On the Numerical Simulation of Diffraction

Julien

Contents

1 Abstract

The theory of the Rayleigh-Sommerfeld, Fresnel and Fraunhofer diffraction integrals is summarized and the regions under which each approximation is valid are derived. Afterwards, numerical implementations of those formulae are proposed and discussed with attention paid to the sampling considerations. Implementation of the Rayleigh-Sommerfeld diffraction using numerical integration technique of Simpson's $\frac{1}{3}^{rd}$ rule, Gauss-Laguerre quadrature and Monte Carlo integration is discussed. Results obtained using Monte Carlo integration in the case of a square aperture are shown in the "very" near field and the "far-field" and the latter is compared with exact results of the Fraunhofer diffraction integral. Numerical implementations of the Fresnel and Fraunhofer diffraction integrals are also described and results are shown for a DFT based evaluation of the Fraunhofer diffraction integral for a square aperture.

2 Introduction

Numerical implementations of diffraction are studied due to the difficulty of obtaining exact solutions to diffraction formulas in many cases and particularly in real world application ([2], [3, ix]). Because diffraction occurs due to lateral confinement of a wave, diffraction theory has applications to many fields which deal with wave propagation and is not limited to the field of optics [1, pp.31-32]. It could be argued that the advent of computational physics has resulted in a paradigm shift in the scientific method [14] and the use of simulation is now common practice in the commercial world to aid in product design [12] and in academia to design experiments [13]. For this reason, when implementing any physical phenomena numerically it is important to understand the implications of the digitization of an algorithm as well as its validity when applied in various scenarios.

This paper studies numerical implementations of optical scalar diffraction; first the approximations inherent in the formulation of the theory of diffraction are discussed, then, approaches to numerical implementation of these formulae are summarized and discussed.

3 Diffraction Regimes

The mathematical theory of scalar diffraction can be formulated by the use of Maxwell's equations, which can be used to derive the Helmholtz equation, which. along with Green's theorem can be used to obtain the Fresnel-Kirchhoff and Rayleigh-Sommerfeld Diffraction formulae. The formulations of the theory of diffraction along with various approximations, lead to the near-field and far-field approximations and are theoretically sufficient to model scalar diffraction at any distance [1]. Knowing the situations in which each formulation in valid can allow for considerable simplifications to be made. The following diagram summarizes the regions of validity of the diffraction approximations discussed in this paper:

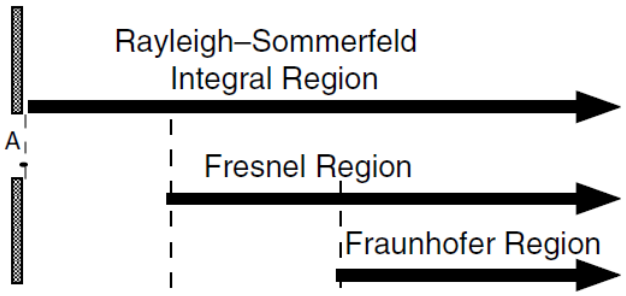

Figure 1: The three diffraction regions [3, p.64].

The near-field can be thought of as the region between the input plane and the Fraunhofer Region (Eq. 13), while the far-field can be considered as the region in which the Fraunhofer approximation can be used [3, p. 63]. As the following subsections will explain, the precise definitions of these regions for a given problem will depend on the optical wavelength λ of the wave, the sizes of the input aperture and the observation plane, as well as the propagation distance between the input plane and the observation plane [3]. A useful quantitative criterion for defining the diffraction regimes is the Fresnel Number [20, p.117]:

$$N_f = \frac{D_2^2}{\lambda z} \tag{1}$$

$$N'_f = \frac{D_1^2}{\lambda z} \tag{2}$$

Where D_2 is the region of interest in the observation plane and D_1 is the characteristic aperture "diameter". When $N_F << 1$ and $N'_F << 1$ the Fraunhofer approximation can be used and otherwise, under the additional constraints discussed later on, the Fresnel approximation can be applied.

