Calculation of the Rayleigh–Sommerfeld diffraction integral by exact integration of the fast oscillating factor

Jan A. C. Veerman, Jurgen J. Rusch, and H. Paul Urbach

Philips Research Laboratories, Professor Holstlaan 4, 5656 AA Eindhoven, The Netherlands

Received August 6, 2004; accepted September 11, 2004

We describe a numerical method that can be used to calculate the propagation of light in a medium of constant (possibly complex) index of refraction n. The method integrates the Rayleigh–Sommerfeld diffraction integral numerically. After an appropriate change of integration variables, the integrand of the diffraction integral is split into a slowly varying and an (often fast) oscillating quadratic factor. The slowly varying factor is approximated by a spline fit, and the resulting Fresnel integrals are subsequently integrated exactly. Although the method is not as fast as methods involving a fast Fourier transform, such as plane-wave propagation or Fresnel approximation, it is accurate over a greater range than these methods. © 2005 Optical Society of America

OCIS codes: 050.0050, 260.0260.

1. INTRODUCTION

Nowadays, storage of information with the use of optical devices such as lasers has many applications. Therefore thorough theoretical knowledge of processes such as near-field recording and optical disc mastering is desirable. This has been one of the reasons for the study of several aspects of light propagation. Several studies have been made of how to calculate light intensities in the focal region.[1–5] At Philips Research Laboratories, detailed studies have been made of the optical disc mastering process[6] and of the scattering of a focused spot by a periodic grating.[7] In the latter study the fields were calculated for distances of only a few wavelengths above the grating. In addition, although the fields were calculated in three dimensions, the shape of the grating was restricted to a one-dimensionally periodic grid, i.e., a grid that was periodic in the x direction and had a uniform shape in the y direction. Currently, a computer program that will calculate the field in and just above an arbitrary pit structure is being developed. In addition, we undertook a study on how to calculate the field accurately over a distance of many wavelengths in a medium of constant index of refraction, as occurs often above such a pit structure. The results of this work are presented here. By combining these two methods, we will be able to calculate the propagation of light from the near field to the intermediate and far field in an accurate manner. The method we used requires the calculation of the Rayleigh–Sommerfeld (RS) diffraction integral. We endeavor to calculate this integral directly, not by using a series expansion. After some transformations of coordinates and integration along the angular coordinate, the problem reduces to that of calculating a number of integrals of the form

$$I = \int_0^{s_1} G(s)\exp(iks^2)\mathrm{d}s, \tag{1.1}$$

which contains, in the general case, a slowly varying function $G(s)$ and a fast oscillating exponential term. This integral is calculated by approximation of $G(s)$ by a cubic spline fit and subsequent exact calculation of the resulting Fresnel integrals. A typical property of this method is that the error made in approximating the diffraction integral is uniform even in the limit $k \to \infty$.

The problem of numerical calculation of a diffraction integral has been the subject of several earlier studies.[3,8–10] Török *et al.* have shown in Ref. 3 that a series expansion is, in general, slower and less accurate than a direct integration of the integral. The method described by d'Arcio *et al.*[10] also involves the computation of Fresnel integrals. Here, however, the entire integrand, including the exponential term, is approximated by a polynomial and a parabolic phasor. This approximation is valid on a set of small two-dimensional rectangles or regions with a parabolic boundary. The resulting integrals are then integrated exactly. In our case, the slowly varying function is approximated separately, without inclusion of the exponential term. Therefore our method is fundamentally different from the one used in Ref. 10. The method we developed is compared with two other methods that are less time consuming but have a narrower range of validity. We show that the present method is valid over a large range.

The organization of this paper is as follows. In Section 2 an overview is given of the plane-wave propagation (PWP) and Fresnel approximation (FA) methods. Section 3 discusses the RS method. Finally in Section 4 the application of the three methods to various spots is discussed. In Appendix A numerical details of the RS method are discussed in more detail. Throughout this paper we assume that the time dependence of the fields is given by the factor $\exp(-i\omega t)$. We omit this factor from all equations.

1084-7529/2005/040636-11$15.00

2. OVERVIEW OF FAST-BUT-LIMITED METHODS

In this section we describe two methods that can be used to calculate the propagation of the field in a fast way but have a limited range of validity as well as limitations with respect to the spots for which the propagation can be calculated.

A. Plane-Wave Propagation Method

In the PWP method[11] the given field at $z = 0$ is decomposed into its plane-wave components by means of a Fourier transform:

$$\mathcal{F}(U)(k_x, k_y) = \frac{1}{4\pi^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} U(x, y, z = 0) \times \exp(-ik_x x - ik_y y)\mathrm{d}x\mathrm{d}y. \quad (2.1)$$

The Fourier transform can be performed by a fast Fourier transform (FFT) algorithm. Subsequently, the plane waves are propagated over a distance z_0 by multiplication with a factor $\exp(ik_z z_0)$ with

$$k_z = \left[\frac{(2\pi n)^2}{\lambda^2} - k_x^2 - k_y^2\right]^{1/2}. \quad (2.2)$$

Here λ is the wavelength in vacuum, and n is the index of refraction. Here the branch of the square root is chosen with the cut along the negative axis and with $\sqrt{x} \geqslant 0$ and with $\sqrt{-x} = i\sqrt{x}$ for $x > 0$. Since n is in the first quadrant, k_z is always in the first quadrant. k_z can be imaginary when k_x and k_y are large, namely, for evanescent waves. When the medium absorbs light, n^2 is complex, and hence k_z is complex also, for all k_x and k_y. The field in the plane $z = z_0$ can be calculated from the field in k space by a backward Fourier transform:

$$U(x, y, z = z_0) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathcal{F}(U)(k_x, k_y) \times \exp(ik_z z_0)\exp(ik_x x + ik_y y)\mathrm{d}k_x\mathrm{d}k_y. \quad (2.3)$$

This method is very fast when a FFT is used, but it also has some disadvantages:

1. The rectangular computational domain is fixed in both (x, y) space and (k_x, k_y) space, once a certain grid has been chosen in the $z = 0$ plane. This means that the region in which the field can be calculated is restricted to this computational domain.
2. The field needs to fall off to zero toward the boundary of the computational domain.
3. Most importantly, when z_0 becomes large, the factor $\exp(ik_z z_0)$ will start to oscillate rapidly, causing aliasing effects. This could in principle be resolved by making the grid in k space denser as z_0 increases and interpolating the function $\mathcal{F}(U)$ on this new dense grid. This is, however, rather cumbersome and will lead to an intolerable increase in computer time and storage requirements.

