An 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

Short title: Diffraction integrals

Abstract.
We present new Bessel series representations for the calculation of the diffraction integral, yielding the point spread function of the optical system, as occurs in the Nijboer-Zernike theory of aberrations. In this analysis we can allow an arbitrary aberration and a defocus part. We present our representations in all 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 wavefront 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, 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.

OCIS codes: 050.1960, 070.2580, 000.3860, 000.4430.

*E-mail: A.J.E.M.Janssen@philips.com

1 Introduction

The Nijboer-Zernike theory of diffraction integrals was intended to produce an analytical result which, with the computational means of that time, leads to a good approximation for the intensity distribution in or close to the focal plane. As indicated already by Nijboer himself, the permitted wavefront aberration should not, in terms of the phase of the exponential factor, exceed a few radians. The amplitude distribution over the wavefront is limited to a uniform one. These two conditions make 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, we propose an extended analytical approach which does not suffer from the small aberration and the amplitude uniformity restriction. 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 be treated also in our new approach along the same lines as 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 Sec. 4 for the case of coma with defocus values up to 2π). In a forthcoming paper[1] we will investigate these issues 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 to the wavelength of the light used. For a summary of the diffraction theory of aberrations and the relevant aspects of the Nijboer-Zernike theory we refer to Ref. 2, Ch. 9, in particular Secs. 9.1–4 and Appendix VII that 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, Ch. 9, a point source of monochromatic light in the object plane of a centred 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

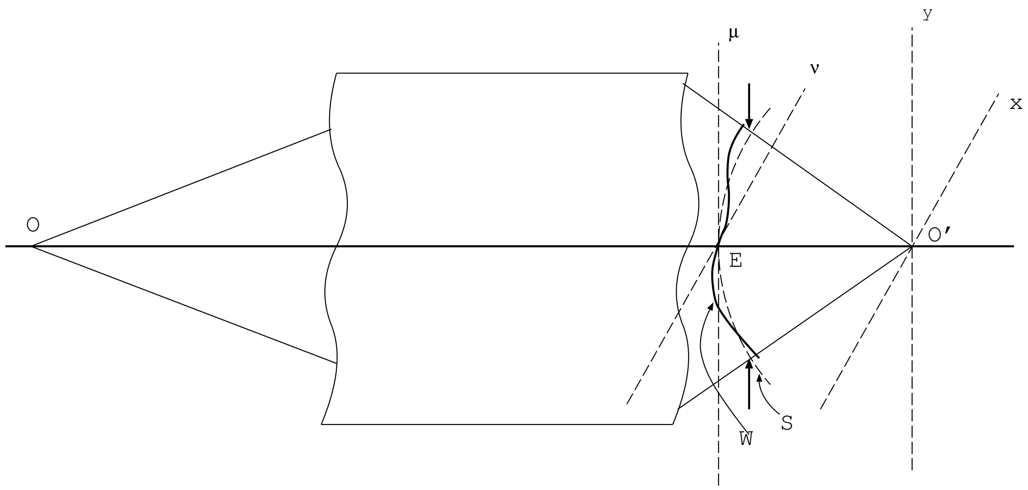

Figure 1: A point source in O emits a spherical wave towards the schematically represented optical system. In the image space, an aberrated wavefront W leaves the exit pupil (center in E) and comes to a focus close to the image plane through O'. The spherical reference wavefront is denoted by S and the wave front aberration is given by the perpendicular distance between S and W. The phase function Φ is derived from W via $\Phi = 2\pi W/\lambda$ with λ the wavelength of the monochromatic radiation. The normalized cartesian pupil coordinates are denoted by (ν, μ), the coordinates (x, y) in the image plane have been normalized with respect to the diffraction unit, λ/NA, with NA 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.

diffraction integrals we are concerned with take the form

$$U(x,y) = \frac{1}{\pi} \iint_{\nu^2+\mu^2 \le 1} \exp(i(\nu^2+\mu^2)f + i\Phi(\nu,\mu)) \exp(2\pi i\nu x + 2\pi i\mu y)\, d\nu d\mu\,. \quad (1)$$

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 (1) is complex-valued, so that non-uniform amplitude distributions can be accommodated as well.

