A Wigner-based ray-tracing method for imaging simulations

B.M. Mout\textsuperscript{a,b}, M. Wick\textsuperscript{b}, F. Bociort\textsuperscript{a} and H.P. Urbach\textsuperscript{a}

\textsuperscript{a}Delft University of Technology, Lorentzweg 1, 2628 CJ Delft, Netherlands; \textsuperscript{b}Carl Zeiss AG, Carl-Zeiss-Straße 22, 73447 Oberkochen, Germany

ABSTRACT

The Wigner Distribution Function (WDF) forms an alternative representation of the optical field. It can be a valuable tool for understanding and classifying optical systems. Furthermore, it possesses properties that make it suitable for optical simulations: both the intensity and the angular spectrum can be easily obtained from the WDF and the WDF remains constant along the paths of paraxial geometrical rays. In this study we use these properties by implementing a numerical Wigner-Based Ray-Tracing method (WBRT) to simulate diffraction effects at apertures in free-space and in imaging systems. Both paraxial and non-paraxial systems are considered and the results are compared with numerical implementations of the Rayleigh-Sommerfeld and Fresnel diffraction integrals to investigate the limits of the applicability of this approach. The results of the different methods are in good agreement when simulating free-space diffraction or calculating point spread functions (PSFs) for aberration-free imaging systems, even at numerical apertures exceeding the paraxial regime. For imaging systems with aberrations, the PSFs of WBRT diverge from the results using diffraction integrals. For larger aberrations WBRT predicts negative intensities, suggesting that this model is unable to deal with aberrations.

Keywords: Wigner Distribution Function, ray-tracing, diffraction, optical simulation

1. INTRODUCTION

The Wigner Distribution Function (WDF) is a transformation of a (complex) field proposed by Eugene Wigner in 1932. It was introduced into optics in 1968 in a paper that linked the Fourier-optical results for the flux of partially coherent light with the formalism of radiometry. A decade later it was demonstrated that the WDF before and after propagation through a paraxial optical system can be related by a coordinate transformation that is exactly the same as the one describing the paths of paraxial rays. Therefore the WDF remains constant along paths followed by paraxial rays. Using this property, a tool was developed to incorporate diffraction effects into computer rendering. Simulations of diffraction due to surfaces (e.g. compact discs) and apertures were presented, but the majority of results were not compared to those of methods commonly used within the optics community.

These publications motivate the present investigation of a simulation method employing the WDF and ray-tracing, to capture diffraction effects. Our goal is to determine the limits of the applicability of such a method for simulating imaging systems.

In Sec. 2 we give a brief overview of the WDF, Sec. 3 describes WBRT and Sec. 4 the results we obtained using this method.

2. WIGNER DISTRIBUTION FUNCTION

Since the simulation method used in this research is based on the WDF, we begin with a brief description of this function and its use in optics. For a more complete overview and proofs of the properties mentioned in this section, the reader is referred to the overview paper by Alonso.

The WDF is a bilinear function given by the following transformation of a scalar field:

\begin{equation}
W(x, \xi) = \int \int T \left( x + \frac{x'}{2} \right) T^* \left( x - \frac{x'}{2} \right) e^{-i \xi \cdot x'} d^2 x'
\end{equation}

where $W$ represents the WDF, $x = (x_1, x_2)$ is a point in a plane in the spatial domain, $\xi = (\xi_1, \xi_2)$ is a point in the two-dimensional reciprocal space, and $T$ is the scalar electric field. All integrals are from negative to positive.
infinity unless specified otherwise. The WDF of some commonly used electric fields are listed in Table 1. Note that all WDFs in Table 1 are real functions. This is a general property of the WDF. The listed functions are also positive, but this is a specific property of the WDF of exponential functions up to second order. All WDFs of higher-order functions have regions with negative values.7 Due to these negative values the WDF is able to account for interference effects.

Table 1: The WDF of some commonly used electric fields. The propagation distance along the optical axis is represented by \( z \), the wavenumber by \( k \) and the two-dimensional spatial frequencies by the vector \( \mathbf{k} \).

