Extended Nijboer–Zernike approach for the computation of optical point-spread functions

Augustus J. E. M. Janssen

Philips Research Laboratories WY-81, 5656 AA Eindhoven, The Netherlands

Received May 24, 2001; revised manuscript received October 1, 2001; accepted October 2, 2001

New Bessel-series representations for the calculation of the diffraction integral are presented yielding the point-spread function of the optical system, as occurs in the Nijboer–Zernike theory of aberrations. In this analysis one can allow an arbitrary aberration and a defocus part. The representations are presented in full detail for the cases of coma and astigmatism. The analysis leads to stably converging results in the case of large aberration or defocus values, while the applicability of the original Nijboer–Zernike theory is limited mainly to wave-front deviations well below the value of one wavelength. Because of its intrinsic speed, the analysis is well suited to supplement or to replace numerical calculations that are currently used in the fields of (scanning) microscopy, lithography, and astronomy. In a companion paper [J. Opt. Soc. Am. A **19**, 860 (2002)], physical interpretations and applications in a lithographic context are presented, a convergence analysis is given, and a comparison is made with results obtained by using a numerical package. © 2002 Optical Society of America

OCIS codes: 050.1960, 070.2580, 000.3860, 000.4430.

1. INTRODUCTION

The Nijboer–Zernike theory of diffraction integrals was intended to produce an analytical result that, with the computational means of that time, led to a good approximation for the intensity distribution in or close to the focal plane. As indicated already by Nijboer himself, the permitted wave-front aberration should not, in terms of the phase of the exponential factor, exceed a few radians. The amplitude distribution over the wave front is limited to a uniform one. These two conditions mean that the analysis is not well suited to the solution of practical problems, where the defocusing and the aberrations may be relatively large and where the uniformity condition on the amplitude is too restrictive. For this reason the analytical method developed by Nijboer has not been extensively used to solve practical problems; numerical integration or Fourier transform methods are used to calculate the image intensity profiles. Neither can one find recent publications in the optics literature aimed at application or further developments of the Nijboer–Zernike theory.

In this paper an extended analytical approach is proposed that does not suffer from the small-aberration and amplitude uniformity restrictions. In the remainder of this paper, we will focus on the exact treatment of large aberrations and/or defocus terms in the diffraction integral. It will become evident that amplitude nonuniformity can also be treated in this new approach along the same lines as the treatment given the (strong-) aberration term in the exponential. One is thus led to the conclusion that with these extensions the Nijboer–Zernike approach has become an interesting alternative to the common numerical methods, both in accuracy and in calculation speed. In fact, numerical experiments with defocus values as large as 2π and/or aberrations of the same order of magnitude show the validity of the extended approach (an example is contained in Section 4 for the case of coma with defocus values up to 2π). In a forthcoming paper[1] these issues will be investigated in more detail.

We first consider the Nijboer–Zernike theory of diffraction integrals containing small aberrations. As is well-known, these diffraction integrals apply to optical systems where the pupil is large compared with the wavelength of the light used. For a summary of the diffraction theory of aberrations and the relevant aspects of the Nijboer–Zernike theory, the reader is referred to Ref. 2, Chap. 9, in particular Secs. 9.1–9.4 and Appendix VII, which summarize Nijboer's thesis;[3] more recent literature on Zernike polynomials and their use and interpretation in the context of aberrated circular optical systems include Refs. 4–7. We thus consider, as in Ref. 2, Chap. 9, a point source of monochromatic light in the object plane of a centered optical system (see Fig. 1), and we assume the distortions to be symmetrical about the meridional plane, so that the aberration functions are even functions of the angular coordinate θ in the exit pupil. In the normalized form that is convenient for our purposes, the diffraction integrals that we are concerned with take the form

$$U(x, y) = \frac{1}{\pi}\iint_{\nu^2+\mu^2\leqslant 1} \exp[i(\nu^2 + \mu^2)f + i\Phi(\nu, \mu)] \times \exp(2\pi i\nu x + 2\pi i\mu y)\mathrm{d}\nu\mathrm{d}\mu. \tag{1}$$

Here $U(x, y)$ is the normalized point-spread function, with x, y the spatial Cartesian coordinates in the image plane, $\Phi(\nu, \mu)$ is the aberration function, with ν, μ the Cartesian coordinates in the exit pupil representing spatial frequencies, and f is a parameter representing defocusing. From a mathematical point of view, there is nothing that prevents us from assuming that Φ in Eq. (1)

0740-3232/2002/050849-09$15.00

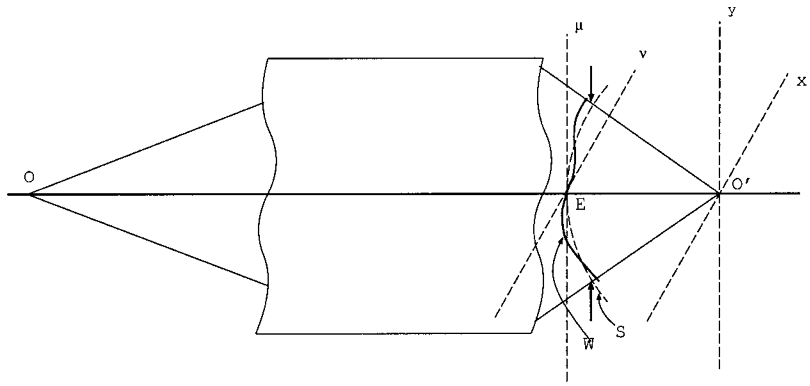

Fig. 1. A point source at O emits a spherical wave toward the schematically represented optical system. In the image space an aberrated wave front W leaves the exit pupil (center at E) and comes to a focus close to the image plane through O′. The spherical reference wave front is denoted by S, and the wave-front aberration is given by the perpendicular distance between S and W. The phase function Φ is derived from W through $\Phi = 2\pi W/\lambda$, where λ is the wavelength of the monochromatic radiation. The normalized Cartesian pupil coordinates are denoted by (ν, μ), and the coordinates (x, y) in the image plane have been normalized with respect to the diffraction unit λ/NA, where NA is the image-side numerical aperture of the optical system. Note that the analysis in this paper is not limited to on-axis object and image points.

is complex valued, so that nonuniform amplitude distributions can be accommodated as well.

