intensities as well. To work in the framework of intensities, first the fields at the object and image are considered,[12,15] and the relation between their angular spectrums is the coherent optical transfer function (COTF). Once the COTF is known, the incoherent optical transfer function (OTF) (whose modulus is the MTF) can be found as the autocorrelation of the COTF.

A basic hypothesis to make these calculations possible is the achievement of local isoplanatism. The fulfillment of the SC is a necessary condition for isoplanatism, and the OSC limits the size of isoplanatic patches and must be analyzed in the framework of field distributions. Afterward the analysis in intensities can be carried out by means of the relation between the COTF and the OTF.

Moreover, in Eq. (3) we consider a field distribution but do not discuss its origin, which depends on the illumination and boundary conditions in the object. A recent numerical approach to this problem with applications to microscope objectives is given by Totzek and Tiziani.[16]

B. Generalized Sine Condition

To obtain the generalized SC by means of Fourier optics,[4,6] we require each Fourier component present in the object [see Eq. (3)] and captured by the optical system to also be present in the image [see Eq. (5)] and to have the same value for any object point $\tilde{Q}$ (in a small patch surrounding Q_ξ) and its corresponding image point $\tilde{Q}'$. This value is

$$\nu_x\delta\xi + \nu_y\delta\eta = \nu_x'\delta\xi' + \nu_y'\delta\eta'. \quad (7)$$

To show why this must hold, let us assume, for simplicity, that the object is a fringe pattern generated by the interference of only two beams, one associated with a spatial frequency ν_{x1} and the other with ν_{x2} ($\nu_{y1} = \nu_{y2} = 0$). If the two frequencies have the same angular spectrum, then $\hat{U}(\nu_x, \nu_y) = A[\delta(\nu_x - \nu_{x1}) + \delta(\nu_x - \nu_{x2})]\delta(\nu_y)$ [$\delta(\ )$ is the delta Dirac function]. With use of Eq. (3), the field distribution at the object is $U(\delta\xi, \delta\eta) = A[\exp(-i2\pi\nu_{x1}\delta\xi) + \exp(-i2\pi\nu_{x2}\delta\xi)]$. The two beams must be captured by the optical system to reproduce this pattern in the image. The beams corresponding to ν_{x1} and ν_{x2} originate beams associated with ν_{x1}' and ν_{x2}', respectively, at the exit. The angular spectrum at the exit is $\hat{U}'(\nu_x', \nu_y') = A'[\delta(\nu_x' - \nu_{x1}') + \delta(\nu_x' - \nu_{x2}')]\delta(\nu_y')$. From Eq. (5), the field at the image (ignoring quadratic and higher-order terms) is $U'(\delta\xi', \delta\eta') = A'[\exp(-i2\pi\nu_{x1}'\delta\xi') + \exp(-i2\pi\nu_{x2}'\delta\xi')]$. If the fringes in the object are well reproduced in the image and if $A' = A$, we must have $\exp(-i2\pi\nu_{x1}'\delta\xi') + \exp(-i2\pi\nu_{x2}'\delta\xi') = \exp(-i2\pi\nu_{x1}\delta\xi) + \exp(-i2\pi\nu_{x2}\delta\xi)$. This is possible for every value of ν_{x2} if $\nu_{x1}'\delta\xi' = \nu_{x1}\delta\xi$. Since ν_{x1} can be any frequency, Eq. (7) must hold.

For a PSOS and with canonical axes,[3] Eq. (7) splits in two; i.e.,

$$\nu_x\delta\xi = \nu_x'\delta\xi', \qquad \nu_y\delta\eta = \nu_y'\delta\eta', \quad (8)$$

and the local transverse magnification matrix $\mathbf{m}$,[6] defined by $(\delta\boldsymbol{\Psi}')^T = \mathbf{m} \times (\delta\boldsymbol{\Psi})^T$ (the superscript T indicates transpose), is diagonal. Calling $m_x = m_{11}$ and $m_y = m_{22}$, we have

$$\delta\xi' = m_x\delta\xi, \qquad \delta\eta' = m_y\delta\eta. \quad (9)$$

To obtain good imagery in the neighborhood of Q_ξ', m_x and m_y must be constant, and combining Eqs. (8) and (9) we get

$$\nu_x' = \nu_x/m_x, \qquad \nu_y' = \nu_y/m_y. \quad (10)$$

Equations (10) are the generalized SC for our PSOS and states that the ratio between a spatial frequency at the image and its corresponding one at the object is equal to a constant that is the inverse of the lateral magnification. According to our previous paper,[6] for a general NCOS the generalized SC in the language of the Fourier optics approach can be paraphrased as follows: Each generalized spatial frequency present in the object and captured by the optical system gives rise in the image to a generalized spatial frequency that differs from it only by an overall constant factor **M**, which is a 2×2 matrix; that is, $\boldsymbol{\nu}' = \mathbf{M} \times \boldsymbol{\nu}$ [where $\mathbf{M} = (\mathbf{m}^T)^{-1}$]. That is, good imagery of the patch is achieved if each spatial frequency at the image is proportional to the corresponding one at the object so that the superposition of frequencies at the object is reproduced at the image. This must also hold for the wave fronts associated with these frequencies [except for the phase factor of Eq. (5) independent of frequency]. When this proportionality fails, there is a misadjustment leading to errors in the image that are evaluated with the OSC.

C. Variation of the Aberration with Field in Terms of the Offense against the Sine Condition

Usually a PSOS does not satisfy the SC [Eqs. (10)]. Considering a hypothetical ray that obeys the SC, i.e., the unit ray vector $(s'_{hx}, s'_{hy}, s'_{hz})$ is such that the associated frequency $\boldsymbol{\nu}'_h$ is

$$\boldsymbol{\nu}'_h = (\nu'_{hx}, \nu'_{hy}) = \left(\frac{\nu_x}{m_x}, \frac{\nu_y}{m_y} \right), \tag{11}$$

the OSC can be written as

$$\delta\boldsymbol{\nu}' = \boldsymbol{\nu}' - \boldsymbol{\nu}'_h = (\delta\nu'_x, \delta\nu'_y) = \left(\nu'_x - \frac{\nu_x}{m_x}, \nu'_y - \frac{\nu_y}{m_y} \right) \neq 0. \tag{12}$$