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

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

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

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 (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 at the right-hand side of (2) can be taken over all integers $n,m \geq 0$ with $n-m$ even and ≥ 0. Also, $R_n^m(\rho)$ is a polynomial in ρ of degree n containing the powers $\rho^n, \rho^{n-2}, ..., \rho^m$ only. Then using polar coordinates for the integral at the right-hand side of (1), we have

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

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

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

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

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

$$= \sum_{n,m}\alpha_{nm}\,R_n^m(\rho)\int_0^{2\pi}\cos m(\theta+\phi)\exp(2\pi i\rho r\cos\theta)\,d\theta \ =$$

$$= \; 2\pi \sum_{n,m} \alpha_{nm} \, i^m \, R_n^m(\rho) \, J_m(2\pi\rho r) \cos m\phi \; . \tag{5}$$

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

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

with integer $n, m \geq 0$ such that $n - m$ is even and ≥ 0. It is a key result in the Nijboer-Zernike theory that

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

Hence the integrals in (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

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

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

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

Then (7) can be used for each of the terms at the right-hand side of (9). This then yields a series representation of the integral in (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 at the right-hand side of (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 in the right-hand side of (4) terms of order $k \geq 2$ have to be included, for then one needs to consider products of $k+1$ Zernike polynomials that 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 has obtained a systematic way for the determination of coefficients in series of the type (9) for products of Zernike polynomials. This is based on the formula, valid for $m, k = 0, 1, \dots$,

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

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

In this paper we remedy this situation by giving explicit Bessel series representations for the integrals in (6), as well as for integrals as in (6) with $R_n^m(\rho)$ replaced with ρ^n where $n-m$ is even and ≥ 0. The integrals involving $R_n^m(\rho)$ are directly useful when the (possibly non-uniform) 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 (2) by expanding any Φ^k at the right-hand side of (4) as a series of terms of the form $\rho^n \cos m(\theta+\phi)$ with integer $n, m \geq 0$ such that $n-m \geq 0$ and even. Full details are given for the cases that the aberration represents coma and astigmatism, respectively.

2 Summary of results

In Sec. 3 we shall explicitly express any of the terms

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

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

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

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

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

For integers $n, m \geq 0$ such that $n-m$ is even and ≥ 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

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

and

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

Thus, combining the results of Appendix A.1, the integral expression (3) for U can be expressed explicitly, in terms of powers of f, the α_{nm} in (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 (14) is rather small, say $p \leq 5$, and that the quantities t_{lj} (despite their complicated appearance) are quite easy to compute. Furthermore, due to the behaviour of the Bessel functions of large order at fixed argument, the convergence of the series in (14) is rapid when f van v are not very large. The author is preparing a companion paper[1] in which, among other things, our approach of evaluating diffraction integrals is compared to 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 $v \leq 20$, $0 \leq m \leq n \leq 14$ and $|f| \leq 2\pi$.

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

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

when k is a non-negative integer. The particular convenient fact that we may consider finite $p \geq 0$ in (14), (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

$$U(x,y) \approx 2 \int\limits_0^1 \rho \exp(if\rho^2)\, J_0(2\pi\rho r)\, d\rho \ +$$

$$+\ 2i \sum_{n,m} \alpha_{nm}\, i^m \cos m\phi \int\limits_0^1 \rho \exp(if\rho^2)\, R_n^m(\rho)\, J_m(2\pi\rho r)\, d\rho \ . \tag{18}$$

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

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

where

$$v_{lj} = (-1)^p (m+l+2j) \binom{m+j+l-1}{l-1} \binom{j+l-1}{l-1} \binom{l-1}{p-j} \Big/ \binom{q+l+j}{l}\,,$$

$$j = 0, 1, \ldots\,, \quad l = 1, 2, \ldots\,, \tag{20}$$

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

We thus see that our results can be considered as 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 basis of the formulas (11–16). Furthermore, when the aberration is so small that first order considerations in (4) suffice, a considerable simplification of the Nijboer-Zernike theory is obtained from formulas (19–20). These formulas (19-20) 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 by no way required that f is real. Hence our results extend to the case that 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 k^{th} order term (11) in the expansion (4), when multiplied by $\rho \exp(if\rho^2)$ and integrated over $\rho \in [0,1]$ as in (3), can be expressed in terms of the T_{nm} in (13).