B. Fresnel Approximation

The FA[11] for propagation of light holds when the light has propagated over many wavelengths away from the initial plane at $z = 0$. [Here "many" is defined by expression (2.6) below.] In this case the following formula can be derived by application of the method of stationary phase to Eq. (2.3)[11]:

$$U(x, y, z_0) \approx -i\exp\left(2\pi i\frac{z_0 n}{\lambda}\right)\left(\frac{n}{z_0\lambda}\right) \times \exp\left[\frac{\pi i n'}{z_0\lambda}(x^2 + y^2)\right] \times \iint \mathcal{F}\left\{U(x', y', z = 0) \times \exp\left[\frac{\pi i n'}{z_0\lambda}(x'^2 + y'^2)\right]\exp\left[\frac{2\pi n''}{xz_0}(x_0 x' + y_0 y')\right]\right\}\left(\frac{xn'}{z_0\lambda}, \frac{yn'}{z_0\lambda}\right)\mathrm{d}x'\mathrm{d}y', \quad (2.4)$$

where

$$n = n' + in'' \quad (2.5)$$

is the complex index of refraction. This equation involves one FFT, and the resulting grid in (x, y) space is linearly diverging with distance z_0. In general, therefore, when the initial field lies well within the region of definition, the resulting field at greater values $z = z_0$ is well contained within the region covered by the mesh. The FA becomes exact when $z_0 \to \infty$. A criterion for good approximation is given in Ref. 11: The FA is large with respect to the next-order term if

$$z_0^3 \gg \frac{\pi n'}{4\lambda}[(x - x')^2 + (y - y')^2]^2_{\max}, \quad (2.6)$$

with x, y, x', y' as in expression (2.4), and the maximum taken over all integration points (x', y'). Generally, this criterion is rather strict, and the FA is often valid for smaller z_0, in particular for paraxial fields such as laser beams.

The combination of both methods might yield a correct result for all values of z_0, provided that an overlap region of z_0 values exists. Detailed study revealed, however, that this is not always the case and depends, e.g., on the field that is used.[12] Thus, for fields that are not well behaved in either (k_x, k_y) or (x, y) space, such as spots that have been cut off by an aperture or spots that fall off to zero slowly in either of these spaces, overlapping regions will not occur. Therefore we developed a third robust method for the calculation of the propagation of a field over arbitrary distances. This method involves the numerical integration of the RS diffraction integral[11] and is able to calculate field propagation accurately for arbitrary distances.

3. INTEGRATION OF THE RAYLEIGH–SOMMERFELD DIFFRACTION INTEGRAL

A. Introduction

This method of integration has been used before[6] to calculate the propagation of a field diffracted by a aperture Ω toward the entry plane of a lens. The program that performed these calculations was called MASTER, after the process of optical disc mastering that was studied with the use of this program. We modified this program to calculate the field propagation for arbitrary z_0. It should be noted beforehand that the RS diffraction integral also, like Eq. (2.3) of the PWP, contains an exponential term that will oscillate rapidly when z_0 increases. In this case, however, we devised a method to calculate the integral containing this term with greater accuracy and without the need to decrease the mesh size with increasing distance z_0.

B. Theory

The RS diffraction integral is given by[11]

$$U(x_0, y_0, z_0) = -\int\int_{\Omega} U(x, y, 0)\frac{2z_0}{|\mathbf{r}_0 - \mathbf{r}|} \times \left(ik - \frac{1}{|\mathbf{r}_0 - \mathbf{r}|}\right)\frac{\exp(ik|\mathbf{r}_0 - \mathbf{r}|)}{4\pi|\mathbf{r}_0 - \mathbf{r}|}\mathrm{d}x\mathrm{d}y, \quad (3.1)$$

where

$$k = \omega\sqrt{\epsilon_0\mu_0 n}. \quad (3.2)$$

Here Ω is the region in the $z = 0$ plane where U is nonzero. Ω could be the entire plane; but in the case of, e.g., an aperture in an opaque screen, Ω would be equal to the aperture. As is well known, Eq. (3.1) is exact when $U(x, y, z = 0)$ is known. We shall assume that Ω is a rectangle with sides parallel to the x and y axes. When initially Ω is the entire plane, Ω is chosen so large that U is negligible in the exterior of Ω. The assumption that Ω is rectangular can be relaxed: The method that will be described next can be generalized to the case in which the boundary of Ω is an arbitrary polygon.

The point of observation (x_0, y_0, z_0) may lie inside or outside of Ω. Note that Eq. (3.1) is different from Eq. (2.3) in Ref. 6. Equation (3.1) is the exact equation, which includes the term of the integrand that is proportional to $1/|\mathbf{r}_0 - \mathbf{r}|^3$. Equation (2.3) in Ref. 6 follows from Eq. (3.1) when $k \gg 1/|\mathbf{r}_0 - \mathbf{r}|$. In our program we want to be able to calculate the field for arbitrary values of $\mathbf{r}_0$ and k, and therefore we use the exact equation. The field $U(x, y, 0)$ could for example have been obtained by solving Maxwell's equations rigorously, and the exact formula [Eq. (3.1)] then gives the field in a plane at arbitrary distance z_0.

We now translate the integration region Ω over the vector $(x_0, y_0, 0)$,

$$\tilde{\Omega} = \Omega - (x_0, y_0, 0), \quad (3.3)$$

and we obtain

$$U(x_0, y_0, z_0) = -\int\int_{\tilde{\Omega}} U(x_0 + x, y_0 + y, 0)\frac{2z_0}{(x^2 + y^2 + z_0^2)^{1/2}} \times \left[ik - \frac{1}{(x^2 + y^2 + z_0^2)^{1/2}}\right] \times \frac{\exp[ik(x^2 + y^2 + z_0^2)^{1/2}]}{4\pi(x^2 + y^2 + z_0^2)^{1/2}}\mathrm{d}x\mathrm{d}y. \quad (3.4)$$

We define polar coordinates r and ϕ around the point of observation (x_0, y_0) by

$$r = (x^2 + y^2)^{1/2}, \quad (3.5)$$

$$x = r\cos\phi, \qquad y = r\sin\phi \quad (3.6)$$

and the variable ρ by

$$\rho = (x^2 + y^2 + z_0^2)^{1/2} - z_0 = (r^2 + z_0^2)^{1/2} - z_0. \quad (3.7)$$

Then

$$r = \rho\left(1 + \frac{2z_0}{\rho}\right)^{1/2}. \quad (3.8)$$

After changing the integration variables to ϕ and ρ, we obtain

$$U(x_0, y_0, z_0) = \int\int_{\tilde{\tilde{\Omega}}} \mathbf{f}(\rho, \phi)\frac{\exp(ik\rho)}{\rho + z_0}\mathrm{d}\phi\mathrm{d}\rho, \quad (3.9)$$

where

$$\mathbf{f}(\rho, \phi) = \frac{z_0 n\exp(ikz_0)}{i\lambda}U\{x_0 + [(\rho + z_0)^2 - z_0^2]^{1/2} \times \cos\phi, y_0 + [(\rho + z_0)^2 - z_0^2]^{1/2} \times \sin\phi, 0\}\left[1 - \frac{1}{ik(\rho + z_0)}\right] \quad (3.10)$$

and $\tilde{\tilde{\Omega}}$ is the image of $\tilde{\Omega}$ under the mapping $(x, y) \rightarrow (\rho, \phi)$. The integral over ρ contains the factor $\exp(ik\rho)$, which performs many oscillations as a function of ρ within $\tilde{\tilde{\Omega}}$ when $z_0 \rightarrow \infty$ is large. In comparison, the function $f(\rho, \phi)$ is slowly varying as a function of ρ, whatever the value of z_0. When the field is well behaved, which should be the case in a physical situation, the integrand is also a slowly varying function of ϕ for fixed ρ. The integration with respect to ϕ can, for fixed values of ρ, be performed to great accuracy with, e.g., the use of a Gaussian integration method.[13]