The OSC is related to the variation of the aberration function with field coordinates. To obtain the variation of W with $\delta\xi'$ in terms of $\delta\nu'_x$, we first consider, for simplicity, light beams in the symmetry plane ($s_y = s_{py} = 0$). If the transverse aberrations for Q_ξ and for $\tilde{Q}$ are zero, we have the situation shown in Fig. 3. A plane wave front emanating from the patch surrounding Q_ξ gives rise at the image to a real locally plane wave front $Q'_\xi J$. We also consider an auxiliary hypothetical plane wave front $Q'_\xi I$. The wave front $Q'_\xi J$ is perpendicular to a general real ray from Q_ξ that does not obey the SC (that is, $\nu'_x = \nu_x/m_x + \delta\nu'_x$), and the wave front $Q'_\xi I$ is perpendicular to the hypothetical ray that obeys the SC. Let δW be the increment in the wave-front aberration that results from the shift of the image point from Q'_ξ to $\tilde{Q}'$, i.e., a distance $\delta\xi'$, when the inclinations of the rays are maintained constant (plane wave fronts). The optical path difference δW is measured along the real ray, perpendicular to the wave front and from the hypothetical wave front to the real one, i.e., $\delta W = n'IJ$. To calculate IJ we consider

Fig. 3. Two locally plane wave fronts arriving at the ideal image point Q'_ξ.

$$s'_x = \frac{J\tilde{Q}'}{\delta\xi'}, \qquad s'_{hx} = \frac{I\tilde{Q}'}{\delta\xi'},$$

$$s'_x - s'_{hx} = \frac{(J\tilde{Q}' - I\tilde{Q}')}{\delta\xi'} = \frac{-\delta W}{n'\,\delta\xi'}. \tag{13}$$

Since $\delta\nu'_x = \nu'_x - \nu'_{hx} = (n'/\lambda)(s'_x - s'_{hx})$, in the limit of small field size ($\tilde{Q}'$ near Q'_ξ) we get

$$\left. \frac{\partial W}{\partial(\delta\xi')} \right|_{\mathbf{u}'} = -\lambda\,\delta\nu'_x, \tag{14}$$

where $\mathbf{u}' = (u'_x, u'_y)$ is the (nongeneralized) spatial frequency associated with the plane wave front perpendicular to the real unit ray vector (s'_x, s'_y, s'_z), that is,

$$\mathbf{u}' = \frac{n'}{\lambda}\mathbf{s}'. \tag{15}$$

Usually in a real PSOS the transverse aberrations for the object points Q_ξ and $\tilde{Q}$ are not zero, and the general real rays from these points arrive at M' and $\tilde{M}'$, respectively, instead of at Q'_ξ and $\tilde{Q}'$ as in Fig. 3. Nevertheless, here we can neglect the variation of transverse aberration with field so that $Q'_\xi M' \approx \tilde{Q}'\tilde{M}'$ and Eq. (14) is valid. Moreover, for beams not necessarily in the symmetry plane, Eq. (14) still holds.[5] Similarly, the variation of W with $\delta\eta'$ can be written in terms of $\delta\nu'_y$. Then the variations of the aberration function with field for $\mathbf{u}'$ constant are

$$\left. \frac{\partial W}{\partial(\delta\xi')} \right|_{\mathbf{u}'} = -\lambda\,\delta\nu'_x, \qquad \left. \frac{\partial W}{\partial(\delta\eta')} \right|_{\mathbf{u}'} = -\lambda\,\delta\nu'_y. \tag{16}$$

4. FIELD DERIVATIVES OF THE WAVE-FRONT ABERRATION

When optical systems are designed, the aberration function is usually considered to depend on aperture [Eqs. (1)] and field [Eqs. (2)] coordinates. Let $W_{Q'_\xi} = W(\boldsymbol{\chi}', 0)$ and $W_{\tilde{Q}'} = W(\boldsymbol{\chi}', \delta\boldsymbol{\Psi}')$ be the aberration functions for a gen-

eral ray from Q_ξ and for the image-space associated ray,[3] respectively, from $\widetilde{Q}$. The aberration function corresponding to point $\widetilde{Q}$ is estimated (up to first order) as follows:

$$W(\chi', \delta\Psi') = W(\chi', 0) + \left.\frac{\partial W}{\partial(\delta\xi')}\right|_{\chi', \delta\Psi'=0} (\delta\xi')$$

$$+ \left.\frac{\partial W}{\partial(\delta\eta')}\right|_{\chi', \delta\Psi'=0} (\delta\eta'). \qquad (17)$$

The expressions for $\partial W/\partial(\delta\xi')|_{\chi', \delta\eta'}$ and for $\partial W/\partial(\delta\eta')|_{\chi', \delta\xi'}$ in terms of variables that are computed through ray tracing can be written in terms of the OSC by use of Eqs. (16). We do this in two steps. In the first step we write $\delta\boldsymbol{\nu}'$ as a function of $\boldsymbol{\chi}'$. In the second we find expressions for the field derivatives of the aberration function for constant aperture coordinates in terms of the variations given in Eqs. (16).

A. Relation between Coordinates and Frequencies

The generalized spatial frequencies in object space can be related to $\boldsymbol{\chi}$ by

$$\boldsymbol{\nu} = \frac{n}{\lambda}(\mathbf{s} - \mathbf{s}_p) = \frac{n}{\lambda(-R)}\boldsymbol{\chi}. \qquad (18)$$

In image space, because of aberrations, we have to apply the approximations $B'M' \approx R'$, $X' - \xi' - \Delta_{\xi'} \approx -R's'_x$, and $Y' - \Delta_{\eta'} = -R's'_y$, yielding

$$\boldsymbol{\nu}' = \frac{n'}{\lambda}(\mathbf{s}' - \mathbf{s}'_p) = \frac{n'}{\lambda R'}(\boldsymbol{\Delta} - \boldsymbol{\chi}') + \mathcal{O}(2), \qquad (19)$$

where we neglect all terms of order 2 or higher in aperture coordinates, signed $\mathcal{O}(2)$ [e.g., $(X'/R')^2$]. Up to the considered order there is a decoupling of the components of the generalized spatial frequencies $\boldsymbol{\nu}'$, the pupil coordinates $\boldsymbol{\chi}'$, and the transverse aberrations $\boldsymbol{\Delta}$; i.e., ν'_x (ν'_y) does not depend on Y' (X') or on $\Delta_{\eta'}$ ($\Delta_{\xi'}$). Defining

$$\boldsymbol{\chi}_m = \left(\frac{X}{m_x}, \frac{Y}{m_y}\right) \qquad (20)$$

and applying Eqs. (18) and (19), we write the OSC of Eq. (12) as

$$\delta\boldsymbol{\nu}' = \frac{(-1)}{\lambda}\left(\frac{n'}{R'}\boldsymbol{\chi}' - \frac{n}{R}\boldsymbol{\chi}_m\right) + \frac{n'}{R'\lambda}\boldsymbol{\Delta}. \qquad (21)$$