We write for Φ as in (2)

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

$$= \sum_{n_1,m_1,\ldots,n_k,m_k} \alpha_{n_1 m_1} \ldots \alpha_{n_k m_k}\, R_{n_1}^{m_1}(\rho) \ldots R_{n_k}^{m_k}(\rho) \;\cdot$$

$$\cdot \int_0^{2\pi} \cos m_1(\theta+\phi) \ldots \cos m_k(\theta+\phi) \exp(2\pi i\rho r\cos\theta)\,d\theta \;. \quad (21)$$

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

$$\begin{aligned} \cos x\cos y &= \tfrac{1}{2}\cos(x+y)+\tfrac{1}{2}\cos(x-y) \;= \\ &= \tfrac{1}{4}\cos(x+y)+\tfrac{1}{4}\cos(x-y) \;+ \\ &\quad + \tfrac{1}{4}\cos(-x+y)+\tfrac{1}{4}\cos(-x-y)\;, \end{aligned} \quad (22)$$

so that

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

The right-hand series in (23) has terms $\cos \nu x$ where ν is a non-negative integer $\le\, m_1+\ldots+m_k$ with the same parity as $m_1+\ldots+m_k$. Furthermore, there holds (see Ref. 9, 9.1.44–45 on p. 361)

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

Next we have from the explicit form

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

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

$$R_{n_1}^{m_1}(\rho) \ldots R_{n_k}^{m_k}(\rho) \quad (26)$$

is a polynomial in ρ of degree $n_1+\ldots+n_k$ with non-zero coefficients of the powers ρ^t only for integers t satisfying

$$m_1+\ldots+m_k \le t \le n_1+\ldots+n_k \quad (27)$$

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

We thus conclude that any term in the right-hand side series in (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 $n, m \geq 0$ and $n-m$ even and ≥ 0 and where the β_{nm} can be explicitly expressed in terms of the α_{nm} using (23–25). Hence we can express the k^{th} order term in the expansions (3-4) of $U(x,y)$,

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

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

4 Explicit result for coma and astigmatism

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

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

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

For the first case in (30) we have

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

Using Ref. 9, 9.1.44–45 on p. 361, for either exponential in the inner integral and carrying out the integration over θ, we have

$$U(x,y)=2\int_0^1 \rho\exp(if\rho^2)\,J_0(\alpha\,\rho^3)\,J_0(2\pi\rho r)\,d\rho\ +$$

$$+\ 4\sum_{j=1}^{\infty}(-1)^j\int_0^1 \rho\exp(if\rho^2)\,J_j(\alpha\,\rho^3)\,J_j(2\pi\rho r)\,d\rho\cos j\phi\ . \tag{32}$$

Next, by using the power series expansion of the j^{th} Bessel function with argument $z=\alpha\rho^3$, see Ref. 9, 9.1.10 on p. 360, we get

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

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

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

where for $m=0,1,...$

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

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

In (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 (30). Then we have

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

Then as above we have

$$\int_0^{2\pi}\exp(i\gamma\rho^2\cos 2(\theta+\phi)\exp(2\pi i\rho r\cos\theta)\,d\theta\ =$$

$$=\ 2\pi\,J_0(\gamma\rho^2)\,J_0(2\pi\rho r)+4\pi\sum_{j=1}^{\infty}(-i)^j\,J_j(\gamma\rho^2)\,J_{2j}(2\pi\rho r)\cos 2j\phi\,. \tag{38}$$

There results

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

where for $m=0,1,...$

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

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

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

In Fig.2 we show $\mid U(x,y)\mid$ for the case of coma, see (30), with $\alpha=1$ and $f=0,\pi/4,\pi,2\pi$, calculated according to (34-36) including the terms with $l=0,1,2,3$.