Note that the region $\tilde{\tilde{\Omega}}$ in the variables (ρ, ϕ) is rather complicated. To be able to integrate over $\tilde{\tilde{\Omega}}$, we determine a number of important circles $l = 1,..., N$, which are defined as follows: A circle with center $(x_0, y_0, 0)$ either intersects a corner point of the grid region or is tangential to one of the sides of this region (see Fig. 1). Additionally, if the point $(x_0, y_0, 0)$ lies inside the grid region, the circle with radius $\rho = 0$ is added. The circles are ordered according to increasing values of ρ. We now have a

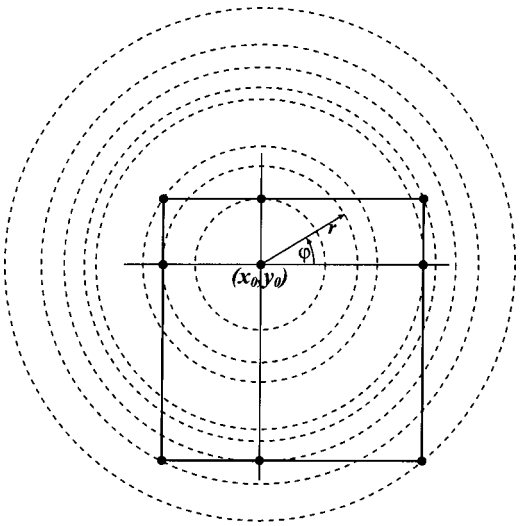

Fig. 1. Important circles for the point (x_0, y_0) in the case of a rectangular diaphragm Ω. The integral over ϕ runs over the overlap region of Ω and the area enclosed in two adjacent circles. Note that when the point (x_0, y_0) lies within Ω the first circle has radius $r = 0$.

number of intervals $\rho_1(l) \leqslant \rho < \rho_2(l)$, $l = 1,...N$, so that for ρ in interval l the circle with center $(x_0, y_0, 0)$ and radius ρ intersects the region Ω in some number, $M(l)$, say, of arcs $\phi_{2j-1} < \phi < \phi_{2j}$, $j = 1,..., M(l)$. We then have

$$U(x_0, y_0, z_0) = \int\int_{\tilde{\Omega}} \mathbf{f}(\rho, \phi) \frac{\exp(ik\rho)}{\rho + z_0} \mathrm{d}\phi \mathrm{d}\rho = \sum_{l=1}^{N} \int_{\rho_1(l)}^{\rho_2(l)} \exp(ik\rho)\mathbf{F}_l(\rho)\mathrm{d}\rho, \quad (3.11)$$

with, for $\rho_1(l) \leqslant \rho < \rho_2(l)$,

$$\mathbf{F}_l(\rho) = \frac{1}{\rho + z_0} \sum_{j=1}^{M(l)} \int_{\phi_{2j-1}(\rho)}^{\phi_{2j}(\rho)} \mathbf{f}(\rho, \phi)\mathrm{d}\phi. \quad (3.12)$$

The function $\mathbf{F}_l(\rho)$ is obtained by integrating over all ϕ segments at fixed ρ. This integration can be performed by Gaussian integration if we can determine the value of $\mathbf{f}(\rho, \phi)$ at the abscissas ϕ_k, $k = 1,..., N_{\mathrm{Gauss}}$, that are needed for the Gaussian integration. Thus we need a good approximation at points $(x, y) = (r(\rho) \times \cos\phi, r(\rho)\sin\phi)$ of the function $\mathbf{f}(\rho(x, y), \phi(x, y))$. Here the points (x, y) do not necessarily need to be grid points where the spot function $U(x, y)$ is known. To determine $\mathbf{f}(\rho, \phi)$ for such points, we approximated $U(x, y)$ by a bicubic spline in x and y and subsequently determined $\mathbf{f}(\rho, \phi)$ with the aid of Eq. (3.10). It appeared that, by use of this spline, the Gauss integral with respect to ϕ was convergent to within 1% when the number of integration points was approximately 20 when U was sufficiently smooth (we tested this for four types of spot; see Subsection 4.B).

After the integral over ϕ has been performed, we obtain the integrals

$$U(x_0, y_0, z_0) = \sum_{l=1}^{N} \int_{\rho_1(l)}^{\rho_2(l)} \mathbf{F}_l(\rho)\exp(ik\rho)\mathrm{d}\rho. \quad (3.13)$$

Just as in the PWP method, we have here an oscillatory factor that will oscillate rapidly as a function of ρ when z_0 becomes large. In this case, however, the integration including this factor can be performed with greater accuracy than in the case of the PWP method. We first note the following: $\mathbf{F}_l(\rho)$, when compared with the factor $\exp(ik\rho)$, is a slowly varying function of ρ that may, however, not be smooth near the end points $\rho = \rho_1(l)$, $\rho = \rho_2(l)$ of the integration interval over ρ. In fact, for $\rho = \rho_1(l)$, the number $M(l)$ of arcs $\phi_{2j-1}(\rho) \leqslant \phi \leqslant \phi_{2j}(\rho)$ of circles with center $(x_0, y_0, 0)$ that intersects $\tilde{\Omega}$ changes. This means that the circles with radius $\rho_1(l)$ or $\rho_2(l)$ or both types of circle are at some point tangent to the boundary of $\tilde{\Omega}$. Suppose that this happens for the circle with radius $\rho_1(l)$. Then for this radius, two points of intersection, corresponding to $\phi = \phi_{2j-1}(\rho)$ and $\phi = \phi_{2j}(\rho)$, say, coincide. Therefore, for ϕ near $\phi_{2j-1}(\rho) = \phi_{2j}(\rho)$, the derivative of $\rho(\phi)$ at the point of intersection will vanish. Hence we have for $\phi \approx \phi_{2j-1}(\rho) = \phi_{2j}(\rho)$:

$$\rho(\phi) \approx \rho_1(l) \times \frac{C}{2}[\phi - \phi_{2j-1}(\rho)]^2, \quad (3.14)$$

where C is the second derivative of ρ with respect to ϕ at $\phi = \phi_{2j-1}(\rho)$. Hence the inverse $\phi(\rho)$ has for $\rho \approx \rho_1(l)$ the expansion

$$\mathbf{F}_l(\rho) = a_0 + a_1(\rho - \rho_1)^{1/2} + a_2(\rho - \rho_1) + a_3(\rho - \rho_1)^{3/2} + \ldots, \quad (3.15)$$

for some a_0, a_1, a_2, a_3, etc. Similarly, for $\rho \approx \rho_2(l)$, we have

$$\mathbf{F}_l(\rho) = b_0 + b_1(\rho_2 - \rho)^{1/2} + b_2(\rho_2 - \rho) + b_3(\rho_2 - \rho)^{3/2} + \ldots. \quad (3.16)$$