B. Field Derivatives of *W* for Constant Aperture Coordinates

The expression for $\partial W/\partial(\delta\xi')|_{\chi', \delta\eta'}$ is obtained by taking into account that the aberration function for a patch surrounding the reference field point Q_ξ can be thought of as function of the variables $(\mathbf{u}', \delta\Psi')$ or $(\boldsymbol{\chi}', \delta\Psi')$ [see Eq. (19) with $(s'_{px}, s'_{py}, s'_{pz})$ constant because Q_ξ is a fixed point]. Then we have[17]

$$\left.\frac{\partial W}{\partial(\delta\xi')}\right|_{\chi', \delta\eta'} = \left.\frac{\partial W}{\partial(\delta\xi')}\right|_{\mathbf{u}'} - \left.\frac{\partial W}{\partial X'}\right|_{Y', \delta\Psi'} \left.\frac{\partial X'}{\partial(\delta\xi')}\right|_{\mathbf{u}'}$$

$$- \left.\frac{\partial W}{\partial Y'}\right|_{X', \delta\Psi'} \left.\frac{\partial Y'}{\partial(\delta\xi')}\right|_{\mathbf{u}'}. \qquad (22)$$

The different terms of Eq. (22) can be evaluated as follows:

$$\left.\frac{\partial W}{\partial(\delta\xi')}\right|_{\mathbf{u}'}.$$

This derivative is related to the OSC by Eqs. (16).

$$\left.\frac{\partial W}{\partial X'}\right|_{Y', \delta\Psi'} \quad \text{and} \quad \left.\frac{\partial W}{\partial Y'}\right|_{X', \delta\Psi'}.$$

These aperture derivatives can be written in terms of variables computed by means of ray tracing with the well-known[3,4] formulas for the transverse aberration; which are

$$\left.\frac{\partial W}{\partial X'}\right|_{Y', \delta\Psi'} = -\frac{n'\Delta_{\xi'}}{R'}, \qquad \left.\frac{\partial W}{\partial Y'}\right|_{X', \delta\Psi'} = -\frac{n'\Delta_{\eta'}}{R'}. \quad (23)$$

$$\left.\frac{\partial X'}{\partial(\delta\xi')}\right|_{\mathbf{u}'} \quad \text{and} \quad \left.\frac{\partial Y'}{\partial(\delta\xi')}\right|_{\mathbf{u}'}.$$

The exit reference sphere coordinates X', Y', and $\widetilde{X}'$, $\widetilde{Y}'$ for the light beams from Q_ξ and $\widetilde{Q}$, respectively, can be calculated from Fig. 2 and Eq. (19). We have $s'_{py} = 0$, $\eta' = 0$, $\delta\eta' = \widetilde{\eta}' \neq 0$, and here $\mathbf{u}' = (n'/\lambda)\mathbf{s}'$ is constant; then (up to first order) the result is

$$X' = \Delta_{\xi'} + \xi' - \frac{u'_x\lambda R'}{n'}, \qquad Y' = \Delta_{\eta'} - \frac{u'_y\lambda R'}{n'},$$

$$\widetilde{X}' = \Delta_{\widetilde{\xi}'} + \widetilde{\xi}' - \frac{u'_x\lambda\widetilde{R}'}{n'}, \qquad \widetilde{Y}' = \Delta_{\widetilde{\eta}'} + \widetilde{\eta}' - \frac{u'_y\lambda\widetilde{R}'}{n'}. \qquad (24)$$

From Fig. 2 we get

$$\frac{\xi'}{R'} = s'_{px}, \qquad R' = [\xi'^2 + (E'O')^2]^{1/2},$$

$$\widetilde{R}' = [(\xi' + \delta\xi')^2 + (\delta\eta')^2 + (E'O')^2]^{1/2}$$

$$\approx R' + \frac{\xi'\delta\xi'}{R'}; \qquad (25)$$

then in the limit of small field size (small $\delta\xi'$) we have $\widetilde{R}' - R' \approx \xi'\delta\xi'/R'$, $(\Delta_{\widetilde{\xi}'} - \Delta_{\xi'}) \approx 0$, and $(\Delta_{\widetilde{\eta}'} - \Delta_{\eta'}) \approx 0$. Therefore from Eq. (24) we obtain

$$\left.\frac{\partial X'}{\partial(\delta\xi')}\right|_{\mathbf{u}'} = 1 - \frac{s'_x\xi'}{R'} = 1 - s'_x s'_{px},$$

$$\left.\frac{\partial Y'}{\partial(\delta\xi')}\right|_{\mathbf{u}'} = -\frac{s'_y\xi'}{R'} = -s'_y s'_{px}. \qquad (26)$$

Substituting Eqs. (16), (23), and (26) into Eq. (22) yields the field derivative as

$$\left.\frac{\partial W}{\partial(\delta\xi')}\right|_{\chi', \delta\eta'} = -\lambda\delta\nu'_x + \frac{n'\Delta_{\xi'}}{R'}(1 - s'_x s'_{px})$$

$$+ \frac{n'\Delta_{\eta'}}{R'}(-s'_y s'_{px}), \qquad (27)$$

which can be written in terms of variables computed by means of ray tracing with use of Eq. (21) for $\delta \nu_x'$.

The expression for $\partial W / \partial (\delta \eta')|_{\chi', \delta \xi'}$ is obtained by considering that Eq. (22) is valid if we interchange $\delta \eta'$ and $\delta \xi'$. From Eq. (24) we have

$$\frac{\partial X'}{\partial (\delta \eta')}\bigg|_{\mathbf{u}'} = 0, \qquad \frac{\partial Y'}{\partial (\delta \eta')}\bigg|_{\mathbf{u}'} = 1, \qquad (28)$$

so that with reasoning similar to that for the other derivative, we obtain

$$\frac{\partial W}{\partial (\delta \eta')}\bigg|_{\chi', \delta \xi'} = -\lambda \, \delta \nu_y' + \frac{n' \Delta_{\eta'}}{R'}. \qquad (29)$$

In Appendix B we compare our results with those of Hopkins.