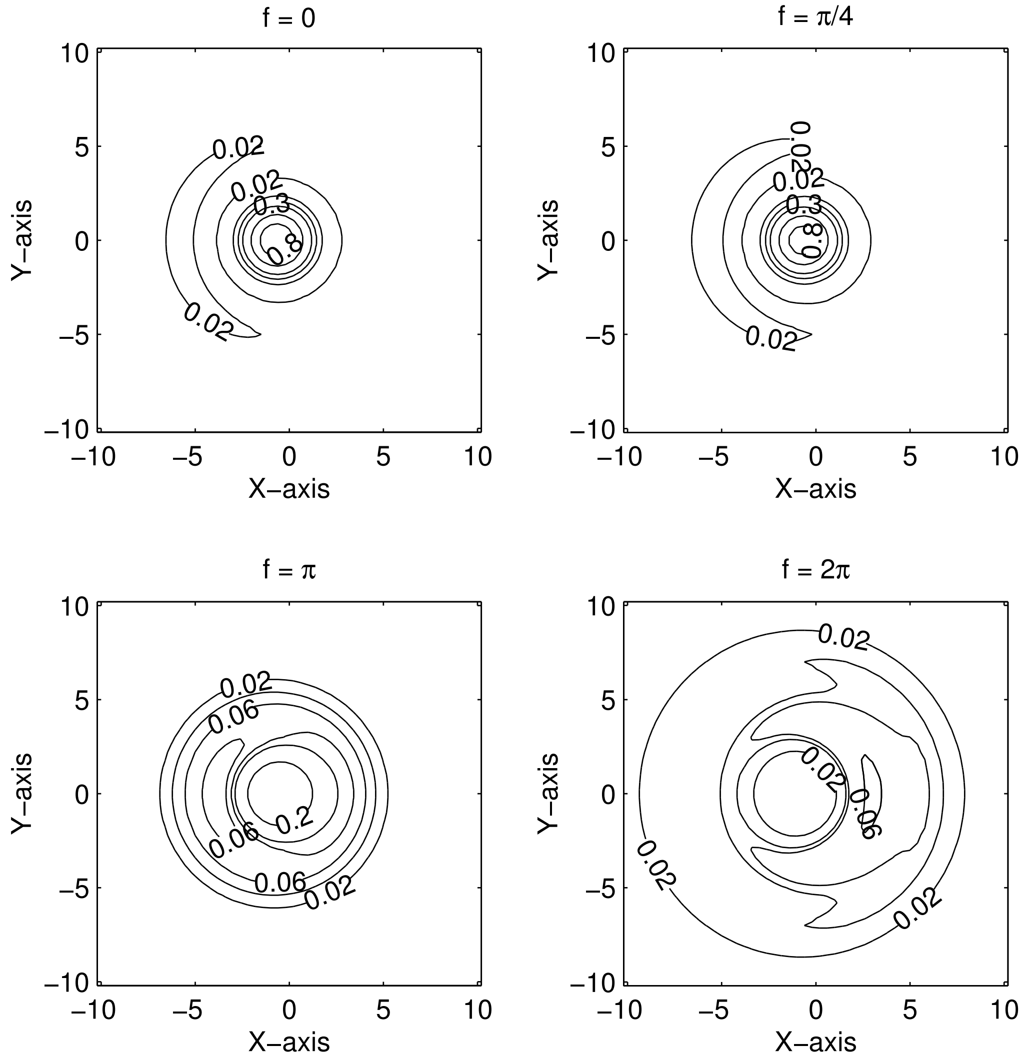

Figure 2: Contour plot of the modulus of the point spread function $U(x, y)$ with aberration $\Phi(\rho, \theta) = \alpha\rho^3 \cos\theta$ (coma), where $\alpha = 1$ and $f = 0, \pi/4, \pi, 2\pi$.

5 Conclusion

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

Appendix A. Derivation of formulas (14) and (19)

A.1 Derivation of formula (14)

We shall now show that for integer $n, m \geq 0$ with $n - m$ even and ≥ 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 $v = 2\pi r$, $p = \frac{1}{2}(n-m)$, $q = \frac{1}{2}(n+m)$ as before, and

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

We start by noting that

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

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

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

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

Then, with H_{nm} given as in (**??**), 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 before.