[Note that when the circle with $\rho = \rho_2(l)$ is *not* tangent to the boundary of $\tilde{\Omega}$, it holds that $b_j = 0$ for all odd j in this expansion.] Hence $F_l(\rho)$ has similar expansions near $\rho \approx \rho_1(l)$ and $\rho \approx \rho_2(l)$, and therefore $F_l(\rho)$ is smooth except possibly for square-root singularities of the derivative $F_l'(\rho)$ at the end points of the integration intervals $\rho_1(l) < \rho < \rho_2(l)$. This lack of smoothness will decrease the accuracy when the integration interval includes both boundaries. To prevent this, we split the integration area into two halves and introduce the new integration variables s:

$$\int_{\rho_1}^{\rho_2} \exp(ik\rho)\mathbf{F}_l(\rho)\mathrm{d}\rho$$
$$= \int_{\rho_1}^{(\rho_1+\rho_2)/2} \exp(ik\rho)\mathbf{F}_l(\rho)\mathrm{d}\rho + \int_{(\rho_1+\rho_2)/2}^{\rho_2} \exp(ik\rho)\mathbf{F}_l(\rho)\mathrm{d}\rho$$
$$= \exp(ik\rho_1)\int_0^{\sqrt{(\rho_2-\rho_1)/2}} \exp(iks^2)\mathbf{F}_{\text{left}}(s)\mathrm{d}s + \exp(ik\rho_2)\left[\int_0^{\sqrt{(\rho_2-\rho_1)/2}} \exp(iks^2)\mathbf{F}_{\text{right}}(s)\mathrm{d}s\right]^*, \tag{3.17}$$

with

$$s = \sqrt{\rho - \rho_1}, \tag{3.18}$$

in the first integral, and

$$s = \sqrt{\rho_2 - \rho}, \tag{3.19}$$

in the second, and with

$$\mathbf{F}_{\text{left}}(s) = 2s\mathbf{F}_l(\rho_1 + s^2), \tag{3.20}$$

$$\mathbf{F}_{\text{right}}(s) = 2s\mathbf{F}_l(\rho_2 - s^2)^*. \tag{3.21}$$

Now consider the integral

$$I = \int_0^{\sqrt{(\rho_2-\rho_1)/2}} \exp(iks^2)G(s)\mathrm{d}s. \tag{3.22}$$

In the above situation, $G(s)$ is a smooth and slowly varying function. The exponential term can fluctuate wildly, however. We now approximate $G(s)$ by a cubic spline $\sigma(s)$. The resulting integral, including the fast oscillating term, then consists of a number of Fresnel integrals, which can be calculated exactly (see Appendix A). It is shown in Appendix B that the approximation of the function by a spline fit results in an approximation of the integral that has a relative error that is asymptotically uniformly bounded with respect to k, when $k \to \infty$. In fact, it is shown in Appendix B that

$$\frac{\left|\int_0^{\sqrt{(\rho_2-\rho_1)/2}} \exp(iks^2)G(s)\mathrm{d}s - \int_0^{\sqrt{(\rho_2-\rho_1)/2}} \exp(iks^2)\sigma(s)\mathrm{d}s\right|}{\left|\int_0^{\sqrt{(\rho_2-\rho_1)/2}} \exp(iks^2)G(s)\mathrm{d}s\right|} \leq C\frac{h^2(h+1)}{\sqrt{k}}, \tag{3.23}$$

where h (>0) is the mesh size used in the definition of the spline and C is a constant that depends on G but is independent of h and k. Thus the fast oscillating factor $\exp(iks^2)$ in the integral is dealt with in an accurate and uniform manner and will therefore not give rise to artifacts in the solution. More precisely, the grid size to approximate G that is needed to obtain a specified accuracy in the calculation of the integral is completely independent of the fast oscillating term $\exp(iks^2)$ but depends only on the function G, which in practice is slowly varying. Hence the grid used in the integration over s of the numerical calculation of the integrals is independent of the value of k; i.e., it can be chosen to be the same for all k.

4. CALCULATION OF PROPAGATION FOR VARIOUS FIELDS

To test the new program, we calculated the propagation of four types of field with it, and we compared the results of the calculations with those of the PWP and FA methods. Below we will give an account of these calculations.

A. Calculation Conditions

All calculations were performed with the following conditions.

- The index of refraction was that of vacuum:

$$n = 1. \tag{4.1}$$

- The wavelength of the light was

$$\lambda = 0.4579 \times 10^{-6}\ \text{m}. \tag{4.2}$$

This is the wavelength of a laser that was also used in Ref. 7.

- The computational rectangle at $z = 0$ was taken to be a square with a width equal to

$$L_x = L_y = L = 0.011424 \times 10^{-3}\ \text{m}. \tag{4.3}$$

This square was also taken to be the diaphragm Ω for the RS calculations. The square is approximately 25 wavelengths wide.

- The propagated fields were calculated for five values of z_0:

$$z_0(i_z) = 0.0075 \times 10^{-3} i_z\ \text{m} \qquad (i_z = 1,...,4),$$
$$z_0(0) = 1 \times 10^{-8}\ \text{m}; \tag{4.4}$$

$z_0(0)$ is taken to be approximately, but not exactly, zero because the FA cannot be solved numerically at $z_0 = 0$, and the method that uses the RS formula, although it has an existing limit for $z \downarrow 0$, runs into trouble when it is solved numerically for $z = 0$ in the same way as for $z > 0$ because the term $|\mathbf{r_0} - \mathbf{r}|$ in Eq. (3.1) can become equal to zero in this case. The maximum value of z_0, $z_0 = 0.030 \times 10^{-3}$ m, is approximately 65 wavelengths long.

• The initial fields were defined on a $N_x \times N_y$ grid, with

$$N_x = N_y = 40, \tag{4.5}$$

$$x(l,j) = \left(l - \frac{N_x}{2}\right)L, \qquad l = 0,..., N_x - 1,$$

$$j = 0,...N_y - 1, \tag{4.6}$$

$$y(l,j) = \left(j - \frac{N_y}{2}\right)L, \qquad l = 0,..., N_x - 1,$$

$$j = 0,...N_y - 1, \tag{4.7}$$

where L is given by Eq. (4.3). These grid values were also the values of the x and y coordinates for which the field values were calculated in the case of the PWP and RS diffraction methods. In the case of the FA, the calculations were performed for points on a diverging grid:

$$x_{\mathrm{FA}}(l, j, i_z) = k_x(l, j)z_0(i_z)\lambda, \tag{4.8}$$

$$y_{\mathrm{FA}}(l, j, i_z) = k_y(l, j)z_0(i_z)\lambda. \tag{4.9}$$

In our case, this means that $|x_{\mathrm{FA}}(l, j, i_z)| > |x(l, j)|$ when $i_z \geqslant 1$.

• After the calculations of the field for the three methods were performed, the deviation of the PWP and FA fields from the RS field was calculated for all grid points for the absolute value as well as for the real and imaginary parts of the fields. Thus we calculated the relative deviation:

$$\epsilon_{\mathrm{rel}}(x, y, z) = \frac{|f_2(x, y, z) - f_1(x, y, z)|}{\|f_1\|_\infty}, \tag{4.10}$$

with $\|f_1\|_\infty$ as the maximum absolute value of the field f_1 for all grid points (x, y, z).