In the Nijboer–Zernike theory one aims at expressing $U(x, y)$ in numerically tractable forms, where one should be well aware that no powerful computational aids such as we have now were available at the time the theory was developed. For values of defocus parameter f and maximum $|\Phi|$ of order well below unity, $U(x, y)$ can be computed with sufficient accuracy by using the results of the Nijboer–Zernike theory, where a modest number of Bessel functions must be evaluated. See Ref. 2, Figs. 9.6 and 9.9, where some contour plots of $|U(x, y)|^2$ are shown that were drawn by Nijboer[3] and Nienhuis[8] in their respective theses by using this approach.

Let us sketch the basic features of the Nijboer–Zernike theory. One starts by expanding the aberration function Φ as

$$\Phi(\nu, \mu) \equiv \Phi(\rho, \theta) = \sum_{n,m} \alpha_{nm} R_n^m(\rho) \cos m\theta, \qquad (2)$$

where (with slight abuse of notation) we use polar coordinates $\rho \exp(i\theta) = \nu + i\mu$ (see Ref. 2, Sec. 9.2.2). Here R_n^m is a Zernike polynomial, so that the expansion in Eq. (2) is orthogonal on the unit circle; see Ref. 2, Sec. 9.2.1 and Appendix VII, for the main properties of the Zernike polynomials. In particular, we have that the series on the right-hand side of Eq. (2) can be taken over all integers n, $m \geqslant 0$ with $n - m$ even and $\geqslant 0$. Also, $R_n^m(\rho)$ is a polynomial in ρ of degree n containing the powers $\rho^n, \rho^{n-2}, \ldots, \rho^m$ only. Then, using polar coordinates for the integral on the right-hand side of Eq. (1), we have

$$U(x, y) = \frac{1}{\pi} \int_0^1 \rho \exp(if\rho^2) \times \left\{ \int_0^{2\pi} \exp[i\Phi(\rho, \theta + \phi)] \times \exp(2\pi i\rho r \cos\theta) \mathrm{d}\theta \right\} \mathrm{d}\rho, \qquad (3)$$

where $x + iy = r\exp(i\phi)$. The inner integral on the right-hand side of Eq. (3) is expanded as

$$\int_0^{2\pi} \exp[i\Phi(\rho, \theta + \phi)] \exp(2\pi i\rho r \cos\theta) \mathrm{d}\theta = \sum_{k=0}^{\infty} \frac{i^k}{k!} \int_0^{2\pi} \Phi^k(\rho, \theta + \phi) \exp(2\pi i\rho r \cos\theta) \mathrm{d}\theta. \qquad (4)$$

As to the first-order term ($k = 1$) on the right-hand side of Eq. (4), we get, by using elementary properties of the Bessel functions,

$$\int_0^{2\pi} \Phi(\rho, \theta + \phi) \exp(2\pi i\rho r \cos\theta) \mathrm{d}\theta = \sum_{n,m} \alpha_{nm} R_n^m(\rho) \int_0^{2\pi} [\cos m(\theta + \phi)] \times \exp(2\pi i\rho r \cos\theta) \mathrm{d}\theta = 2\pi \sum_{n,m} \alpha_{nm} i^m R_n^m(\rho) J_m(2\pi\rho r) \cos m\phi. \qquad (5)$$

Thus, as to first-order terms, evaluation of U in Eq. (3) requires computation of the integrals

$$\int_0^1 \rho \exp(if\rho^2) R_n^m(\rho) J_m(2\pi\rho r) \mathrm{d}\rho, \qquad (6)$$

with integers n, $m \geqslant 0$ such that $n - m$ is even and $\geqslant 0$. It is a key result in the Nijboer–Zernike theory that

$$\int_0^1 \rho R_n^m(\rho) J_m(2\pi\rho r) \mathrm{d}\rho = (-1)^{(n-m)/2} \frac{J_{n+1}(v)}{v}, \qquad v = 2\pi r. \qquad (7)$$

Hence the integrals in expression (6) admit a simple expression when the defocus parameter f vanishes. For $f \neq 0$ it is customary in the Nijboer–Zernike theory to use Bauer's identity and some further properties of the Zernike polynomials, yielding

$$\exp(if\rho^2) = \exp\left(\frac{1}{2}if\right)\sqrt{\frac{\pi}{f}}\sum_{s=0}^{\infty}(2s + 1)i^s J_{s+1/2}\left(\frac{1}{2}f\right)R_{2s}^0(\rho). \tag{8}$$

The next step for evaluating the integrals in expression (6) is then to write the products $R_n^m(\rho)\,R_{2s}^0(\rho)$, which turn up when Eq. (8) is inserted into expression (6), as a finite series:

$$R_n^m(\rho)R_{2s}^0(\rho) = \sum_p A_p R_p^m(\rho). \tag{9}$$

Then Eq. (7) can be used for each of the terms on the right-hand side of Eq. (9). This then yields a series representation of the integrals in expression (6) that converges sufficiently fast, so that only a modest number of terms is enough to accurately approximate the integrals, provided that f is not too large.

The problematic point in the Nijboer–Zernike approach is the determination of the coefficients A_p in the series on the right-hand side of Eq. (9). While one can show that representations (9) do exist for any allowed tuple n, m, s, the explicit determination of the A_p is not easy; when n and m are not too large, these A_p can be determined (and this is in fact done this way) by using the explicit form of the Zernike polynomials. This same problem becomes even more serious when, on the right-hand side of Eq. (4), terms of order $k \geqslant 2$ have to be included, for then one needs to consider products of $k + 1$ Zernike polynomials, which should be represented as series of the form (9) with an appropriate m, so as to be able to use the basic formula (7). Interestingly, after this paper was accepted, the author obtained a systematic way to determine coefficients in series of the Eq. (9) type for products of Zernike polynomials. This is based on the formula, valid for $m, k = 0, 1, \ldots,$

$$\rho^{m+2k} = \sum_{p=0}^{k} \frac{m + 2p + 1}{m + p + k + 1} \frac{\binom{k}{p}}{\binom{m + k + p}{p}} R_{m+2p}^m(\rho), \tag{10}$$

together with the explicit representation of Zernike polynomials as finite series of monomials (see Section 3). Nevertheless, even with this coefficient problem solved in principle, the whole procedure remains rather cumbersome.