<table>
<thead>
<tr>
<th>Description</th>
<th>Electric field</th>
<th>Wigner Distribution Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plane wave</td>
<td>( T(x) = Ae^{ik\cdot x} )</td>
<td>( W(x, \xi) = (2\pi)^2</td>
</tr>
<tr>
<td>Point source</td>
<td>( T(x) = \delta(x - x_0) )</td>
<td>( W(x, \xi) =</td>
</tr>
<tr>
<td>Paraxial spherical wave</td>
<td>( T(x; z) = A \frac{\exp\left[ik\left(z + \frac{</td>
<td>x</td>
</tr>
<tr>
<td>Gaussian beam</td>
<td>( T(x) = A \exp\left[-\left(\frac{</td>
<td>x</td>
</tr>
</tbody>
</table>

A useful property of the WDF is that its marginals give the angular spectrum of plane waves:

\[
|\tilde{T}(\xi)|^2 = \iint W(x, \xi)d^2x, \tag{2}
\]

and the intensity in a plane:

\[
|T(x)|^2 = \frac{1}{(2\pi)^2} \iint W(x, \xi)d^2\xi. \tag{3}
\]

In linear optics, the electric field after propagation through an optical system, can be calculated using an integral kernel:

\[
T_o(y) = \iint T_i(x)h(x, y)d^2x. \tag{4}
\]

where \( y \) are the spatial coordinates in the output plane and the indices \( o \) and \( i \) indicate the output and input plane. When the kernel \( h \) is a second order phase function, the WDF of the output field \( T_o \) can be obtained by a coordinate shift of the WDF of the input field. The effect of the Fresnel propagator8

\[
h(x, y) = \frac{e^{i\frac{\lambda z}{2\pi}x^2}}{1\lambda z} e^{i\frac{\lambda z}{2\pi}|x-y|^2}d^2x, \tag{5}\]

is for instance given by

\[
W_o(y, \eta) = W_i(y - \eta \frac{\lambda z}{2\pi}, \eta). \tag{6}\]

The spatial frequency in the output plane is represented by \( \eta \). The thin paraxial lens described by

\[
h(y, x) = e^{-i\frac{\lambda z}{2\pi}x^2}\delta(y - x), \tag{7}\]

leads to the following WDF in the output plane:

\[
W_o(y, \eta) = W_i(y, \eta + \frac{k}{f}y). \tag{8}\]
The coordinate transformations of the WDF due to Fresnel propagation and thin lenses are the same as the coordinate transformations of paraxial rays induced by propagation and lenses. All that is needed is the relation between the directional cosines in the input and output plane, \( p \) and \( q \), and the corresponding spatial frequency variables:

\[
\begin{align*}
    p &= \frac{\xi}{k}, \\
    q &= \frac{\eta}{k}.
\end{align*}
\]

More general, phase kernels of second order (as described by the Collins integral\(^9\)) perform a coordinate transformation on the WDF that is identical to the coordinate transformation of the ray matrix formalism on paraxial rays.

The effect of apertures and masks, within the thin element approximation, on the electric field is given by:

\[
T_o(x) = B(x)T_in(x). \tag{10}
\]

The effect on the WDF is a multiplication in spatial and convolution in frequency domain of the input WDF and the WDF of the mask (\( W_B \)):

\[
W_o(y, \eta) = \frac{1}{(2\pi)^2} \int \int W_{in}(y, \xi)W_B(y, \eta - \xi)d^2\xi. \tag{11}
\]

For a rectangular aperture of width \( 2w \) and height \( 2v \) the WDF is given by

\[
W_B(x, \xi) = \begin{cases} 
    \frac{4}{\xi_1\xi_2} \sin \left[ \xi_1 2(w - |x_1|) \right] \sin \left[ \xi_2 2(v - |x_2|) \right] & \text{if } |x_1| \leq w \text{ and } |x_2| \leq v \\
    0 & \text{otherwise},
\end{cases} \tag{12}
\]

for a circular aperture of radius \( a \) the WDF is\(^10\)

\[
W \left[ \begin{pmatrix} x_1 \\ 0 \end{pmatrix}, \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} \right] = 16a^2 R \left[ \exp[-i2\xi_1 x_1] \int_{x_1/a}^1 \frac{\sin \left[ 2a\xi_2 \sqrt{1 - s^2} \right]}{2a\xi_2} \exp[i2a\xi_1 s] ds \right], \tag{13}
\]

where \( R \) denotes the real part of the argument.