According to Eqs. (27) and (29) the field derivatives of the wave-front-aberration function are related both to the OSC and to transverse aberrations. These equations can be used to define the size of the isoplanatic patch that allows a given tolerance in the variation of the aberration function [see Eq. (17)]. That is, if $p_x > 1$ and $p_y > 1$ are conveniently predefined numbers and if bars indicate absolute value, the requirements $|\delta \xi' \, \partial W / \partial (\delta \xi')| = 1/p_x |W(\chi', 0)|$ and $|\delta \eta' \, \partial W / \partial (\delta \eta')| = 1/p_y |W(\chi', 0)|$ yield the allowed values of $\delta \xi'$ and $\delta \eta'$.

Furthermore, when systems are designed, various merit functions, such as the wave-front-aberration variance, the Strehl ratio, the COTF, and the MTF[5,12,15] are in common use. All these functions depend on the aberration function W, and this is such that its field derivatives depend on the OSC [Eqs. (27) and (28)]. Then the OSC is related to the field derivatives of these functions. If a merit function and its field derivative are evaluated for the reference image point, then [with use of Eq. (17)] the merit function can be estimated for neighboring image points. Which merit function is adequate depends on which performance is required for the system under consideration. To study decentration tolerances for a PSOS, it is probably more useful in many applications (for example for box-in-box overlay measurements) to define a parameter that characterizes the symmetry of the image.

5. EQUATIONS AVAILABLE AND SYMMETRY OF ABERRATIONS

With use of Eqs. (12), (18), and (19), Eqs. (27) and (29) can be also written in terms of the direction cosines corresponding to real rays. When we take this into account and use Eqs. (23), remembering that X', Y', $\delta \xi'$, and $\delta \eta'$ are canonical coordinates, and if square brackets denote optical pathlength, the complete set of equations available for each ray traced from Q_ξ is

$$W = [Q_\xi E'] - [Q_\xi B'],$$

$$\frac{\partial W}{\partial X'}\bigg|_{Y', \delta \xi', \delta \eta'} = -\frac{n' \Delta_{\xi'}}{R'},$$

$$\frac{\partial W}{\partial Y'}\bigg|_{X', \delta \xi', \delta \eta'} = -\frac{n' \Delta_{\eta'}}{R'},$$

$$\frac{\partial W}{\partial (\delta \xi')}\bigg|_{X', Y', \delta \eta'} = -n'(s_x' - s_{px}') + \frac{n}{m_x}(s_x - s_{px})$$

$$+ \frac{n' \Delta_{\xi'}}{R'} - \frac{n' s_{px}'}{R'}(\Delta_{\xi'} s_x' + \Delta_{\eta'} s_y'),$$

$$\frac{\partial W}{\partial (\delta \eta')}\bigg|_{X', Y', \delta \xi'} = -n' s_y' + \frac{n}{m_y} s_y + \frac{n' \Delta_{\eta'}}{R'}. \qquad (30)$$

In these equations the different behavior of the field derivatives in the directions along and perpendicular to the symmetry plane when the system is a PSOS is evident. For a COS we have the same equations formally, but from the symmetry properties it results that $\partial W / \partial (\delta \eta')$ does not yield an equation independent of the other four.[4] For a PSOS these five equations are independent and completely useful. According to Eq. (17), the aberration function for image points in the neighborhood of Q_ξ' can be estimated (up to first order) only by tracing rays from the reference object point Q_ξ and using Eqs. (30).

It is often useful to study separately the influence of even (subscript e) and odd (subscript d) aberrations present at Q_ξ' on the aberrations at the neighboring image points. For example, for the box-in-box overlay measurements in the microscope industry there are systematic errors in the measured positions that depend on the symmetry of the aberrations.

In Eq. (17), $W(\chi', 0)$ and its field derivatives depend only on (X', Y'). Since the plane $Y' = 0$ is the symmetry plane, all aberrations are even functions of Y'. Then to analyze the parity we consider only the dependence on X'. With vector notation, Eq. (17) is

$$w(X') = w_i(X') + [\boldsymbol{\Omega}(X') + \mathbf{A}(X')](\delta \boldsymbol{\Psi}')^T, \quad (31)$$

where

$$w(X') = W(\chi', \delta \boldsymbol{\Psi}'), \qquad w_i(X') = W(\chi', 0),$$

$$\Omega_x(X') = -\lambda \, \delta \nu_x', \qquad \Omega_y(X') = -\lambda \, \delta \nu_y',$$

$$A_x(X') \approx \frac{n' \Delta_{\xi'}}{R'}, \qquad A_y(X') = \frac{n' \Delta_{\eta'}}{R'}. \qquad (32)$$

In the equation for $A_x(X')$ we neglect the term $[(n' s_{px}')/R'](\Delta_{\xi'} s_x' + \Delta_{\eta'} s_y')$ in comparison with $(n' \Delta_{\xi'})/R'$ because $s_x' \leqslant 1$ and $s_{px}' \ll 1$. In Eq. (31), $\boldsymbol{\Omega}(X') = [\Omega_x(X'), \Omega_y(X')]$ and $\mathbf{A}(X') = [A_x(X'), A_y(X')]$ are row vectors and $(\delta \boldsymbol{\Psi}')^T = (\delta \xi', \delta \eta')^T$ is a column vector. We consider three items:

i.   Aberration function

$$w(X') = w_e(X') + w_d(X'). \qquad (33)$$

The function $w_e(X')$ contains the even terms in X' of the aberration-function expansion[18]: spherical aberration, astigmatism, and field curvature. The function $w_d(X')$ contains odd terms in X': distortion and coma.

According to Eq. (31) an even aberration function at Q_ξ introduces an even aberration function at $\widetilde{Q}'$. For ex-ample, for a COS and an on-axis object, the spherical aberration present at the on-axis image point is also present at a neighboring point.