We claim that for $l = 1, 2, ...$

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

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

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

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

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

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

The right-hand side of (A9) agrees with the right-hand side of (A7) for the case $l = 1$ when we also observe that

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

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

$$\frac{m+2j+l}{q+1} (-1)^j \frac{\binom{p}{j} \binom{m+j+l-1}{l-1}}{\binom{q+j+l}{q+1}} =$$

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

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

$$F_{nm}^{(l+1)}(v) = \int_0^v t\,F_{nm}^{(l)}(t)\,dt = \sum_{j=0}^{\infty} \frac{\Gamma(-p+j)}{\Gamma(-p)} \frac{\Gamma(j+l)}{\Gamma(j+1)} \frac{\Gamma(q+1)}{\Gamma(q+j+l+1)} \ .$$

$$\cdot \binom{m+j+l-1}{l-1} (m+2j+l) \int_0^v t^{m+l+1} \, J_{m+2j+l}(t)\, dt \ . \tag{A12}$$

Next (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

$$\int_0^v t^{n+l+1} \, J_{m+2j+l}(t)\, dt \ =$$

$$= \ v^{n+l+1} \sum_{k=0}^{\infty} \frac{(m+2(j+k)+l+1)\,\Gamma(-p+j+k)\,\Gamma(q+j+l+1)}{\Gamma(-p+j)\,\Gamma(q+l+j+k+2)} \ \cdot$$

$$\cdot \ J_{m+2(j+k)+l+1}(v) \ . \tag{A13}$$

Inserting (A13) into (A12) and simplifying we obtain

$$\begin{aligned} F_{nm}^{(l+1)}(v) \ &= \ v^{n+l+1} \sum_{j,k=0}^{\infty} \frac{\Gamma(-p+j+k)}{\Gamma(-p)} \ \frac{\Gamma(j+l)}{\Gamma(j+1)} \ \frac{\Gamma(q+1)}{\Gamma(q+j+k+l+2)} \ \cdot \\ &\cdot \binom{m+j+l-1}{l-1} (m+2j+l)(m+2(j+k)+l+1) \ \cdot \\ &\cdot \ J_{m+2(j+k)+l+1}(v) \ . \end{aligned} \tag{A14}$$

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

$$\begin{aligned} F_{nm}^{(l+1)}(v) \ &= \ v^{n+l+1} \sum_{s=0}^{\infty} \frac{\Gamma(-p+s)}{\Gamma(-p)} \ \frac{\Gamma(q+1)}{\Gamma(q+s+l+2)} \,(m+2s+l+1) \ \cdot \\ &\cdot \ J_{m+2s+l+1}(v) \ \cdot \\ &\cdot \sum_{j=0}^{s} \frac{\Gamma(j+l)}{\Gamma(j+1)} \binom{m+j+l-1}{l-1} (m+2j+l) \ . \end{aligned} \tag{A15}$$

Next we consider the series on the last line of (A15). There holds

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

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

$$\sum_{j=0}^{s} (j+l-1) \ldots (j+1)(m+j+l-1) \ldots (m+j+1)(m+2j+l) \ =$$

$$= \frac{1}{l}\,(s+l)\ldots(s+1)(m+s+l)\ldots(m+s+1) \tag{A17}$$

which is conveniently used in the induction step.

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

$$\begin{aligned} F_{nm}^{(l+1)}(v) &= v^{n+l+1} \sum_{s=0}^{\infty} \frac{\Gamma(-p+s)}{\Gamma(-p)}\, \frac{\Gamma(q+1)}{\Gamma(q+s+l+2)}\, \frac{\Gamma(s+l+1)}{\Gamma(s+1)} \,\cdot \\ &\quad \cdot \binom{m+s+l}{l}\,(m+2s+l+1)\,J_{m+2s+l+1}(v)\ , \end{aligned} \tag{A18}$$

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

A.2 Derivation of formula (19)

We shall now show that for integer $n, m \geq 0$ with $n-m$ even and ≥ 0

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

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

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

To show (A19–A20) we have from (25)

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

with the $T_{n-2s,m}$ as in (13). We next insert formula (14) into the right-hand side of (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

$$\begin{aligned} &\int_0^1 \rho\, R_n^m(\rho)\exp(if\rho^2)\,J_m(2\pi\rho r)\,d\rho = \\ &= \exp(if)\sum_{l=1}^{\infty}(-2if)^{l-1}\Big\{\sum_{s=0}^{p}\sum_{j=0}^{p-s}\frac{(-1)^s\,(n-s)!\,t_{lj}(n-2s,m)}{s!\,(q-s)!\,(p-s)!}\,\frac{J_{m+l+2j}(v)}{v^l}\Big\}\ , \end{aligned} \tag{A22}$$