• The calculation of $\epsilon_{\mathrm{rel}}(x, y, z)$ for the FA in comparison with the RS method was performed for $i_z \geqslant 1$ after linear interpolation of the field values of the Fresnel grid onto the grid points of the RS (and PWP) method. This method yields results that are most reliable when the (diverging) Fresnel grid is approximately the same size as that of the RS method. In our case this is true for $i_z = 1$–3. For $i_z = 4$, the Fresnel grid is much larger than the RS grid. In this case the deviation of the Fresnel field from the RS field will not only be caused by differences in the calculations but also by the deviation due to the linear interpolation. In this case the calculated values of $\epsilon_{\mathrm{rel}}(x, y, z)$ may serve as an upper boundary.

• The calculational fields, as well as the real, imaginary, and absolute values of ϵ_{rel} were subsequently visualized with the use of the visualization package AMIRA.[14] Thus a direct comparison of the three types of calculation could be made.

• The calculations were performed on a dual Pentium 2.8-GHz, 4-Gb computer, running the Linux operating system (kernel level 2.4.20 with SMP support). The CPU time for the calculations was approximately 10 s for the PWP and FA methods and approximately 5 min for the RS method, in all cases for five planes of 20×20 points.

B. Initial Fields

We used four types of field for the initial field at $z_0 = 0$:

• Gaussian spot (Fig. 2):

$$f_{\mathrm{Gauss}} = \pi \frac{a}{(140)^2} \exp\left[-\frac{a}{2}\left(\frac{x^2}{L^2} + \frac{y^2}{L^2}\right)\right], \tag{4.11}$$

with L given by Eq. (4.3) and the dimensionless a given by

$$a = 256. \tag{4.12}$$

• Band-limited Gaussian spot (also referred to as the focal spot). This spot was obtained by propagation of a Gaussian spot to its focal plane behind a lens, by use of the original MASTER program. The Gaussian spot was of the form (4.11) but had width $L = 2.0 \times 10^{-3}$ m, and was propagated over a distance $d_0 = 300 \times 10^{-3}$ m to a lens with numerical aperture = 0.92 and focal distance $f = 1.95 \times 10^{-3}$ m. The propagated spot had a size far larger than the lens; therefore it was essentially as if a plane wave had been used. The resulting spot had a width that was comparable with that of the Gauss spot mentioned above but was different in shape. It still had a smooth form, however, and fell off to zero well within the calculational box.

• Bloc spot (Fig. 3): This field was given by the step function

$$f_{\mathrm{bloc}} = 1 \quad \text{if } \frac{x^2}{L^2} + \frac{x^2}{L^2} < \frac{100}{N_x^2},$$

$$f_{\mathrm{bloc}} = 0 \quad \text{elsewhere.} \tag{4.13}$$

Here L and N_x are given by Eqs. (4.3) and (4.5), respectively.

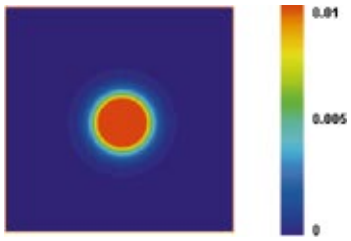

Fig. 2. Absolute value of the field for the Gauss spot for $i_z = 0$ ($z_0 = 0$) (the initial field).

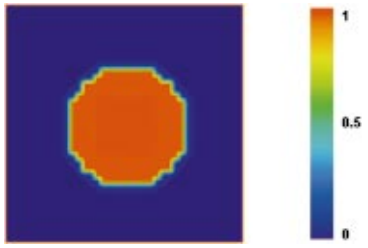

Fig. 3. Absolute value of the field for the bloc spot for $i_z = 0$ ($z_0 = 0$) (the initial field). Owing to the coarseness of the grid, the spot is not entirely circular.

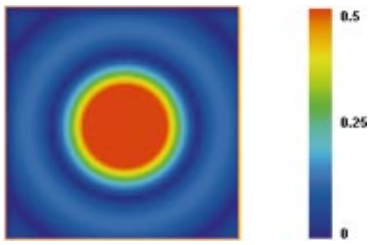

Fig. 4. Absolute value of the field for the Airy spot for $i_z = 0$ ($z_0 = 0$) (the initial field). The central spot is well contained within the box. The sidelobes fall partially outside the box.

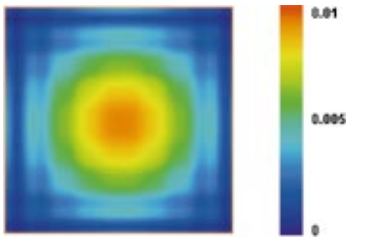

Fig. 5. Absolute value of the field for the Gauss spot for $i_z = 4$ ($z_0 = 64\lambda$) as calculated by the PWP method. In this case the spot shows aliasing effects when the edge of the box is approached.

- Airy spot (Fig. 4): This field was given by the Airy pattern

$$f_{\text{Airy}} = \frac{2J_1(\text{arg})}{\text{arg}},$$

$$\text{arg} = 10\pi N_x \left(\frac{x^2}{L^2} + \frac{y^2}{L^2} \right)^{1/2}, \tag{4.14}$$

where L and N_x are again given by Eqs. (4.3) and (4.5) and $J_1(x)$ is the Bessel function of the first kind.

The first two spots are smooth functions that fall off to zero well within the calculational box. For these spots, all methods are expected to work well within, at least, some range of z values. The bloc spot has a sharp edge. This can present problems when the spot is Fourier transformed because it will yield a long-range tail in k space, which will, after propagation, not be transformed backward correctly. Therefore we expect trouble in the case of the PWP method. The Airy spot has slowly decreasing oscillations for $x, y \to \infty$ and is, from its first oscillation onward, not entirely contained within the calculational box. This could cause difficulties for all methods of calculation.

C. Results of Calculations

Below we discuss the results of the PWP, FA, and RS calculations with the initial conditions mentioned in the previous sections. In particular, we discuss the behavior of $\epsilon_{\text{rel}}(x, y, z)$ of Eq. (4.10).