ii.   Aperture derivatives of the aberration function

$$\mathbf{A}(X') = \mathbf{A}_e(X') + \mathbf{A}_d(X'). \qquad (34)$$

$A_x(X')$ is the derivative of W with respect to X' and changes the parity, whereas $A_y(X')$ is the derivative of W with respect to Y' and does not [see Eqs. (32) and (30)], so
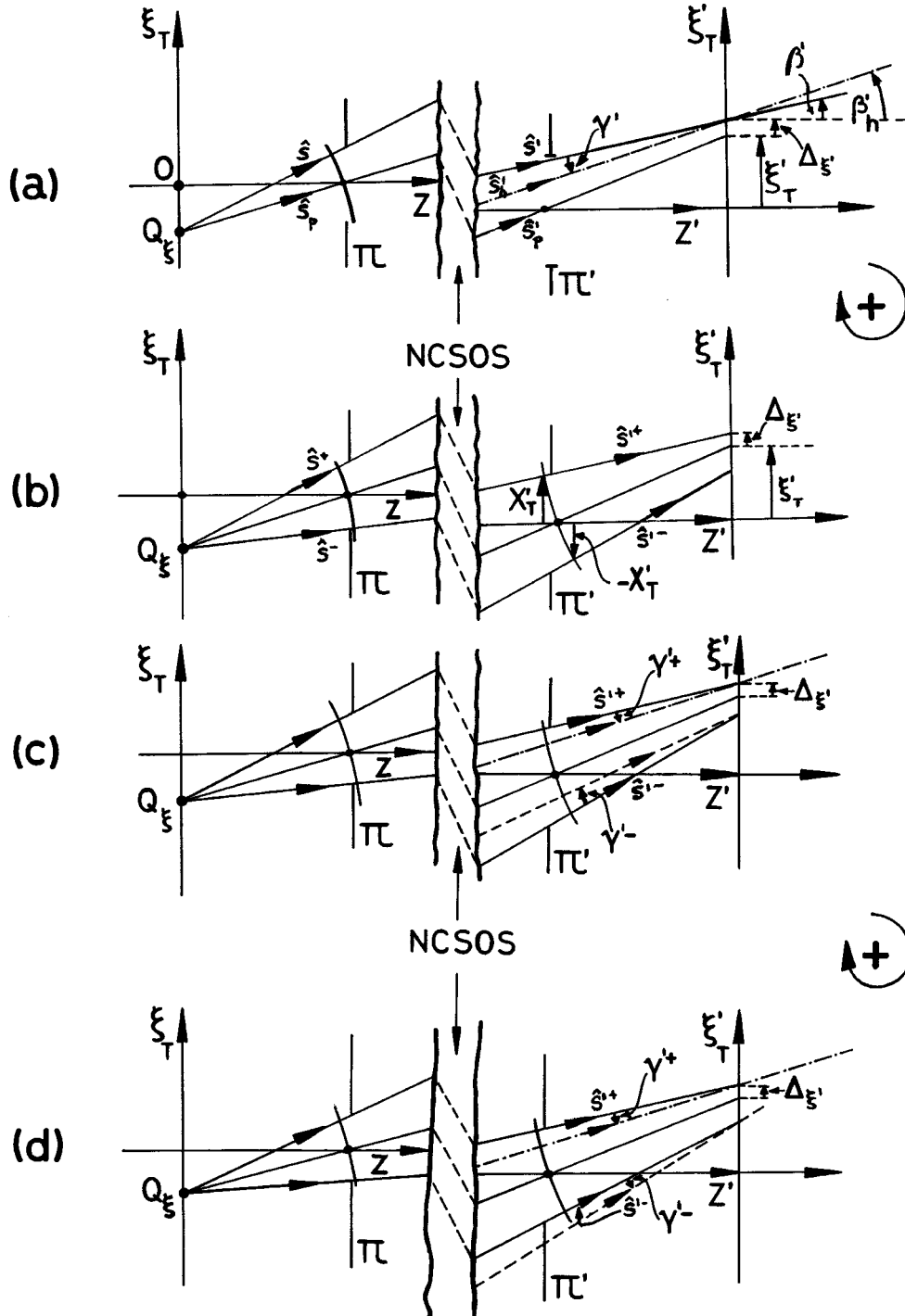
Fig. 4.   OSC: (a) Real and hypothetical rays at the exit that subtend an angle γ'; (b) unit ray vectors $(s_x'^{\,+}, s_y'^{\,+}, s_z'^{\,+})$ and $(s_x'^{\,-}, s_y'^{\,-}, s_z'^{\,-})$; (c) $\gamma'^{\,+} \approx -\gamma'^{\,-}$ (coma at a neighboring point); (d) $\gamma'^{\,+} \approx \gamma'^{\,-}$ (spherical aberration, astigmatism, and field curvature).

$$A_{x,e}(X') = -\left(\frac{\partial W}{\partial X'}\right)_e = -\frac{\partial W_d}{\partial X'},$$

$$A_{x,d}(X') = -\left(\frac{\partial W}{\partial X'}\right)_d = -\frac{\partial W_e}{\partial X'},$$

$$A_{y,e}(X') = -\left(\frac{\partial W}{\partial Y'}\right)_e = -\frac{\partial W_e}{\partial Y'},$$

$$A_{y,d}(X') = -\left(\frac{\partial W}{\partial Y'}\right)_d = -\frac{\partial W_d}{\partial Y'}. \tag{35}$$

According to Eqs. (35), for an even aberration function at Q_ξ we can have $A_{x,d}(X')$ and $A_{y,e}(X')$ different from zero, and then [see Eq. (31)] both even and odd aberration functions are introduced at $\tilde{Q}'$ (similarly for an odd aberration). For example, for the well-known case of a COS and an on-axis object, if for the on-axis object point there is spherical aberration, then this aberration generates coma at a neighboring field point.

   iii.   Offense against the sine condition

$$\mathbf{\Omega}(X') = \mathbf{\Omega}_e(X') + \mathbf{\Omega}_d(X'). \tag{36}$$

For a better understanding of the OSC we consider two rays from Q_ξ in the symmetry plane: the real one and the hypothetical one with $(s'_{hx}, s'_{hy}, s'_{hz})$ satisfying the SC [see Fig. 4(a)]. Then $\Omega_y(X') = -\lambda \delta\nu'_y = 0$ and $\Omega_x(X') = -\lambda \delta\nu'_x = n'(-s'_x + s'_{hx})$, and we have

$$\Omega_{x,e}(X') = \tfrac{1}{2}[\Omega_x(X') + \Omega_x(-X')],$$

$$\Omega_{x,d}(X') = \tfrac{1}{2}[\Omega_x(X') - \Omega_x(-X')]. \tag{37}$$

We refer to γ' as the angle from the ray vector (s'_x, s'_y, s'_z) to $(s'_{hx}, s'_{hy}, s'_{hz})$ and choose the mathematical sign convention that angles count positive when they are counterclockwise. For a small OSC, $\gamma' \approx \sin\beta'_h - \sin\beta'$ (in this case $\sin\beta' = s'_x$), and we can write $\Omega_x(X') \approx n'\gamma'$. For rays with exit reference sphere coordinates $X' > 0$ and $-X' < 0$, the unit ray vectors at the exit are termed $(s'^{+}_x, s'^{+}_y, s'^{+}_z)$ and $(s'^{-}_x, s'^{-}_y, s'^{-}_z)$, respectively [see Fig. 4(b)]. If $X' > 0$ we have