In this paper the above situation is remedied by giving explicit Bessel-series representations for the integrals in expression (6), as well as for integrals as those in expression (6) but with $R_n^m(\rho)$ replaced with ρ^n, where $n - m$ is even and $\geqslant 0$. The integrals involving $R_n^m(\rho)$ are directly useful when the (possibly nonuniform) aberration $A\exp(i\Phi)$, rather than Φ itself, has been given in the form of a Zernike series. The integrals involving ρ^n can be used when Φ has been given in the form of a Zernike series, as in Eq. (2), by expanding any Φ^k on the right-hand side of Eq. (4) as a series of terms of the form $\rho^n \cos m(\theta + \phi)$ with integers $n, m \geqslant 0$ such that $n - m \geqslant 0$ and even. Full details are given for the two cases in which the aberration represents coma and astigmatism.

2. SUMMARY OF RESULTS

In Section 3 we shall explicitly express any of the terms

$$\int_0^{2\pi} \Phi^k(\rho,\ \theta + \phi)\exp(2\pi i\rho r\cos\theta)\mathrm{d}\theta \tag{11}$$

that occur in the right-hand series in Eq. (4) in terms of the functions

$$\rho^n J_m(2\pi\rho r)\cos m\phi \tag{12}$$

with integers $n, m \geqslant 0$ such that $n - m$ is even and $\geqslant 0$. Then, in Appendix A.1, we consider the integrals

$$T_{nm} := \int_0^1 \rho^{n+1}\exp(if\rho^2)J_m(2\pi\rho r)\mathrm{d}\rho. \tag{13}$$

For integers $n, m \geqslant 0$ such that $n - m$ is even and $\geqslant 0$, we show the explicit result

$$T_{nm} = \exp(if)\sum_{l=1}^{\infty}(-2if)^{l-1}\sum_{j=0}^{p} t_{lj}\frac{J_{m+l+2j}(v)}{v^l}, \tag{14}$$

where

$$v = 2\pi r, \qquad p = \tfrac{1}{2}(n - m), \qquad q = \tfrac{1}{2}(n + m), \tag{15}$$

$$t_{lj} = (-1)^j \frac{m + l + 2j}{q + 1}\binom{p}{j} \times \binom{m + j + l - 1}{l - 1} \Big/ \binom{q + l + j}{q + 1},$$
$$j = 0, 1, \ldots, \quad l = 1, 2, \ldots. \tag{16}$$

Thus, when the results of Appendix A.1 are combined, the integral expression (3) for U can be expressed explicitly, in terms of powers of f, the α_{nm} in Eq. (2), and Bessel functions, in a way that is, in principle, practicable for computerization. We also may note that in many cases the p occurring in Eq. (14) is rather small, say $p \leqslant 5$, and that the quantities t_{lj} (despite their complicated appearance) are quite easy to compute. Furthermore, because of the behavior of the Bessel functions of large order at fixed argument, the convergence of the series in Eq. (14) is rapid when f and v are not very large. The author is preparing a companion paper[1] in which, among other things, this approach of evaluating diffraction integrals is compared with other approaches from a numerical point of view. In that paper it is shown that, as a rule of thumb, one needs some 25 terms in the series over l to get absolute accuracy of order 10^{-6} for all $v \leqslant 20$, $0 \leqslant m \leqslant n \leqslant 14$, and $|f| \leqslant 2\pi$.

The formulas (14)–(16) continue to hold for integers n, $m \geqslant 0$ with $n - m$ even and < 0, except that the summations over j in Eq. (14) should be extended to $+\infty$ and the binomials in Eq. (16) involving p should be read as

$$\binom{p}{k} = \frac{p(p - 1)\cdots(p - k + 1)}{k!}, \tag{17}$$

when k is a nonnegative integer. The particularly convenient fact that we may consider finite $p \geqslant 0$ in Eqs. (14) and (16) is a consequence of the fact that we have developed Φ in terms of Zernike polynomials.

As already said, we have for U the first-order approximation

$$U(x,\ y) \approx 2\int_0^1 \rho \exp(if\rho^2) J_0(2\pi\rho r)\mathrm{d}\rho + 2i\sum_{n,m} \alpha_{nm} i^m \cos m\phi \times \int_0^1 \rho \exp(if\rho^2) R_n^m(\rho) J_m(2\pi\rho r)\mathrm{d}\rho. \tag{18}$$

In Appendix A.2 we shall show that, for integers n, $m \geqslant 0$ with $n - m$ even and $\geqslant 0$,

$$\int_0^1 \rho R_n^m(\rho)\exp(if\rho^2) J_m(2\pi\rho r)\mathrm{d}\rho = \exp(if)\sum_{l=1}^{\infty} (-2if)^{l-1}\sum_{j=0}^{p} v_{lj} \frac{J_{m+l+2j}(v)}{l v^l}, \tag{19}$$

where

$$v_{lj} = (-1)^p (m + l + 2j)\binom{m + j + l - 1}{l - 1} \times \binom{j + l - 1}{l - 1}\binom{l - 1}{p - j} \bigg/ \binom{q + l + j}{l}, \quad j = 0, 1, \ldots,\quad l = 1, 2, \ldots, \tag{20}$$

and v, p, q are as in Eq. (15). In the particular case in which $f = 0$, we have that only the term with $l = 1$ is present on the right-hand side of Eq. (19). For $l = 1$ the series over j on the right-hand side of Eq. (19) reduces to only the term with $j = p$ because of the occurrence of the binomial $\binom{l-1}{p-j}$ on the right-hand side of Eq. (20). It thus appears that the basic integral (7) in the Nijboer–Zernike theory occurs as a special case of Eq. (19) with $f = 0$. We may also note that the first integral on the right-hand side of relation (18) can be dealt with by using Eq. (19) and taking $n = m = 0$; this is so, since $R_0^0 \equiv 1$.