With these basic properties of the WDF in place, we can now develop a Wigner-based simulation method.

### 3. METHOD

The key idea behind the simulation method we use is to transport the WDF over geometrical rays by ray-tracing. The similarity between the coordinate transformations of the WDF and of geometrical rays upon passing through a paraxial optical system justifies this approach for paraxial systems. In Ref. 11 it is stated that the effect of third-order aberrations on the WDF can be modeled by a coordinate transformation as well. This suggests that our approach could also work for systems with aberrations.

For all simulations in this research the WDF of the input field is calculated analytically. This is possible for some commonly used electric fields like plane waves, point sources and Gaussian beams (see Table 1). For general fields numerical calculations of the WDF are required.

Within the four-dimensional phase-space of the input plane, a domain of interest is chosen based on two criteria: First of all the WDF should have a significant non-zero value. Secondly, the rays associated with the phase-space coordinates should pass through the aperture of the optical system.

Within this domain of interest, a random phase-space coordinate is chosen using a Monte-Carlo procedure. This coordinate is related to a geometrical optical ray with the same spatial coordinates, the directional cosines
determined by Eq. (9) and a weight proportional to the WDF at those coordinates. The ray is transported through the optical system up to the aperture using a geometrical optical ray tracer.

The effect of an aperture on the WDF is a multiplication in spatial and convolution in frequency coordinates of the input WDF and the WDF of the aperture (Eq. (11)). A ray can be seen as a double delta function in frequency and spatial coordinates. Using Eq. (11) the WDF after an aperture incident by a ray is:

\[
W_o(y, \eta) = \frac{1}{(2\pi)^2} \int \int A_r \delta(y - y_r) \delta(\xi - \xi_r) W_B(y, \eta - \xi) d^2 \xi.
\]

\[
= \frac{A_r}{(2\pi)^2} \delta(y - y_r) W_B(y, \eta - \xi_r) \tag{14}
\]

where \(y_r\) and \(\xi_r\) are the phase-space coordinates of the ray and \(A_r\) is the weight of the ray. The WDF after the aperture is still confined to one spatial coordinate, but has been spread out over the frequency coordinates. This WDF is sampled using a Monte-Carlo procedure similar to the one used for the input plane. The result is a set of rays starting from the same point, going in different directions with a weight determined by Eq. (14).

The ray-like double delta function as used in Eq. (14) is not a WDF of a physical scalar electric field. Therefore one cannot view the simulation procedure described above as breaking up the problem into subproblems with elementary solutions, as is for instance the case when propagating a field using a plane wave decomposition. The WDF after the aperture is calculated by performing a coordinate transformation of the input WDF (using ray-tracing) followed by a convolution with the WDF of the aperture (by generating new rays). These two steps are evaluated using a Monte-Carlo algorithm, and every sampling point of this Monte-Carlo integral consists of a ray-like double delta function. The resulting WDF can therefore correspond to a physical electric field, even if the individual sampling points are unphysical.

After the aperture the WDF is further transported using ray-tracing up to the output plane. In the output plane the weight of a ray, corresponding to the value of the WDF at a specific phase-space coordinate, is added to value of the pixel it intersects with. This is a discrete, Monte-Carlo evaluation of the frequency marginal shown in Eq. (3), combined with an integration over the pixel area. The result is the integral of the intensity over the pixel. A proof of the analytical equivalence of this procedure to the Fresnel integral is given in Appendix A.

The procedure described in this section is continued until the intensity distribution in the output plane converges. The process can easily be distributed over multiple processors or even multiple computers.

All WBRT simulations for this publications are performed using MATLAB.