$$\Omega_x(X') = n'(s'^{+}_{hx} - s'^{+}_x) \approx n'\gamma'^{+},$$

$$\Omega_x(-X') = n'(s'^{-}_{hx} - s'^{-}_x) \approx n'\gamma'^{-}. \tag{38}$$

In Fig. 4(c) we have $\gamma'^{+} \approx -\gamma'^{-}$, and according to Eqs. (38), $\Omega_x(X') = -\Omega_x(-X')$; that is, $\Omega_{x,d}(X') \neq 0$ and $\Omega_{x,e}(X') = 0$ [see Eqs. (37)], so that there is an *odd* OSC. On the other hand, in Fig. 4(d) we have $\gamma'^{+} \approx \gamma'^{-}$, and according to Eqs. (38), $\Omega_x(X') = \Omega_x(-X')$; that is [see Eqs. (37)], we have $\Omega_{x,d}(X') = 0$ and $\Omega_{x,e}(X') \neq 0$, so there is an even OSC.

The odd OSC [as in Fig. 4(c)] makes no contribution to the even aberration function at $\tilde{Q}'$, while the even OSC [as in Fig. 4(d)] makes no contribution to the odd aberration function at $\tilde{Q}'$ [see Eq. (31)]. For example, for the well-known case of a COS and an axial object, we can have only $\gamma'^{+} \approx -\gamma'^{-}$; that is, $\Omega_{x,d}(X') \neq 0$ and $\Omega_{x,e}(X') = 0$, so that there is only an odd OSC, and it introduces coma.

6. CONCLUSION

The generalized sine condition for a perturbated symmetrical optical system can be obtained by means of Fourier optics, requiring the spatial frequencies present in the image to be proportional to those at the object. The unfulfillment of this condition is related to the field derivatives of the wave-front-aberration function. The offense against the generalized sine condition can be used to analyze decentration tolerances when microscope objectives are designed and produced. This analysis is performed assuming a perturbation in the system, tracing rays and using the formulas of Section 5.

APPENDIX A: GENERAL DESCRIPTION OF COORDINATE SYSTEMS

Following Hopkins[3] (except that we interchange X with Y to follow the notation of previous papers[4]), for our PSOS we define as a starting point three types of coordinate systems: base, local, and canonical. These coordinates are shown in Fig. 5 (for simplicity we draw the symmetry plane alone). The surfaces of the object and the receiving screens coincide with the surfaces S and S', respectively, of the reference COS. We consider aperture reference planes at the entrance and the exit, which are termed Π and Π', respectively. For the coordinates we have the following.

1. Base Coordinates
We consider an object point O_o in the symmetry plane (our results do not depend on the localization of O_o, and O_o may or may not coincide with the central point $\tilde{O}_o$), and we choose O'_o as its image obtained with the PSOS on the surface S'. We choose axes O_oE_o and $E'_oO'_o$ perpendicular to S and S', respectively, and therefore $E'_oO'_o$ is parallel to O_oE_o. (For a NCOS, $E'_oO'_o$ is usually not the emergent ray corresponding to O_oE_o.) The points E_o and E'_o are the intersection of the axes O_oE_o and $E'_oO'_o$ with the planes Π and Π', respectively. The base coordinates at the object are (ξ_o, η_o, ζ_o) with origin at point O_o, ζ_o along O_oE_o and ξ_o in the symmetry plane. The base coordinates at the image are $(\xi'_o, \eta'_o, \zeta'_o)$ with origin at O'_o, ζ'_o along the prolongation of $E'_oO'_o$ and ξ'_o parallel to ξ_o (η'_o and η_o are perpendicular to the symmetry plane). The base coordinates at the entrance reference sphere are (X_o, Y_o, Z_o) with origin at E_o, Z_o along O_oE_o and (X_o, Y_o) parallel to (ξ_o, η_o). The base coordinates at the exit reference sphere are (X'_o, Y'_o, Z'_o) with origin at E'_o.

2. Local Coordinates
We take into account an object point Q_ξ that lies on the axis ξ_o (Q_ξ can be different from O_o). We consider a base ray $Q_\xi E$. For a general NCOS the base ray is the central ray of the beam in any conveniently defined sense,[3] and for a COS the base ray is usually termed the principal ray (to avoid confusion we reserve the term principal ray for a COS and adopt the term base ray for a NCOS). The base ray intersects the plane Π at point E and emerges to cut the plane Π' at point E' and the image surface S' at Q'_ξ, which is considered[3] to be the geometrical image of $\tilde{Q}_\xi$ (usually the base ray is not perpendicular to S and S').
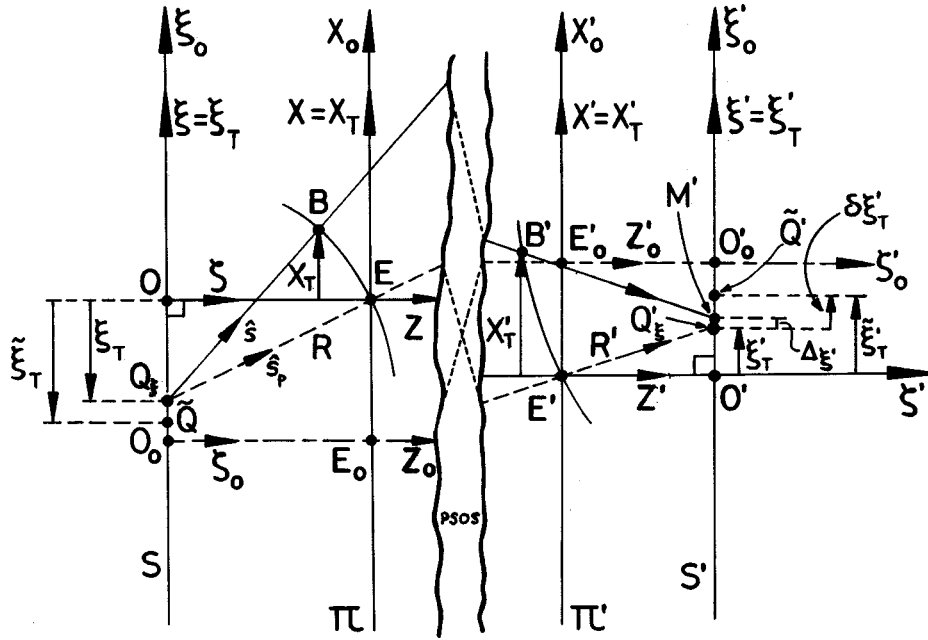
Fig. 5.   Coordinate systems for a PSOS: O_o, object point in the symmetry plane with image at O_o'; (ξ_o, η_o, ζ_o) $[(\xi_o', \eta_o', \zeta_o')]$, base coordinates at the object (image) with origin at O_o (O_o'); (X_o, Y_o, Z_o) $[(X_o', Y_o', Z_o')]$, base coordinates at the entrance (exit) reference sphere with origin at E_o (E_o'); Q_ξ, object point on the axis ξ_0; (ξ, η, ζ) $[(\xi', \eta', \zeta')]$, local coordinates at the object (image) with origin at O (O') for the light beam from Q_ξ; (X, Y, Z) $[(X', Y', Z')]$, local coordinates at the entrance (exit) reference sphere with origin at E (E'); (X_T, Y_S) $[(X_T', Y_S')]$, canonical axes in the reference spheres; (ξ_T, η_S) $[(\xi_T', \eta_S')]$, canonical coordinates in the object (image).