We thus see that these results can be considered a completion of the Nijboer–Zernike theory in the sense that, in principle, the diffraction integral can be computed effectively for all aberrations of modest to relatively large size on the basis of the formulas (11)–(16). Furthermore, when the aberration is so small that first-order considerations in Eq. (4) suffice, a considerable simplification of the Nijboer–Zernike theory is obtained from formulas (19) and (20). These formulas are, in addition, directly applicable when one has developed the full aberration $A\exp(i\Phi)$, rather than Φ, in a Zernike series.

We finally note that in the results just presented it is in no way required that f be real. Hence these results extend to the case in which we have Gaussian beam profiles as well as defocusing in the exit pupil.

3. EXPRESSING THE DIFFRACTION INTEGRAL IN TERMS OF T_{nm}

In this section we show how the kth-order term (11) in the expansion (4), when multiplied by $\rho\exp(if\rho^2)$ and integrated over $\rho \in [0,1]$ as in Eq. (3), can be expressed in terms of the T_{nm} in Eq. (13).

We write the following for Φ as in Eq. (2):

$$\int_0^{2\pi} \Phi^k(\rho,\ \theta + \phi)\exp(2\pi i\rho r\cos\theta)\mathrm{d}\theta = \sum_{n_1, m_1, \ldots, n_k, m_k} \alpha_{n_1 m_1}\cdots\alpha_{n_k m_k} R_{n_1}^{m_1}(\rho)\cdots R_{n_k}^{m_k}(\rho) \times \int_0^{2\pi} [\cos m_1(\theta + \phi)]\cdots[\cos m_k(\theta + \phi)] \times \exp(2\pi i\rho r\cos\theta)\mathrm{d}\theta. \tag{21}$$

To deal with the integral on the right-hand side of Eq. (21), we repeatedly apply the formula

$$\cos x\cos y = \tfrac{1}{2}\cos(x + y) + \tfrac{1}{2}\cos(x - y) = \tfrac{1}{4}\cos(x + y) + \tfrac{1}{4}\cos(x - y) + \tfrac{1}{4}\cos(-x + y) + \tfrac{1}{4}\cos(-x - y), \tag{22}$$

so that

$$\cos m_1 x\cdots\cos m_k x = \frac{1}{2^k}\sum_{\varepsilon_1, \ldots, \varepsilon_k = \pm 1} \cos\left(\left|\sum_{l=1}^{k} \varepsilon_l m_l\right| x\right). \tag{23}$$

The right-hand series in Eq. (23) has terms $\cos \nu x$, where ν is a nonnegative integer $\leqslant m_1 + \cdots + m_k$ with the same parity as that of $m_1 + \cdots + m_k$. Furthermore, there holds (see Ref. 9, Eqs. 9.1.44 and 9.1.45 on p. 361) the relationship

$$\int_0^{2\pi} [\cos \nu(\theta + \phi)]\exp(2\pi i\rho r\cos\theta)\mathrm{d}\theta = 2\pi i^{\nu}\cos \nu\phi\, J_{\nu}(2\pi\rho r). \tag{24}$$

Next we have from the explicit form

$$R_n^m(\rho) = \sum_{s=0}^{p} \frac{(-1)^s (n - s)!\rho^{n-2s}}{s!(q - s)!(p - s)!}, \quad p = \frac{1}{2}(n - m),\ q = \frac{1}{2}(n + m), \tag{25}$$

where $n - m$ is even and $\geqslant 0$, that

$$R_{n_1}^{m_1}(\rho)\cdots R_{n_k}^{m_k}(\rho) \tag{26}$$

is a polynomial in ρ of degree $n_1 + \cdots + n_k$ with nonzero coefficients of the powers ρ^t for only integers t satisfying

$$m_1 + \cdots + m_k \leqslant t \leqslant n_1 + \cdots + n_k \tag{27}$$

and having the same parity as that of either side of relation (27).

We thus conclude that any term on the right-hand-side series in Eq. (21) is a finite series of terms of the form

$$\beta_{nm}\rho^n J_m(2\pi\rho r)\cos m\phi \tag{28}$$

with $n, m \geqslant 0$ and $n - m$ even and $\geqslant 0$ and where the β_{nm} can be explicitly expressed in terms of the α_{nm} by using Eqs. (23)–(25). Hence we can express the kth-order term in the expansions (3) and (4) of $U(x, y)$,

$$\int_0^1 \rho \exp(if\rho^2)\left[\int_0^{2\pi} \Phi^k(\rho, \theta + \phi)\exp(2\pi i\rho r\cos\theta)\mathrm{d}\theta\right]\mathrm{d}\rho, \tag{29}$$

explicitly in terms of the T_{nm} of Eq. (13), as claimed.

4. EXPLICIT RESULT FOR COMA AND ASTIGMATISM

We shall now develop explicit representations of $U(x, y)$ in terms of the T_{nm} for the cases in which

$$\Phi(\rho, \theta) = \alpha\rho^3\cos\theta, \qquad \Phi(\rho, \theta) = \gamma\rho^2\cos 2\theta \tag{30}$$

(coma and astigmatism, respectively); for these special cases there are some shortcuts in the program outlined in Section 3.

For the first case in Eq. (30), we have

$$U(x, y) = \frac{1}{\pi}\int_0^1 \rho\exp(if\rho^2)\left[\int_0^{2\pi}\exp[i\alpha\rho^3\cos(\theta + \phi)] \times \exp(2\pi i\rho r\cos\theta)\mathrm{d}\theta\right]\mathrm{d}\rho. \tag{31}$$

Using Ref. 9, Eqs. 9.1.44 and 9.1.45 on p. 361, for either exponential in the inner integral, and carrying out the integration over θ, we have

$$\begin{aligned} U(x, y) &= 2\int_0^1 \rho\exp(if\rho^2)J_0(\alpha\rho^3)J_0(2\pi\rho r)\mathrm{d}\rho \\ &\quad + 4\sum_{j=1}^{\infty}(-1)^j\int_0^1 \rho\exp(if\rho^2)J_j(\alpha\rho^3) \\ &\quad \times J_j(2\pi\rho r)\mathrm{d}\rho\cos j\phi. \end{aligned} \tag{32}$$