- Gauss and focal spot: In this case the (AMIRA) pictures of the absolute values of the spots showed good agreement for $i_z \leq 2$, i.e., $z_0 \leq 32\lambda$ in the case of the PWP–RS comparison. The agreement of the absolute values was good everywhere in the case of the FA–RS comparison. The pictures of the real and imaginary parts of the spots showed good agreement for $z_0 \leq 32\lambda$, in the cases of both the PWP–RS and the PWP–FA comparisons. For $i_z \geq 3$, i.e., $z_0 \geq 48\lambda$, the spots had expanded closely toward the edges of the calculational box. Here the PWP method showed some aliasing effects (Fig. 5), making the agreement with the RS method (Fig. 6) less good. The pictures for the real and imaginary parts go through some oscillations. The maxima of the oscillations were for $z_0 \geq 48\lambda$ less pronounced in the case of the FA than for the other two methods. This is probably due to round-off errors in the linear interpolation of this method. The values of $\epsilon_{\text{rel}}(x, y, z)$ were of $O(1\%)$ for $i_z \leq 2$ ($z_0 \leq 32\lambda$) in the case of the FA–RS comparison. This was also the case in the PWP–RS comparison. $\epsilon_{\text{rel}}(x, y, z)$ became, in both PWP–RS and FA–RS comparisons, of $O(3\%-5\%)$ when $i_z \geq 3$ ($z_0 \geq 48\lambda$). In the PWP–RS case this was due to the aliasing effects, and in the FA–RS case it was due to the round-off errors of the linear interpolation.
- Bloc spot: This shape was obtained only approximately (Fig. 3). The PWP method showed a clear "square-like" symmetry in the spot (Fig. 7). The spot was smoother and was more circular in the case of the RS method (Fig. 8). When we look at Eqs. (4.13), we see that the shape of this spot should be circular. However, since we used a rather coarse grid, this shape was obtained only approximately. In addition, this spot has a sharp edge, which can cause difficulties for a FFT. This was reflected in the shape of the spots: the PWP method showed a clear squarelike symmetry in the spot. The spot was smoother and more circular in the case of the RS

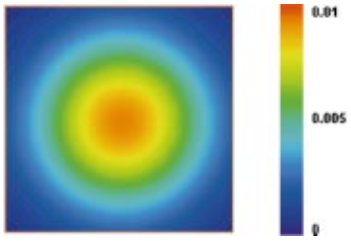

Fig. 6. Absolute value of the field for the Gauss spot for $i_z = 4$ ($z_0 = 64\lambda$) as calculated by the RS method. Although the spot nears the edge of the calculational box, the spot remains smooth.

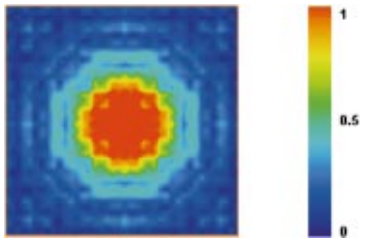

Fig. 7. Absolute value of the field for the block spot for $i_z = 2$ ($z_0 = 32\lambda$) calculated by the PWP method. The squarelike symmetry remains present.

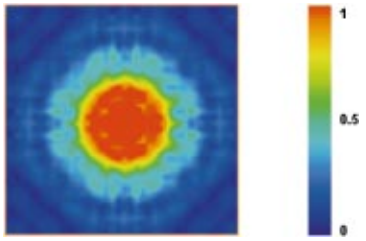

Fig. 8. Absolute value of the field for the bloc spot for $i_z = 2\,(z_0 = 32\lambda)$ calculated by the RS method. The spot has propagated and becomes smoother than the initial spot.

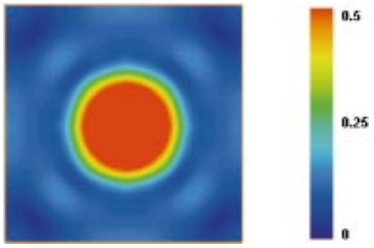

Fig. 9. Absolute value of the field for the Airy spot for $i_z = 2\,(z_0 = 32\lambda)$, calculated by the PWP method. The central spot agrees well with that of the RS method. For the sidelobes, a clear difference is visible.

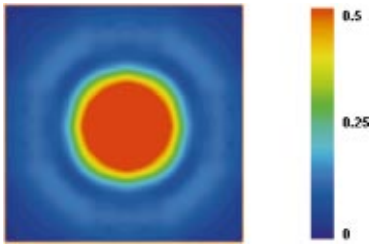

Fig. 10. Absolute value of the field for the Airy spot for $i_z = 2\,(z_0 = 32\lambda)$, calculated by the RS method. Owing to the cutoff in the calculational box, the sidelobes are not exactly circular.

method. Finally, the FA calculated the spot to be even smoother and more circular, for all $i_z \geqslant 1\,(z_0 \geqslant 16\lambda)$. The smoothness of the spot increased when z_0 increased. However, the squarelike symmetry remained visible for all z_0 in the case of the PWP method. ϵ_{rel} was, in the case of the PWP–RS comparison, of $O(3\%-8\%)$ and showed a regular pattern. In the case of the FA–RS comparison, ϵ_{rel} was also of $O(3\%-8\%)$. Here too a regular pattern was visible in the errors; however, in this case it was more circular.

- Airy spot: For this spot the agreement between the FA and RS methods was quite good for $i_z \leqslant 3\,(z_0 \leqslant 48\lambda)$. ϵ_{rel} was of $O(3\%)$ in this case. However, although the central spot was still circular, the outside lobes showed a more hexagonlike shape. Apparently, the fact that the outside lobes fall partially outside the calculational box had an effect here. For $i_z = 4$, ϵ_{rel} became of $O(5\%-7\%)$ near the center of the spot but remained of $O(3\%)$ at the outside. In the case of the PWP method (Fig. 9), the agreement with the RS method (Fig. 10) was quite good for the central spot, but for the sidelobes the agreement was very bad. ϵ_{rel} was of $O(1\%)$ at the center but became of $O(7\%)$ near the edges of the box. Apparently, the fact that the sidelobes of the spot are not well contained within the box also causes trouble in the backward FFT.

D. Discussion

We compared the propagation of four fields by using the three calculational methods mentioned above. It appeared that, when the fields are sufficiently well behaved, there is good agreement between the three methods in their common regions of validity. The RS method has an unlimited region of validity in all cases. Because the RS method is quite time consuming, the PWP and FA methods are still of interest in regions where they can be validly applied, however.

APPENDIX A: INTEGRATION OF DIFFRACTION INTEGRAL

In this appendix we describe how the integral

$$I = \int_0^{\sqrt{(\rho_1-\rho_2)/2}} \exp(iks^2) G(s)\mathrm{d}s \tag{A1}$$

[see Eq. (3.22)] can be expressed in a number of Fresnel integrals that can be used to determine I with great accuracy, even when the exponential term is oscillating fast.

The integration interval is divided into $\nu + 1$ points:

$$0 < s_0 < s_1 < \cdots < s_\nu = \sqrt{(\rho_1 - \rho_2)/2}. \tag{A2}$$

Now $G(s)$ is approximated on this grid by a cubic spline $\sigma(s)$, so that, for $s_{i-1} < s < s_i$,[13]

$$\sigma(s) = D_i(s - s_i)^3 + C_i(s - s_i)^2 + B_i(s - s_i) + A_i, \tag{A3}$$

for some numbers D_i, C_i, B_i, and A_i, $i = 1,\ldots,\nu$. The coefficients D_i and C_i can be expressed in terms of the second derivatives m_i, $i = 0,\ldots,\nu$, by requiring that on every interval i the second derivative of σ is continuous, i.e., that

$$\lim_{s\downarrow s_i} \sigma''(s) = \lim_{s\uparrow s_i} \sigma''(s) = m_i; \qquad i = 1,\ldots,\nu - 1, \tag{A4}$$

$$\lim_{s\downarrow 0} \sigma''(s) = m_0; \qquad \lim_{s\uparrow s_\nu} \sigma''(s) = m_\nu. \tag{A5}$$