The local coordinates at the object (ξ, η, ζ) are such that the origin is point O (defined as the foot of the perpendicular from E to the plane S), the axis ζ is along OE, and the axis ξ is parallel to ξ_0.   Similarly, the local coordinates at the image are (ξ', η', ζ') with origin at point O'.   The local coordinates at the entrance reference sphere are (X, Y, Z) with origin at E, with (X, Y) parallel to (ξ, η) and the axis Z on the prolongation of OE.   Similarly, the local coordinates at the exit reference sphere are (X', Y', Z') with origin at E'.

3. Canonical Coordinates

According to Hopkins's basic theorem,[3] the canonical coordinates at the entrance and exit reference spheres, (X_T, Y_S) and (X_T', Y_S'), respectively, are obtained by rotating the local coordinates (X, Y) and (X', Y') around the Z and Z' axes, respectively, so that all the rays entering the optical system with $Y_S = 0$ emerge with $Y_S' = 0$ and those entering with $X_T = 0$ emerge with $X_T' = 0$.   For our PSOS we choose local axes in such a way that the plane $Y = 0$ is the symmetry one, so (up to first order) the rays that lie in the symmetry plane at the entrance also lie in the symmetry plane at the exit, whereas the rays entering the system in a plane perpendicular to the symmetry one must lie in a plane perpendicular to the symmetry one.   Therefore the canonical axes (X_T, Y_S) and (X_T', Y_S') coincide, respectively, with the local axes (X, Y) and (X', Y') defined above.   The canonical coordinates in the object and the image, (ξ_T, η_S) and (ξ_T', η_S'), must be parallel[3] to (X_T, Y_S) and (X_T', Y_S'), respectively, and therefore for our PSOS they are identical to (ξ, η) and (ξ', η'), respectively.

APPENDIX B: COMPARISON WITH HOPKINS FORMULAS

1. Relation between Coordinates and Spatial Frequencies

Equations (18) and (19) can be compared with the Hopkins results.[3]   For an object point Q (which he names $\tilde{Q}$) he considers direction cosines of the general and the base rays (L, M, N) and $(\tilde{L}, \tilde{M}, \tilde{N})$, respectively, defines $\tilde{\lambda} = L - \tilde{L}$ and $\tilde{\mu} = M - \tilde{M}$ [then $\tilde{\lambda}(n/\lambda) = \nu_x$, $\tilde{\mu}(n/\lambda) = \nu_y$], and has $\tilde{\lambda} = L - \tilde{L} = -X/R$ and $\tilde{\mu} = M - \tilde{M} = -Y/R$, which coincide with our Eq. (18).   For the image point Q' he writes in his Eq. (3.15) the canonical coordinates X_T', $\tilde{\lambda}_T'$, Y_S' and $\tilde{\mu}_S'$ in terms of (X_T, Y_S).   From this equation, for a general optical system one can obtain that both $(\nu_x'\lambda)/n' = \tilde{\lambda}_T'$ and $(\nu_y'\lambda)/n' = \tilde{\mu}_S'$ are proportional to both X_T' and Y_S'.   From our Eq. (19) we have for our PSOS that ν_y' is independent of X_T' because all the rays in the fan on the plane $Y_S' = 0$, which focus at the tangential image point (here $\Delta_{\eta'} = 0$), have $s_y' = 0$ and for the base ray $s_{py}' = 0$, so that $\nu_y' = 0$.   Moreover, ν_x' is independent of Y_S' because the rays in the fan such that $X_T' = 0$ focus in the sagittal image point (here $\Delta_{\xi'} = 0$), which is on the base ray, and have $s_x' \neq 0$ but (up to first order) $s_x' \approx s_{px}'$ and $\nu_x' = 0$ for any Y_S'.   In Eq. (19) we also have that ν_x' (ν_y') is proportional not only to X_T' (Y_S') but also to $\Delta_{\xi'}$ $(\Delta_{\eta'})$.   Then our expressions coincide with those of Hopkins except for transverse aberrations.

2. Generalized Sine Condition

We obtain the SC by starting from Fourier optics, whereas Hopkins uses geometrical optics.   Using Eqs.

(18), (19), and (9), our generalized SC [Eqs. (10)] can be written as

$$\frac{(\widetilde{\xi}'_T - \xi'_T)n'}{R'}(X'_T - \Delta_{\xi'}) = \frac{(\widetilde{\xi}_T - \xi_T)n}{R}X_T,$$

$$\frac{(\widetilde{\eta}'_S - \eta'_S)n'}{R'}(Y'_S - \Delta_{\eta'}) = \frac{(\widetilde{\eta}_S - \eta_S)n}{R}Y_S. \qquad (B1)$$