Next, by using the power-series expansion of the jth Bessel function with argument $z = \alpha\rho^3$ (see Ref. 9, Eqs. 9.1.10 on p. 360), we get

$$\int_0^1 \rho\exp(if\rho^2)J_j(\alpha\rho^3)J_j(2\pi\rho r)\mathrm{d}\rho = \left(\frac{1}{2}\alpha\right)^j\sum_{k=0}^{\infty}\frac{(-\frac{1}{4}\alpha^2)^k}{k!(k + j)!}T_{3j+6k,j}. \tag{33}$$

Then, on collecting terms in the expansion of U with equal powers of α, we finally obtain

$$U(x, y) = \sum_{l=0}^{\infty} C_l\alpha^l, \tag{34}$$

where, for $m = 0, 1, \ldots,$

$$C_{2m} = 2\left(-\frac{1}{4}\right)^m\sum_{k=0}^{m}\varepsilon_k\frac{(-1)^k T_{6m,2k}}{(m - k)!(m + k)!}\cos 2k\phi, \tag{35}$$

$$C_{2m+1} = -2\left(-\frac{1}{4}\right)^m\sum_{k=0}^{m}\frac{(-1)^k T_{6m+3,2k+1}}{(m - k)!(m + k + 1)!} \times \cos(2k + 1)\phi. \tag{36}$$

In Eq. (35) we have used Neumann's symbol ε, so that $\varepsilon_0 = 1$, $\varepsilon_n = 2$, $n \neq 0$.

We next consider the second choice for Φ in Eqs. (30). Then we have

$$U(x, y) = \frac{1}{\pi}\int_0^1 \rho\exp(if\rho^2) \times \left\{\int_0^{2\pi}\exp[i\gamma\rho^2\cos 2(\theta + \phi)]\exp(2\pi i\rho r\cos\theta)\mathrm{d}\theta\right\}\mathrm{d}\rho. \tag{37}$$

Then, as above, we have

$$\begin{aligned} \int_0^{2\pi}&\exp[i\gamma\rho^2\cos 2(\theta + \phi)]\exp(2\pi i\rho r\cos\theta)\mathrm{d}\theta \\ &= 2\pi J_0(\gamma\rho^2)J_0(2\pi\rho r) \\ &\quad + 4\pi\sum_{j=1}^{\infty}(-i)^j J_j(\gamma\rho^2)J_{2j}(2\pi\rho r)\cos 2j\phi. \end{aligned} \tag{38}$$

The result is

$$U(x, y) = \sum_{l=0}^{\infty} D_l\gamma^l,$$

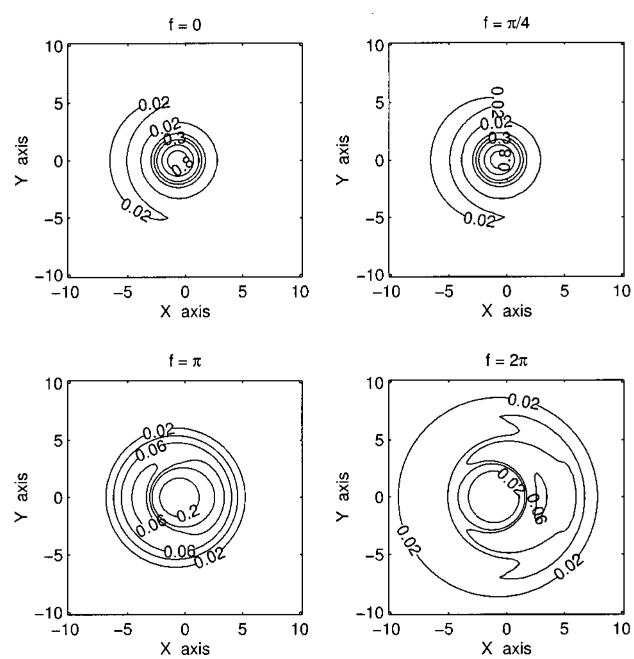

Fig. 2. Contour plot of the modulus of the point-spread function $U(x, y)$ with aberration $\Phi(\rho, \theta) = \alpha\rho^3\cos\theta$ (coma), where $\alpha = 1$ and $f = 0, \pi/4, \pi, 2\pi$.

where, for $m = 0, 1, \ldots,$

$$D_{2m} = 2\left(-\frac{1}{4}\right)^m \sum_{k=0}^{m} \varepsilon_k \frac{T_{4m,4k}}{(m-k)!(m+k)!} \cos 4k\phi, \tag{39}$$

$$D_{2m+1} = -2i\left(-\frac{1}{4}\right)^m \sum_{k=0}^{m} \frac{T_{4m+2,4k+2}}{(m-k)!(m+k+1)!} \times \cos 2(2k+1)\phi, \tag{40}$$

where we have again used Neumann's symbol in Eq. (39).

In Fig. 2 $|U(x, y)|$ is shown for the case of coma [see Eqs. (30)] with $\alpha = 1$ and $f = 0, \pi/4, \pi, 2\pi$, calculated according to Eqs. (34)–(36) including the terms with $l = 0, 1, 2, 3$.

5. CONCLUSION

An extension has been given of the Nijboer–Zernike theory of the diffraction integral over the exit pupil, producing the point-spread function in the image plane. The new approach is valid for large aberration and defocus values and also permits a nonuniform amplitude distribution over the wavefront in the exit pupil. As such, the new approach provides an interesting tool for the analysis of image plane intensity distributions encountered in fields such as microscopy, lithography, or astronomical observation. In a companion paper[1] physical interpretations and applications in a lithographic context of this approach are presented, a convergence analysis is given, and a comparison is made with results obtained by using a numerical package.