Then the 2ν coefficients $B_1,\ldots,B_\nu$ and $A_1,\ldots,A_\nu$ can be expressed in terms of m_i and $G_i \equiv G(s_i)$, $i = 0,\ldots,\nu$, by requiring that on every interval i we have

$$\lim_{s\downarrow 0} \sigma(s) = G_0; \qquad \lim_{s\uparrow s_\nu} \sigma(s) = G_\nu, \tag{A6}$$

$$\lim_{s\downarrow s_i} \sigma(s) = \lim_{s\uparrow s_i} \sigma(s) = G_i, \qquad i = 1,\ldots,\nu - 1. \tag{A7}$$

Writing

$$h_i = s_i - s_{i-1}, \tag{A8}$$

we thus obtain for $s_i < s < s_{i-1}$[13]:

$$\sigma(s) = \frac{m_{i+1} - m_i}{6h_{i+1}} s^3 + \frac{m_i s_{i+1} - m_{i+1} s_i}{2h_{i+1}} s^2 + \left[-\frac{m_i s_{i+1}^2 - m_{i+1} s_i^2}{2h_{i+1}} + \frac{G_{i+1} - G_i}{h_{i+1}} - \frac{h_{i+1}(m_{i+1} - m_i)}{6} \right] s + \left[\frac{m_i s_{i+1}^3 - m_{i+1} s_i^3}{6h_{i+1}} - \frac{G_{i+1} - G_i}{h_{i+1}} s_i + \frac{h_{i+1}(m_{i+1} - m_i) s_i}{6} + G_i - \frac{h_{i+1}^2}{6} m_i \right]. \tag{A9}$$

The second derivatives m_i, $i = 0,\ldots,\nu$, are still unknown. We can determine them by solving a linear system, obtained by requiring that the first derivative of σ is continuous in the interior points $s_1,\ldots,s_{\nu-1}$ and by requiring that

$$\sigma'(s_0) = G'(s_0); \qquad \sigma'(s_\nu) = G'(s_\nu). \tag{A10}$$

Now define the integrals

$$I_i^\mu = \int_{s_i}^{s_{i+1}} \exp(iks^2) s^\mu \mathrm{d}s, \qquad \mu = 0,1,\ldots. \tag{A11}$$

Then, by use of $k = 2\pi/\lambda$,

$$I_i^0 = \frac{\sqrt{\lambda}}{2} \left\{ C\left(\frac{2s_{i+1}}{\sqrt{\lambda}}\right) - C\left(\frac{2s_i}{\sqrt{\lambda}}\right) + i\left[S\left(\frac{2s_{i+1}}{\sqrt{\lambda}}\right) - S\left(\frac{2s_i}{\sqrt{\lambda}}\right) \right] \right\}, \tag{A12}$$

$$I_i^1 = \frac{\exp(iks_{i+1}^2) - \exp(iks_i^2)}{2ik}, \tag{A13}$$

where $C(x)$ and $S(x)$ are the Fresnel integrals:

$$C(x) = \int_0^x \cos\frac{\pi t^2}{2} \mathrm{d}t, \qquad S(x) = \int_0^x \sin\frac{\pi t^2}{2} \mathrm{d}t. \tag{A14}$$

For $\mu = 2, 3,\ldots$, the integrals I_i^μ can be computed from the recursion

$$I_i^\mu = \frac{s_{i+1}^{\mu-1} \exp(iks_{i+1}^2) - s_i^{\mu-1} \exp(iks_i^2)}{2ik} - \frac{\mu - 1}{2ik} I_i^{\mu-2}. \tag{A15}$$

Hence we can approximate the integral in Eq. (A1) by

$$I = \sum_{i=0}^{\nu-1} \left\{ \frac{m_{i+1} - m_i}{6h_{i+1}} I_i^3 + \frac{m_i s_{i+1} - m_{i+1} s_i}{2h_{i+1}} I_i^2 + \left[-\frac{m_i s_{i+1}^2 - m_{i+1} s_i^2}{2h_{i+1}} + \frac{G_{i+1} - G_i}{h_{i+1}} - \frac{h_{i+1}(m_{i+1} - m_i)}{6} \right] I_i^1 + \left[\frac{m_i s_{i+1}^3 - m_{i+1} s_i^3}{6h_{i+1}} - \frac{G_{i+1} - G_i}{h_{i+1}} s_i + \frac{h_{i+1}(m_{i+1} - m_i) s_i}{6} + G_i - \frac{h_{i+1}^2}{6} m_i \right] I_i^0 \right\}. \tag{A16}$$

With this formalism the integrals over the ρ intervals in Eq. (3.9) can all be calculated. It is important to note that for every interval $(\rho_1(i), \rho_2(i))$ the number of grid points $\mu + 1$ used in the spline approximation is determined dynamically, on the basis of a criterion that takes account of the expected steepness of the functions $F_{\mathrm{left}}(s)$ and $F_{\mathrm{right}}(s)$ and furthermore takes account of the length $[\rho_2(i) - \rho_1(i)]/2$ of the integration interval in Eq. (3.9). Once chosen, the number of grid points is not further increased iteratively.

APPENDIX B: UNIFORM ERROR ESTIMATE OF THE SPLINE INTEGRAL

Consider the integral (3.22):

$$I = \int_0^{\sqrt{(\rho_1 - \rho_2)/2}} \exp(iks^2) G(s) \mathrm{d}s. \tag{B1}$$

We will show in this appendix that the method we use to approximate this integral, i.e., the approximation of $G(s)$ by a spline fit $\sigma(s)$ described in Appendix A, will give a result that is a good approximation even when $k \to \infty$. To this end, we will consider the integral

$$I_G \equiv \int_0^1 \exp(iks^2) G(s) \mathrm{d}s \tag{B2}$$

(we choose the upper bound of the integral equal to 1 for convenience) and give an error estimate of the spline approximation of this integral:

$$I_\sigma \equiv \int_0^1 \exp(iks^2) \sigma(s) \mathrm{d}s, \tag{B3}$$

when $k \to \infty$.

First, suppose that we have a function $f(s)$ on $[0, 1]$ with $f(0) = 0$. Then

$$I_f \equiv \int_0^1 \exp(iks^2) f(s)\mathrm{d}s = \frac{1}{2ik}\int_0^1 \frac{f(s)}{s}\frac{\mathrm{d}}{\mathrm{d}s}\exp(iks^2)\mathrm{d}s$$
$$= \frac{1}{2ik}\frac{f(s)}{s}\exp(iks^2)\Big|_0^1 - \frac{1}{2ik}\int_0^1\left[\frac{f'(s)}{s} - \frac{f(s)}{s^2}\right]\exp(iks^2)\mathrm{d}s$$
$$= \frac{1}{2ik}[f(1)\exp(ik) - f'(0)] + \frac{1}{2ik}\int_0^1\left[\frac{f(s) - sf'(s)}{s^2}\right]\exp(iks^2)\mathrm{d}s. \quad \text{(B4)}$$