where

$$t_{lj}(n-2s,m) \ =$$

$$= \ (-1)^j \, \frac{m+l+2j}{q-s+1} \binom{p-s}{j} \binom{m+j+l-1}{l-1} \Big/ \binom{j+q-s+l}{q-s+1} \ =$$

$$= \ (-1)^j \, (m+l+2j) \binom{m+j+l-1}{l-1} \frac{(p-s)!\,(q-s)!}{(p-j-s)!\,(j+q-s+l)!} \ \cdot$$

$$\cdot \ \frac{(j+l-1)!}{j!} \ . \qquad \text{(A23)}$$

On simplifying we then find that the quantity in { } at the right-hand side of (A22) is given by

$$\{\ \} \ = \ \sum_{s=0}^{p} \sum_{j=0}^{p-s} (-1)^j \, (m+l+2j) \binom{m+j+l-1}{l-1} \frac{(j+l-1)!}{j!} \ \cdot$$

$$\cdot \ \frac{J_{m+l+2j}(v)}{v^l} \ \frac{(-1)^s \, (n-s)!}{s!\,(p-j-s)!\,(j+q-s+l)} \ . \qquad \text{(A24)}$$

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

$$\{\ \} \ = \ \sum_{j=0}^{p} (-1)^j \, (m+l+2j) \binom{m+j+l-1}{l-1} \frac{(j+l-1)!}{j!} \ \frac{J_{m+l+2j}(v)}{v^l} \ \cdot$$

$$\cdot \ \sum_{s=0}^{p-j} \frac{(-1)^s \, (n-s)!}{s!\,(p-j-s)!\,(j+q-s+l)!} \ . \qquad \text{(A25)}$$

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

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

The last identity in (A26) follows at once from the definition of the binomials.

We shall show now that for $k=0,1,...,p$ and integer $i<k$ we have

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

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

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

where

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

and where Δ^k denotes the $k^{\rm th}$ power of the backward difference operator Δ, defined as

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

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

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

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

where

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

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

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

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

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

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

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

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

Therefore, the quantity in (A22) between { } equals

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

where

$$\begin{aligned} v_{lj} &= (-1)^p\,(m+l+2j)\binom{m+j+l-1}{l-1}\frac{(j+l-1)!}{j!}\ \cdot \\ &\quad \cdot \binom{l-1}{p-j}\frac{(q+j)!}{(q+l+j)!}\,l\ = \\ &= (-1)^p\,(m+l+2j)\binom{m+j+l-1}{l-1}\binom{j+l-1}{l-1}\binom{l-1}{p-j}\Big/ \\ &\qquad\qquad \Big/\binom{q+l+j}{l}\ , \end{aligned} \tag{A37}$$

and this is what we wanted to prove.

Acknowledgements

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

Author contact information

A.J.E.M. Janssen, Philips Research Laboratories, WY-81
Prof. Holstlaan 4, 5656 AA Eindhoven, The Netherlands.
Tel.: +31 40 274 2402
Fax.: +31 40 274 4675
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 opticalpoint spread functions", (submitted/to appear) J. Opt. Soc. Am. A.

2. M. Born and E. Wolf, *Principles of Optics* (Pergamon Oxford, UK, 1985).

3. B.R.A. Nijboer, *The Diffraction Theory of Aberrations* (Thesis University of Groningen, The Netherlands, 1942).

4. R.J. Noll, "Zernike polynomials and atmospheric turbulence", J. Opt. 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 wave- front sensor: A 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* (Thesis University of Groningen, The Netherlands, 1948).

9. M. Abramowitz and I.A. Stegun, *Handbook of Mathematical Functions* (Dover, New York, 1970).

Figure captions

Figure 1:
A point source in O emits a spherical wave towards the schematically represented optical system. In the image space, an aberrated wavefront W leaves the exit pupil (center in E) and comes to a focus close to the image plane through $O^{'}$. The spherical reference wavefront 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 via $\Phi = 2\pi W/\lambda$ with λ the wavelength of the monochromatic radiation. The normalized cartesian pupil coordinates are denoted by (ν, μ), the coordinates (x, y) in the image plane have been normalized with respect to the diffraction unit, λ/NA, with NA 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.

Figure 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 $f = 0, \pi/4, \pi, 2\pi$.