APPENDIX A: DERIVATION OF FORMULAS (14) AND (19)

1. Derivation of Formula (14)

It will now be shown that, for integers $n, m \geqslant 0$ with $n - m$ even and $\geqslant 0$,

$$T_{nm} = \exp(if)\sum_{l=1}^{\infty}(-2if)^{l-1}\sum_{j=0}^{p} t_{lj}\frac{J_{m+l+2j}(v)}{v^l}, \tag{A1}$$

where $v = 2\pi r$, $p = \frac{1}{2}(n-m)$, and $q = \frac{1}{2}(n+m)$, as in Section 2, and

$$t_{lj} = (-1)^j \frac{m+l+2j}{q+1}\binom{p}{j} \times \binom{m+j+l-1}{l-1} \Big/ \binom{q+l+j}{q+1}, \quad j = 0, 1, \ldots, \quad l = 1, 2, \ldots. \tag{A2}$$

We start by noting that

$$T_{nm} = \left(\frac{1}{2\pi r}\right)^{n+2}\int_0^v t^{n+1}\exp(i\beta t^2)J_m(t)\mathrm{d}t =: \left(\frac{1}{2\pi r}\right)^{n+2} H_{nm}(v), \tag{A3}$$

where $v = 2\pi r$ and $\beta = f/(2\pi r)^2$.

We consider, more generally, real values of n and m such that $n + m + 1 > 0$, and we define a sequence of functions $F_{nm}^{(l)}(v)$ by

$$F_{nm}^{(0)}(v) = v^n J_m(v), \tag{A4}$$

$$F_{nm}^{(l)}(v) = \int_0^v t F_{nm}^{(l-1)}(t)\mathrm{d}t, \qquad l = 1, 2, \ldots. \tag{A5}$$

Then, with H_{nm} given as in Eq. (A3), it follows by partial integrations that

$$H_{nm}(v) = \exp(i\beta v^2)\sum_{l=1}^{\infty}(-2i\beta)^{l-1}F_{nm}^{(l)}(v), \tag{A6}$$

where $\beta = f/(2\pi r)^2$ as above.

It is claimed that, for $l = 1, 2, \ldots,$

$$F_{nm}^{(l)}(v) = v^{n+l}\sum_{j=0}^{\infty}\frac{m+2j+l}{q+1} \times (-1)^j \frac{\binom{p}{j}\binom{m+j+l-1}{l-1}}{\binom{q+j+l}{q+1}} J_{m+2j+l}(v). \tag{A7}$$

Here we have set $p = \frac{1}{2}(n-m)$, $q = \frac{1}{2}(n+m)$ as usual, and the binomials that occur on the right-hand side of Eq. (A7) are given as

$$\binom{\alpha}{k} = \frac{\alpha(\alpha-1)\cdots(\alpha-k+1)}{k!} = \frac{\Gamma(\alpha+1)}{\Gamma(\alpha-k+1)\Gamma(k+1)} \tag{A8}$$

for $\alpha \in \mathbb{R}$ and $k = 0, 1, \ldots$ with the usual precautions for the Γ-functions on the far right-hand side of Eq. (A8) when α is a negative integer. From Eq. (A7) result (A1) follows from Eqs. (A3) and (A6) by using that $v = 2\pi r$ and $\beta = f/(2\pi r)^2$.

To show Eq. (A7), we start from Ref. 9, Eq. 11.1.1 on p. 480, with $\mu = n + 1$, $\nu = m$, which takes the form [$p = \frac{1}{2}(n-m)$, $q = \frac{1}{2}(n+m)$]

$$\int_0^v t^{n+1}J_m(t)\mathrm{d}t = v^{n+1}\sum_{k=0}^{\infty}\frac{(m+2k+1)\Gamma(-p+k)\Gamma(q+1)}{\Gamma(-p)\Gamma(q+k+2)} \times J_{m+2k+1}(v). \tag{A9}$$

We can see that the right-hand side of Eq. (A9) agrees with the right-hand side of Eq. (A7) for the case $l = 1$ by observing that

$$(-1)^j\binom{p}{j} = \frac{-p(-p+1)\cdots(-p+j-1)}{j!} = \frac{\Gamma(-p+j)}{\Gamma(-p)j!}. \tag{A10}$$

We furthermore note that, for arbitrary $l = 1, 2, \ldots,$

$$\frac{m+2j+l}{q+1}(-1)^j\frac{\binom{p}{j}\binom{m+j+l-1}{l-1}}{\binom{q+j+l}{q+1}}$$
$$=\frac{\Gamma(-p+j)}{\Gamma(-p)}\frac{\Gamma(j+l)}{\Gamma(j+1)}\frac{\Gamma(q+1)}{\Gamma(q+j+l+1)}\times\binom{m+j+l-1}{l-1}(m+2j+l). \quad \text{(A11)}$$

Now assume that Eq. (A7) holds for a certain $l = 1, 2, \ldots$. Then

$$F_{nm}^{(l+1)}(v)=\int_0^v tF_{nm}^{(l)}(t)\mathrm{d}t$$
$$=\sum_{j=0}^{\infty}\frac{\Gamma(-p+j)}{\Gamma(-p)}\frac{\Gamma(j+l)}{\Gamma(j+1)}\frac{\Gamma(q+1)}{\Gamma(q+j+l+1)}\times\binom{m+j+l-1}{l-1}(m+2j+l)\times\int_0^v t^{m+l+1}J_{m+2j+l}(t)\mathrm{d}t. \quad \text{(A12)}$$

Next, Eq. (A9) (with $n' = n + l$, $m' = m + 2j + l$ instead of n, m, so that we have $p' = p - j$, $q' = q + j + l$ instead of p, q) yields

$$\int_0^v t^{n+l+1}J_{m+2j+l}(t)\mathrm{d}t = v^{n+l+1}\sum_{k=0}^{\infty}\frac{[m+2(j+k)+l+1]\Gamma(-p+j+k)\Gamma(q+j+l+1)}{\Gamma(-p+j)\Gamma(q+l+j+k+2)}J_{m+2(j+k)+l+1}(v). \quad \text{(A13)}$$

Inserting Eq. (A13) into Eq. (A12) and simplifying, we obtain

$$F_{nm}^{(l+1)}(v)=v^{n+l+1}\sum_{j,k=0}^{\infty}\frac{\Gamma(-p+j+k)}{\Gamma(-p)}\frac{\Gamma(j+l)}{\Gamma(j+1)}\times\frac{\Gamma(q+1)}{\Gamma(q+j+k+l+2)}\binom{m+j+l-1}{l-1}\times(m+2j+l)\times[m+2(j+k)+l+1]J_{m+2(j+k)+l+1}(v). \quad \text{(A14)}$$