Hopkins defines normalized canonical coordinates at the reference spheres and at the field (x_T, y_S), (x'_T, y'_S) and (G_T, H_S), (G'_T, H'_S), respectively (here we interchange T and S because we interchange X and Y to follow the notation of our previous work[5]). Hopkins has

$$x_T = \frac{X_T}{\widetilde{\widetilde{X}}_T}, \qquad\qquad y_S = \frac{Y_S}{\widetilde{\widetilde{Y}}_S},$$

$$x'_T = \frac{X'_T}{\widetilde{\widetilde{X}}'_T}, \qquad\qquad y'_S = \frac{Y'_S}{\widetilde{\widetilde{Y}}'_S},$$

$$G_T = \frac{(\widetilde{\xi}_T - \xi_T)n\widetilde{\widetilde{X}}_T}{R}, \qquad H_S = \frac{(\widetilde{\eta}_S - \eta_S)n\widetilde{\widetilde{Y}}_S}{R},$$

$$G'_T = \frac{(\widetilde{\xi}'_T - \xi'_T)n'\widetilde{\widetilde{X}}'_T}{R'}, \qquad H'_S = \frac{(\widetilde{\eta}'_S - \eta'_S)n'\widetilde{\widetilde{Y}}'_S}{R'}. \qquad (B2)$$

($\widetilde{\widetilde{X}}_T$ and $\widetilde{\widetilde{Y}}_S$ are the coordinates at the entrance reference sphere of the marginal rays corresponding to $Y_S = 0$ and $X_T = 0$, respectively, and the quantities $\widetilde{\widetilde{X}}'_T = \widetilde{a}_{11}\widetilde{\widetilde{X}}_T$ and $\widetilde{\widetilde{Y}}'_S = \widetilde{a}_{31}\widetilde{\widetilde{Y}}_S$ are employed as normalization factors for the exit reference sphere.) From Eqs. (40) and (41), our SC written in terms of normalized canonical coordinates is

$$G'_T\left(x'_T - \frac{\Delta_{\xi'}}{\widetilde{\widetilde{X}}'_T}\right) = G_T x_T, \qquad H'_S\left(y'_S - \frac{\Delta_{\eta'}}{\widetilde{\widetilde{Y}}'_S}\right) = H_T y_S. \qquad (B3)$$

This equation differs from the Hopkins SC, $x'_T = x_T$ and $y'_S = y_S$ (with $G_T = G'_T$ and $H_S = H'_S$), only in terms that are proportional to transverse aberrations.

3. Field Derivatives of the Wave-Front-Aberration Function

The field derivatives of Eqs. (27) and (29) can also be written in terms of the normalized canonical coordinates of Eqs. (41) since $\delta\nu'_x$ (or $\delta\nu'_y$) are related to X'_T and X_T (or Y'_S and Y_S) by Eq. (21). We have $\partial W/\partial(\delta\xi'_T) = (n'\widetilde{\widetilde{X}}'_T/R')(\partial W/\partial G'_T)$ and $\partial W/\partial(\delta\eta'_S) = (n'\widetilde{\widetilde{Y}}'_S/R') \times (\partial W/\partial H'_S)$; we define, as does Hopkins,[3] reduced transverse aberrations $\Delta_{G'} = n'\widetilde{\widetilde{X}}'_T\Delta_{\xi'}/R'$ and $\Delta_{H'} = n'\widetilde{\widetilde{Y}}'_S\Delta_{\eta'}/R'$ and consider[3] $m_x = n\widetilde{\widetilde{X}}_T R'/n'\widetilde{\widetilde{X}}'_T R$ and $m_y = n\widetilde{\widetilde{Y}}_S R'/n'\widetilde{\widetilde{Y}}'_S R$ [see Eqs. (9) and (41)] so that we obtain

$$\left.\frac{\partial W}{\partial G'_T}\right|_{x'_T,y'_S,H'_S} = x'_T - x_T - \frac{R's'_{px}}{n'}\left(\frac{\Delta_{G'}s'_x}{\widetilde{\widetilde{X}}'^2_T} + \frac{\Delta_{H'}s'_y}{\widetilde{\widetilde{X}}'_T\widetilde{\widetilde{Y}}'_S}\right),$$

$$\left.\frac{\partial W}{\partial H'_S}\right|_{x'_T,y'_S,G'_T} = y'_S - y_S. \qquad (B4)$$

These equations differ only in the transverse-aberration terms that are proportional to s'_{px} from Eq. (6.20) of Hopkins's paper,[3] which are $\partial W/\partial G'_T = x'_T - x_T$ and $\partial W/\partial H'_S = y'_S - y_S$.

REFERENCES

1. A. Cox, *System of Optical Design* (Focal, New York, 1964).
2. S. A. Comastri and J. M. Simon, "Aberration function dependence on field, a way to obtain better profit from ray tracing," Optik (Stuttgart) **69**, 135–140 (1985).
3. H. H. Hopkins, "Image formation by a general optical system. 1. General theory," Appl. Opt. **24**, 2491–2505 (1985).
4. S. A. Comastri and J. M. Simon, "Ray tracing, aberration function and spatial frequencies," Optik (Stuttgart) **66**, 175–190 (1984).
5. S. A. Comastri and J. M. Simon, "Field derivative of the variance of wavefront aberration," J. Mod. Opt. **36**, 1073 (1989).
6. J. M. Simon and S. A. Comastri, "Image-forming systems: matrix formulation of the optical invariant via Fourier optics," J. Mod. Opt. **43**, 2533–2541 (1996).
7. M. Mansuripur, "Abbés sine condition," Opt. Photonics News **9**(2), 56–60 (1998).
8. R. Blendowske, "Fourier optical approach to the sine condition," Opt. Photonics News **9**(2), 6 (1998).
9. R. Blendowske and W. Vollrath, "Strehl ratio split for production-limited optics," Opt. Photonics News **8**(5), (1997).
10. R. Blendowske, U. Voigt, and W. Vollrath, "Delta optics: theoretical aspects of design and production" Sci. Technol. Inf. **10**, 147–155 (1993).
11. S. A. Comastri, "Pupil exploration and calculation of vignetting," Optik (Stuttgart) **85**, 173–176 (1990).
12. J. W. Goodman, *Introduction to Fourier Optics* (McGraw-Hill, New York, 1968).
13. S. A. Comastri, "Optical systems and spatial frequencies," Optik (Stuttgart) **105**, 129–133 (1997).
14. R. S. Longhurst, *Geometrical and Physical Optics* (Longman, London, 1973).
15. M. Born and B. Wolf, *Principles of Optics* (Pergamon, Oxford, 1987).
16. M. Totzek and H. J. Tiziani, "Interference microscopy of sub λ structures: a rigorous computation method and measurements," Opt. Commun. **136**, 61–74 (1997).
17. J. Rey Pastor, P. Pi Calleja, and C. Trejo, *Analisis Matematico* (Kapelusz, Argentina, 1957), Vol. II, p. 158.
18. V. Mahajan, "Zernike circle polynomials and optical aberrations of systems with circular pupils," Appl. Opt. , 8121–8124 (1994).