4. RESULTS

We use Wigner-Based Ray-Tracing (WBRT) to simulate free-space propagation with and without an aperture and to simulate the propagation through simple optical systems consisting of a singlet and an aperture. The results are compared to the results from the analytical solution for the propagation of a paraxial Gaussian beam, the Fresnel diffraction integral and the Rayleigh-Sommerfeld exit pupil diffraction integral of a ray tracer respectively.

4.1 Free-Space Propagation

In this section we present the results of WBRT for systems without lenses.

4.1.1 Gaussian beam propagation

The plane at which the Gaussian beam has the narrowest beam waist is chosen as the \(z = 0\) plane. In this plane the electric field of the Gaussian beam is given by:

\[
T(x) = A \exp \left[ - \left( \frac{|x|}{w_0} \right)^2 \right], \tag{15}
\]
Figure 1: A Gaussian beam with a minimal waist of 1 mm and a wavelength of 600 nm after propagating over 100 m calculated using WBRT and compared to the results obtained using Eq. (17).

where \( \mathbf{x} = (x_1, x_2) \) is the position in the plane and \( w_0 \) is the waist size of the beam. The WDF of this field is:

\[
W(\mathbf{x}, \xi) = |A|^2 \frac{2\sqrt{2}}{\sqrt{\pi}} w_0^2 \exp \left[ -2 \frac{|\mathbf{x}|^2}{w_0^2} \right] \exp \left[ -\frac{w_0^2}{2} |\xi|^2 \right].
\]  

(16)

The intensity of the field after propagating over a distance \( z \) is given by:

\[
I(\mathbf{x}, z) = A^2 \left( \frac{w_0}{w(z)} \right)^2 \exp \left[ -\frac{2|\mathbf{x}|^2}{w(z)^2} \right]
\]  

(17)

with

\[
w(z) = w_0 \sqrt{1 + \left( \frac{z \lambda}{\pi w_0^2} \right)^2}
\]  

(18)

the beam waist size. The WDF of the Gaussian beam at \( z = 0 \) falls off quickly for coordinates \((\mathbf{x}, \xi)\) further away from the origin in phase-space. Therefore only the central part of the phase-space needs to be sampled.

The results for a WBRT simulation of the propagation over 100 m of a Gaussian beam with a minimal waist \((w_0)\) of 1 mm and a wavelength of 600 nm are shown in Figure 1. The results obtained using Wigner-Based Ray-Tracing are in good agreement with the analytical result provided by Eq. (17).

4.1.2 Diffraction at an aperture in free-space

We use Wigner-Based Ray-Tracing to simulate diffraction at circular and rectangular apertures for point sources and plane waves. The results are compared to those of a numerical implementation of the Fresnel diffraction integral.

The Fresnel integral is given by

\[
T_o(y) = \frac{\exp \left[ ikz \right]}{i\lambda z} \iint T_{in}(\mathbf{x}) \exp \left[ \frac{ik}{2z} |y - \mathbf{x}|^2 \right] d^2x,
\]  

(19)

where \( z \) is the distance from input to the output plane. For the Kirchhoff boundary conditions (used throughout this research) and for

\[
\frac{|y - \mathbf{x}|^4}{8z^3} \ll \lambda,
\]  

(20)
it is considered to be an accurate model.

One of the simulated systems consists of a point source followed by a square aperture, the other by a plane wave incident on a circular aperture. The parameters of the two simulated systems are listed in Table 2.

Table 2: The parameters of the systems used for free-space diffraction simulations.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>system 1</th>
<th>system 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>source</td>
<td>point source</td>
<td>plane wave</td>
</tr>
<tr>
<td>distance to aperture</td>
<td>500 mm</td>
<td>-</td>
</tr>
<tr>
<td>aperture shape</td>
<td>square</td>
<td>circular</td>
</tr>
<tr>
<td>aperture size</td>
<td>1x1 mm</td>
<td>1 mm (radius)</td>
</tr>
<tr>
<td>propagation after aperture</td>
<td>1000 mm</td>
<td>100 mm</td>
</tr>
<tr>
<td>wavelength</td>
<td>600 nm</td>
<td>600 nm</td>
</tr>
</tbody>
</table>

The results of the simulations are presented in Figure 2 and Figure 3. Both results are in good agreement with the results obtained using the Fresnel diffraction integral.