Collecting terms in the double series on the right-hand side of Eq. (A14) with the same value of $j + k$, we arrive at

$$F_{nm}^{(l+1)}(v)=v^{n+l+1}\sum_{s=0}^{\infty}\frac{\Gamma(-p+s)}{\Gamma(-p)}\frac{\Gamma(q+1)}{\Gamma(q+s+l+2)}\times(m+2s+l+1)\times J_{m+2s+l+1}(v)\sum_{j=0}^{s}\frac{\Gamma(j+l)}{\Gamma(j+1)}\times\binom{m+j+l-1}{l-1}(m+2j+l). \quad \text{(A15)}$$

Next, we consider the series on the last two lines of Eq. (A15). It holds that

$$\sum_{j=0}^{s}\frac{\Gamma(j+l)}{\Gamma(j+1)}\binom{m+j+l-1}{l-1}(m+2j+l)=\frac{\Gamma(s+l+1)}{\Gamma(s+1)}\binom{m+s+l}{l}. \quad \text{(A16)}$$

This identity is easily established by induction on $s = 0, 1, \ldots,$ where we also note that Eq. (A16) is equivalent to

$$\sum_{j=0}^{s}(j+l-1)\cdots(j+1)(m+j+l-1)\cdots(m+j+1)(m+2j+l)=\frac{1}{l}(s+l)\cdots(s+1)(m+s+l)\cdots(m+s+1), \quad \text{(A17)}$$

which is conveniently used in the induction step.

Going back to Eq. (A15), we then see that we have shown that

$$F_{nm}^{(l+1)}(v)=v^{n+l+1}\sum_{s=0}^{\infty}\frac{\Gamma(-p+s)}{\Gamma(-p)}\frac{\Gamma(q+1)}{\Gamma(q+s+l+2)}\times\frac{\Gamma(s+l+1)}{\Gamma(s+1)}\binom{m+s+l}{l}\times(m+2s+l+1)J_{m+2s+l+1}(v), \quad \text{(A18)}$$

and then comparing Eq. (A18) with Eq. (A11), we see that it is the desired formula (A7) for $F_{nm}^{(l+1)}(v)$.

2. Derivation of Formula (19)

It will now be shown that, for integers $n, m \geq 0$ with $n - m$ even and ≥ 0,

$$\int_0^1 \rho R_n^m(\rho)\exp(if\rho^2)J_m(2\pi\rho r)\mathrm{d}\rho = \exp(if)\sum_{l=1}^{\infty}(-2if)^{l-1}\sum_{j=0}^{p}v_{lj}\frac{J_{m+l+2j}(v)}{lv^l}, \quad \text{(A19)}$$

where $v = 2\pi r$, $p = \frac{1}{2}(n-m)$, and $q = \frac{1}{2}(n+m)$ as above, and

$$v_{lj}=(-1)^p(m+l+2j)\binom{m+j+l-1}{l-1}\times\binom{j+l-1}{l-1}\binom{l-1}{p-j}\Big/\binom{q+l+j}{l},$$
$$j=0,1,\ldots,\quad l=1,2,\ldots. \quad \text{(A20)}$$

To arrive at Eqs. (A19) and (A20), we have from Eq. (25) that

$$\int_0^1 \rho R_n^m(\rho)\exp(if\rho^2)J_m(2\pi\rho r)\mathrm{d}\rho = \sum_{s=0}^{p} \frac{(-1)^s(n-s)!}{s!(q-s)!(p-s)!} T_{n-2s,m}, \quad \text{(A21)}$$

with the $T_{n-2s,m}$ as in Eq. (13). We next insert formula (14) into the right-hand side of Eq. (A21) (with $n' = n - 2s$ instead of n, so that we have $p' = p - s$, $q' = q - s$ instead of p, q), and we get

$$\int_0^1 \rho R_n^m(\rho)\exp(if\rho^2)J_m(2\pi\rho r)\mathrm{d}\rho = \exp(if)\sum_{l=1}^{\infty} (-2if)^{l-1} \times \left[\sum_{s=0}^{p}\sum_{j=0}^{p-s} \frac{(-1)^s(n-s)!t_{lj}(n-2s, m)}{s!(q-s)!(p-s)!} \frac{J_{m+l+2j}(v)}{v^l}\right], \quad \text{(A22)}$$

where

$$t_{lj}(n-2s, m) = (-1)^j \frac{m+l+2j}{q-s+1}\binom{p-s}{j} \times \binom{m+j+l-1}{l-1}\Big/\binom{j+q-s+l}{q-s+1} = (-1)^j(m+l+2j)\binom{m+j+l-1}{l-1} \times \frac{(p-s)!(q-s)!}{(p-j-s)!(j+q-s+l)!} \times \frac{(j+l-1)!}{j!}. \quad \text{(A23)}$$

On simplifying, we then find that the quantity in [ ] on the right-hand side of Eq. (A22) is given by

$$[\;] = \sum_{s=0}^{p}\sum_{j=0}^{p-s} (-1)^j(m+l+2j) \times \binom{m+j+l-1}{l-1}\frac{(j+l-1)!}{j!} \times \frac{J_{m+l+2j}(v)}{v^l} \frac{(-1)^s(n-s)!}{s!(p-j-s)!(j+q-s+l)}. \quad \text{(A24)}$$

We next write the double series over s and j on the right-hand side of Eq. (A24) as $\Sigma_{j=0}^{p}\Sigma_{s=0}^{p-j}$, and we obtain

$$[\;] = \sum_{j=0}^{p} (-1)^j(m+l+2j)\binom{m+j+l-1}{l-1} \times \frac{(j+l-1)!}{j!}\frac{J_{m+l+2j}(v)}{v^l} \times \sum_{s=0}^{p-j} \frac{(-1)^s(n-s)!}{s!(p-j-s)!(j+q-s+l)!}. \quad \text{(A25)}$$

Now noting that $n = p + q$, we can write the series over s on the last line of Eq. (A25) as $S(p-j, n, p-j-l)$, where, for $k = 0, 1, \ldots, p$ and integer $i < k$, we have written

$$S(k, n, i) = \sum_{s=0}^{k} \frac{(-1)^s(n-s)!}{s!(k-s)!(n-s-i)!} = \frac{1}{k!}\sum_{s=0}^{k} (-1)^s\binom{k}{s}\frac{(n-s)!}{(n-s-i)!}. \quad \text{(A26)}$$

The last equality in Eq. (A26) follows immediately from the definition of the binomials.