By application of this result to $f(s) = g(s) - g(0)$, it follows that for arbitrary $g(s)$ with $g(0) \neq 0$

$$I_g \equiv \int_0^1 \exp(iks^2) g(s)\mathrm{d}s$$
$$= g(0)\int_0^1 \exp(iks^2)\mathrm{d}s + \int_0^1 \exp(iks^2)[g(s) - g(0)]\mathrm{d}s$$
$$= g(0)\left(\frac{\pi}{2k}\right)^{1/2}\left\{C\left[\left(\frac{2k}{\pi}\right)^{1/2}\right] + iS\left[\left(\frac{2k}{\pi}\right)^{1/2}\right]\right\} + \frac{g(1) - g(0)}{2ik}\exp(ik) - \frac{g'(0)}{2ik} + \frac{1}{2ik}\int_0^1 \frac{g(s) - g(0) - sg'(s)}{s^2}\exp(iks^2)\mathrm{d}s, \quad \text{(B5)}$$

where $C(x)$ and $S(x)$ are the Fresnel integrals as given in Eqs. (A14). We have, for $k \to \infty$,

$$C\left[\left(\frac{2k}{\pi}\right)^{1/2}\right] + iS\left[\left(\frac{2k}{\pi}\right)^{1/2}\right] \sim \frac{\exp[i(\pi/4)]}{\sqrt{2}} - \frac{(1+i)}{\sqrt{2\pi k}}\exp(ik) + O\left(\frac{1}{k^{3/2}}\right). \quad \text{(B6)}$$

We now apply Eqs. (B4) and (B5) to

$$g(s) = G(s), \quad \text{(B7)}$$

$$f(s) = G(s) - \sigma(s), \quad \text{(B8)}$$

where $\sigma(s)$ is a cubic spline fit to $G(s)$. Let h be the maximum distance between the grid points. Then we have for the spline fit[13]:

$$\max_{0\leqslant s\leqslant 1}|f^{(j)}(s)| \leqslant C\|G\|_{4,\infty}h^{4-j}; \qquad j = 0,1,2,3, \quad \text{(B9)}$$

where

$$\|G\|_{4,\infty} = \max_{0\leqslant s\leqslant 1}[|G(s)| + |G^{(1)}(s)| + |G^{(2)}(s)| + |G^{(3)}(s)| + |G^{(4)}(s)|]. \quad \text{(B10)}$$

Therefore we have, for the absolute error,

$$|I_G - I_\sigma| = |I_{G-\sigma}| = \left|\int_0^1 \exp(iks^2)[G(s) - \sigma(s)]\mathrm{d}s\right| \leqslant \|G - \sigma\|_{4,\infty}h^4, \quad \text{(B11)}$$

for all k. However, for $k \to \infty$, we have $I_G \to 0$. Therefore, for $k \to \infty$, the relative error should be considered. We have, for $k \to \infty$,

$$I_G \sim \frac{G(0)}{2}\left(\frac{\pi}{k}\right)^{1/2}\exp\left(i\frac{\pi}{4}\right) + O\left(\frac{1}{k}\right), \quad \text{(B12)}$$

$$I_{G-\sigma} \sim \frac{\sigma'(0) - G'(0)}{2ik} + \frac{1}{2ik}\int_0^1 \frac{G(s) - \sigma(s) - s[G'(s) - \sigma'(s)]}{s^2} \times \exp(iks^2)\mathrm{d}s. \quad \text{(B13)}$$

Now apply the mean value theorem to the function

$$p(s) = G(s) - \sigma(s) - s[G'(s) - \sigma'(s)]. \quad \text{(B14)}$$

Note that $p(0) = 0$. Then it follows that for every $s \in [0,1)$ there is a $t(s) \in [0,s]$ such that

$$G(s) - \sigma(s) - s[G'(s) - \sigma'(s)] = [\sigma''(t(s)) - G''(t(s))]t(s)s. \quad \text{(B15)}$$

Hence, from Eq. (B10),

$$\frac{|G(s) - \sigma(s) - s[G'(s) - \sigma'(s)]|}{s^2} \leqslant C\|G\|_{4,\infty}h^2. \quad \text{(B16)}$$

Because expression (B9) implies that

$$|G'(0) - \sigma'(0)| \leqslant C\|G\|_{4,\infty}h^3, \quad \text{(B17)}$$

it follows from expressions (B13), (B15), and (B16) that

$$|I_{G-\sigma}| \leqslant C\|G\|_{4,\infty}\frac{h^2(h+1)}{2k}. \quad \text{(B18)}$$

Hence, for $k \to \infty$,

$$\frac{|I_{G-\sigma}|}{|I_G|} \sim \frac{C\|G\|_{4,\infty}}{\sqrt{\pi}G(0)}\frac{h^2(h+1)}{\sqrt{k}}; \qquad k \to \infty. \quad \text{(B19)}$$

Thus approximating I_G by I_σ, with σ a cubic spline fit, gives an error that is asymptotically uniformly bounded with respect to k.

The author's e-mail address is jan.a.c.veerman @philips.com.

REFERENCES

1. J. J. Stamnes, *Waves in Focal Regions* (Institute of Physics, Bristol, UK, 1986).
2. M. Mansuripur, "Distribution of light at and near the focus of high-numerical-aperture objectives," J. Opt. Soc. Am. A **3**, 2086–2093 (1986).
3. P. Török, S. J. Hewlett, and P. Varga, "On the series expansion of high-aperture, vectorial diffraction integrals," J. Mod. Opt. **44**, 493–503 (1997).
4. C. J. R. Sheppard and P. Török, "Efficient calculation of electromagnetic diffraction in optical systems using a multipole expansion," J. Mod. Opt. **44**, 803–818 (1997).
5. S. Stallinga, "Axial birefringence in high-numerical-aperture optical systems and the light distribution close to focus," J. Opt. Soc. Am. A **18**, 2846–2859 (2001).
6. P. L. M. Put, H. P. Urbach, R. D. Morton, and J. J. Rusch, "Resolution limit of optical disc mastering," Jpn. J. Appl. Phys. **36**, 539–548 (1997).
7. J. M. Brok and H. P. Urbach, "Rigorous model of the scattering of a focused spot by a grating and its application in optical recording," J. Opt. Soc. Am. A **20**, 256–272 (2003).
8. T. Gravelsaeter and J. J. Stamnes, "Diffraction by circular apertures. 1: Method of linear phase and amplitude approximation," Appl. Opt. **21**, 3644–3651 (1982).
9. J. J. Stamnes, "Hybrid integration technique for efficient and accurate computation of diffraction integrals," J. Opt. Soc. Am. A **6**, 1330–1342 (1989).
10. L. d'Arcio, J. J. M. Braat, and H. J. Frankema, "Numerical evaluation of diffraction integrals for apertures of complicated shape," J. Opt. Soc. Am. A **11**, 2664–2674 (1994).
11. See, e.g., J. W. Goodman, *Introduction to Fourier Optics* (McGraw-Hill, New York, 1996).
12. J. A. C. Veerman, "Evaluation of spot propagation calculations in 1D," Internal Note, 2001.
13. J. Stoer, *Einführung in die Numerische Mathematik* (Springer Verlag, Berlin, 1979).
14. See http://www.tgs.com.