### 4.2 Singlet

In order to see if WBRT can be applied to optical systems with lenses, we simulate a variety of systems consisting of a lens directly followed by a circular aperture. By varying the aperture radius between 2 mm and 150 mm, the numerical aperture varies between $4 \cdot 10^{-3}$ and 0.3 and the Fresnel number between 10 and $10^5$ (the propagation distance after the lens was always around 0.5 m). Both spherical and aspherical singlets are used.

The results from WBRT are compared to those of the Rayleigh-Sommerfeld exit pupil diffraction integral (RSED) implemented in a ray-tracing based optics simulation tool. Within this method a discretization of the electric field in the exit pupil is calculated using ray-tracing. This field is then propagated to the image plane by using the Rayleigh-Sommerfeld diffraction integral. Exit pupil diffraction integrals like the RSED are commonly used to calculate PSFs.

Some typical results of WBRT are shown in Fig. 4, 5, 6 and 7. Apart from the results obtained using WBRT, the plots also show results from the RSED for the optical system and a corresponding ideal (aberration-free)
Figure 3: The far-field diffraction pattern of a plane wave passing through a circular aperture, simulated using WBRT. The cross-section shows the result from a numerical Fresnel integral as a reference. Intensities are normalized with respect to the maximum intensity. The system parameters are listed in Table 2.

Figure 4: The PSF of a singlet with an aperture radius of 10 mm and a numerical aperture of 0.02 calculated using WBRT. The cross-section also shows the cross sections of PSFs calculated using the Rayleigh-Sommerfeld exit pupil diffraction integral (RSED) for the singlet and a corresponding aberration-free system.

system. The results from WBRT correspond to those from the RSED for systems where the RSED of the optical system and the ideal system overlap (see Fig. 4 and 6a). When the PSFs of the optical system and the ideal system are different (indicating the presence of aberrations), WBRT does not reproduce the results of the RSED (see Fig. 5 and 6b). Moreover, these results frequently included regions of negative intensity. When simulating an aberration-free optical system using WBRT, the results match those of the RSED for numerical apertures up to 0.3 and slowly diverge from them when increasing the aperture further (see Fig. 7). Note that for the three presented systems, the results of WBRT start to diverge from those of the RSED at completely different numerical apertures and Fresnel numbers.

These results suggest that WBRT can reproduce the PSFs of the RSED if no aberrations are present and the numerical aperture is below 0.3.
Figure 5: Cross-sections (averaged) of the PSFs of singlets with aperture radii of 17 mm and 20 mm and a numerical aperture of 0.04. The results from WBRT and the RSED are shown, as well as the RSED results for the corresponding aberration-free systems.

Figure 6: Cross-sections (averaged) of the PSFs of an aspherical singlet with aperture radii of 100 and 150 mm. The numerical apertures are 0.20 and 0.29. The results from WBRT and the RSED are shown, as well as the RSED results for the corresponding aberration-free systems.
Figure 7: Cross-sections (averaged) of the PSFs of a spherical wave converging on an aperture with a radius of 150 mm and 200 mm respectively calculated using WBRT and the RSED. The numerical apertures of the systems are 0.29 and 0.37.
5. DISCUSSION

We use a Wigner-based method where lenses and propagation are represented by their geometric optical coordinate changes (WBRT) and investigate whether it is an adequate tool for simulating paraxial and non-paraxial systems. The results show that this is the case in a distinct regime. The results of WBRT are in good agreement with those of the reference methods for paraxial free-space propagation and for PSF calculations of aberration-free singlets with numerical apertures of 0.2 or lower. For a spherical wave converging on an aperture, this was the case up to a numerical aperture of 0.3. This numerical aperture is far beyond the paraxial domain. However, as soon as aberrations influence the PSF, the results of WBRT are no longer in agreement with those of the exit pupil diffraction method. Furthermore, WBRT frequently predicts unphysical negative intensities. This suggests that WBRT may be suitable for simulating non-paraxial propagation, but is unable to account for both aberrations and diffraction.