We shall see now that, for $k = 0, 1, \ldots, p$ and integer $i < k$, we have

$$S(k, n, i) = \begin{cases} 0, & 0 \leqslant i < k \\ (-1)^k\binom{k-i-1}{k}\dfrac{(n-k)!}{(n-i)!}, & i < 0 \end{cases}. \quad \text{(A27)}$$

To that end, we observe that from the second form of $S(k, n, i)$ on the right-hand side of Eq. (A26), we have

$$S(k, n, i) = \frac{1}{k!}(\Delta^k f)(n), \quad \text{(A28)}$$

where

$$f(x) = \frac{\Gamma(x+1)}{\Gamma(x-i+1)}, \qquad x > \min(-1, i-1), \quad \text{(A29)}$$

and where Δ^k denotes the kth power of the backward-difference operator Δ, defined as

$$(\Delta g)(x) = g(x) - g(x-1) \quad \text{(A30)}$$

for functions g with x, $x-1$ in the domain of g.

In the case in which i is an integer with $0 \leqslant i < k$, we have that f in Eq. (A29) is a polynomial in x of degree $i < k$, whence Eq. (A28) vanishes. In the case in which $t := -i$ is a positive integer, we have

$$f(x) = \frac{1}{(x+1)\cdots(x+t)} = \frac{(-1)^{t-1}}{(t-1)!}(\Delta^{t-1}g)(x), \quad \text{(A31)}$$

where

$$g(x) = \frac{1}{x+t}. \quad \text{(A32)}$$

In Eq. (A31) the first identity follows at once from Eq. (A29), while the second identity follows easily by induction with respect to $t = 1, 2, \ldots$. It thus follows that when $i = -t$ is a negative integer,

$$S(k, n, i) = \frac{1}{k!}\left(\Delta^k\left[\frac{(-1)^{t-1}}{(t-1)!}\Delta^{t-1}g\right]\right)(n) = \frac{(-1)^{t-1}}{k!(t-1)!}(\Delta^{k+t-1}g)(n). \quad \text{(A33)}$$

Next, we use the fact that Δ commutes with the shift operator $f(x) \to f(x-k)$, together with Eq. (A31), which holds for arbitrary positive integers t, and we get

$$S(k, n, i) = \frac{(k+t-1)!}{k!(t-1)!} \frac{(-1)^k}{(n-k+1)\cdots(n+t)}$$
$$= (-1)^k \binom{k+t-1}{k} \frac{(n-k)!}{(n+t)!}, \qquad \text{(A34)}$$

and this is the second case, $i = -t < 0$, in Eq. (A27).

In Eq. (A25) we need $S(k, n, i)$ with $k = p - j$, $i = p - j - l$. Thus Eq. (A27) yields

$$S(p-j, n, p-j-l) = \begin{cases} 0, & p-j \geqslant l \\ (-1)^{p-j} \binom{l-1}{p-j} \dfrac{(q+j)!}{(q+l+j)!}, & p-j < l \end{cases}. \qquad \text{(A35)}$$

Therefore the quantity in Eq. (A22) between [ ] equals

$$[\ ] = \sum_{j=0}^{p} v_{lj} \frac{J_{m+l+2j}(v)}{l v^l}, \qquad \text{(A36)}$$

where

$$v_{lj} = (-1)^p (m+l+2j) \binom{m+j+l-1}{l-1} \frac{(j+l-1)!}{j!} \times \binom{l-1}{p-j} \frac{(q+j)!}{(q+l+j)!} l$$
$$= (-1)^p (m+l+2j) \binom{m+j+l-1}{l-1} \binom{j+l-1}{l-1} \times \binom{l-1}{p-j} \Big/ \binom{q+l+j}{l}, \qquad \text{(A37)}$$

and this is what we wanted to prove.

ACKNOWLEDGMENTS

The author thanks his colleagues J. Braat, W. Coene, and P. Dirksen for useful discussions and advice during the preparation of this paper. The author also thanks the referees for their useful comments.

A. J. E. M. Janssen can be contacted at the address on the title page or by phone, 31-40-274-2402; fax, 31-40-274-4675; and e-mail, A.J.E.M.Janssen@philips.com.

REFERENCES

1. J. J. M. Braat, P. Dirksen, and A. J. E. M. Janssen, “Assessment of an extended Nijboer–Zernike approach for the computation of optical point-spread functions,” J. Opt. Soc. Am. A **19**, 858–870 (2002).
2. M. Born and E. Wolf, *Principles of Optics* (Pergamon, Oxford, UK, 1985).
3. B. R. A. Nijboer, “The diffraction theory of aberrations,” Ph.D. dissertation (University of Groningen, Groningen, The Netherlands, 1942).
4. R. J. Noll, “Zernike polynomials and atmospheric turbulence,” J. Opt. Soc. Am. **66**, 207–211 (1976).
5. J. Y. Wang and D. E. Silva, “Wave-front interpretation with Zernike polynomials,” Appl. Opt. **19**, 1510–1518 (1980).
6. S. C. Biswas and J.-E. Villeneuve, “Diffraction of a laser beam by a circular aperture under the combined effect of three primary aberrations,” Appl. Opt. **25**, 2221–2232 (1986).
7. M. A. A. Neil, M. J. Booth, and T. Wilson, “New modal wavefront sensor: theoretical analysis,” J. Opt. Soc. Am. A **17**, 1098–1107 (2000).
8. K. Nienhuis, “On the influence of diffraction on image formation in the presence of aberrations,” Ph.D. thesis (University of Groningen, Groningen, The Netherlands, 1948).
9. M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions* (Dover, New York, 1970).