Within its domain of validity, several applications of WBRT exist: It could be used as an alternative to numerical implementations of the Collins (or Fresnel) integral for paraxial optical simulations. It might also be applied in computer graphics, where ray-tracing is used to render two-dimensional images of three-dimensional scenes, provided one uses aberration-free camera models.

Our results show borders of the domain of validity of WBRT, within which it can be used as a tool for diffraction calculations.

APPENDIX A. WBRT AND THE FRESNEL INTEGRAL

In this section we show that the intensity, after propagation over a distance $z$, predicted by WBRT and the Fresnel integral are analytically equivalent. The starting point for the derivation is Eq. (3) which gives the intensity distribution from the WDF in the output plane:

$$ I(y) = \frac{1}{(2\pi)^2} \int\int W_o(y, \eta) d^2\eta, $$

where $W_o$ is the WDF in the output plane. Since we use rays to transport the WDF, $W_o$ at position $y$, is determined by the input WDF ($W_i$) at position $x = y - \eta^z/k$. This can be expressed using a delta function. Incorporating this delta function and using the definition of the WDF, we can write:

$$ I(y) = \frac{1}{(2\pi)^2} \int\int\int W_i(x, \eta) \delta \left[ x + \frac{z^z}{k} - y \right] d^2xd^2\eta, $$

$$ = \frac{1}{(2\pi)^2} \int\int\int T \left( x + \frac{x'}{2} \right) T^* \left( x - \frac{x'}{2} \right) \exp \left[ -i\eta \cdot x' \right] \delta \left[ x + \frac{\eta^z}{k} - y \right] d^2x'd^2\eta, $$

$$ = \frac{1}{(2\pi)^2} \int\int \left[ \frac{k}{z} \right]^2 T \left( x + \frac{x'}{2} \right) T^* \left( x - \frac{x'}{2} \right) \exp \left[ -i\frac{k}{z} (y - x) \cdot x' \right] d^2x'd x. $$

Performing a change of variables to

$$ x_+ = x + \frac{x'}{2}, $$

$$ x_- = x - \frac{x'}{2}, $$

results in

$$ I(y) = \frac{1}{(2\pi)^2} \int\int \left[ \frac{k}{z} \right]^2 T (x_+) T^* (x_-) \exp \left[ -i\frac{k}{z} (y - 1/2(x_+ + x_-)) \cdot (x_+ - x_-) \right] d^2x_+ d^2x_-.$$

Note that one can rewrite

$$ (2y - (x_+ + x_-)) \cdot (x_+ - x_-) = |y - x_-|^2 - |y - x_+|^2.$$
Using this and writing out the wave number \(k\) gives

\[
I(y) = \frac{1}{(2\pi)^2} \int \int \int \int \left| T(x_+) T^*(x_-) \exp \left[ -i \frac{\pi}{\lambda z} (|y - x_+|^2 - |y - x_-|^2) \right] \right| d^2x_+ d^2x_-,
\]

(27)

\[
= \int \int \frac{1}{\lambda z} T(x_+) \exp \left[ i \frac{\pi}{\lambda z} |y - x_+|^2 \right] d^2x_+ \\
\int \int \frac{1}{\lambda z} T^*(x_-) \exp \left[ -i \frac{\pi}{\lambda z} |y - x_-|^2 \right] d^2x_-,
\]

(28)

\[
= \left| \int \int \frac{1}{\lambda z} T(x_+) \exp \left[ i \frac{\pi}{\lambda z} |y - x_+|^2 \right] d^2x_+ \right|^2.
\]

(29)

Comparing this result to Eq. (19) shows that this is the intensity of the electric field \(T\), after propagation over a distance \(z\), using the Fresnel integral.

**ACKNOWLEDGMENTS**

The research leading to these results has received funding from the People Programme (Marie Curie Actions) of the European Union’s Seventh Framework Programme FP7/2007-2013/ under REA grant agreement number 608082.

**REFERENCES**