3.1 Huygens-Fresnel Principle and Rayleigh-Sommerfeld Diffraction Formula

The Rayleigh-Sommerfeld diffraction formula (Eq. 6) can be used to relate the complex amplitude of a field at a planar aperture to that on an observation plane. It is a mathematical statement of the Huygens-Fresnel Principle wherein the field in the observation plane is given by a linear superposition of diverging spherical waves [3, p.58]. The Rayleigh-Sommerfeld equation can be derived starting from the Helmholtz equation: which is valid in non-dispersive mediums [3, p.42]:

$$(\boldsymbol{\nabla}^2 + k^2)U = 0 \tag{3}$$

The Helmholtz equation can be derived from Maxwell's equations [1, p.36] with the assumptions that the propagation medium is linear. isotropic and nondispersive [1, p.37]. The solutions to Eq. 3 are the time-independent complex amplitudes of the vector components of the electric field, hence the term *scalar diffraction theory*. The Helmholtz equation is converted to an integral equation using Green's theorem such that the solution to the differential equation, which is in this case the complex amplitude U at some arbitrary point P_0, is given by:

$$U(P_0) = \frac{1}{4\pi} \iint_{\Sigma} (U\frac{\partial G}{\partial n} - G\frac{\partial U}{\partial n})ds \tag{4}$$

Where Σ indicates that the integral is taken over the aperture. The Rayleigh-Sommerfeld solution uses the Green's function:

$$G(P_1) = \frac{e^{jkr_{01}}}{r_{01}} - \frac{e^{jkr_{01}}}{\tilde{r}_{01}}$$

Along with the requirement that the observation distance be many wavelengths from the aperture,

$$r_{01} >> \lambda \tag{5}$$

which allows us to neglect the $\frac{1}{r_{01}}$ term from the derivative (See Eq. 4) of the green's function [1, p. 43]. We can thus obtain the Rayleigh-Sommefeld diffraction formula:

$$U(P_0) = \frac{1}{i\lambda} \iint_{\Sigma} (U(P_1)\frac{e^{ikr_{01}}}{r_{01}} cos(\theta)ds \tag{6}$$

3.2 Fresnel Approximation

The Fresnel diffraction integral can be obtained starting from the Rayleigh-Sommerfeld diffraction integral, and replacing $cos(\theta)$[1] by:

$$cos(\theta) = \frac{z}{r_{01}}$$

We obtain,

$$U(P_0) = \frac{1}{i\lambda}\int\int_{\Sigma}(U(P_1)\frac{e^{ikr_{01}}}{r_{01}}\frac{z}{r_{01}}ds \tag{7}$$

Where,

$$\begin{aligned} r_{01} &= (z^2 + (x-\xi)^2 + (y-\nu)^2)^{0.5} \\ &= z(1 + (\frac{x-\xi}{z})^2 + (\frac{y-\eta}{z})^2)^{0.5} \end{aligned}$$

Where ξ and η are the transverse coordinates in the aperture plane and x and y are the transverse coordinates in the observation plane. To arrive at the Fresnel approximation, we make a binomial expansion in r_{01}:

$$r_{01} = (z^2 + (x-\xi)^2 + (y-\nu)^2)^{0.5} \tag{8}$$

$$\approx z(1 + \frac{1}{2}(\frac{x-\xi}{z})^2 + \frac{1}{2}(\frac{y-\eta}{z})^2) \tag{9}$$

For the r_{01}^2 in the denominator of Eq. 7 we keep the first term of the expansion and for the r_{01} in the exponential we retain the terms shown above. This is adequate as long as [1, p.69]:

$$z^3 >> \frac{\pi}{4\lambda}\big[(x-\xi)^2 + (y-\eta)^2\big]^2_{max} \tag{10}$$

The Fresnel diffraction integral can be shown to be a solution of the paraxial helmholtz equation [2] [19],[3, pp.72-73] and it is obvious that in this context the binomial approximation is equivalent to making a small angle approximation and hence the paraxial approximation, because we have assumed that:

$$(\frac{x-\xi}{z})^2 + (\frac{y-\eta}{z})^2 << 1$$

Which is equivalent to requiring that the angle $\sin(\theta) = \frac{R}{z} \approx \theta$ from a point on the aperture plane to a point on the observation plane be small in the sense that $R << z$:

$$\begin{aligned} (\frac{x-\xi}{z})^2 + (\frac{y-\eta}{z})^2 &= \frac{(x-\xi)^2 + (y-\eta)^2}{z^2} \\ &\equiv \frac{R^2}{z^2} \\ &\implies |\frac{R}{z}| << 1 \end{aligned}$$

[1] Note: This is a necessity when directly integrating the R-S diffraction numerically.

[2] In obtaining the paraxial approximation the slowly-varying envelope approximation is made wherein we require that the spatial amplitude of the electric field and its derivative do not vary appreciably over distances of the order of a wavelength and hence we note that those constraints apply to the validity of a given numerical solution.

Returning now to Eq. 10, which says that the phase shift induced by the most significant dropped term in the expansion be much less than π [20, p. 47]:

$$\frac{z}{8}\theta^4 k << \pi$$
$$\frac{z}{4\lambda}\theta^4 << 1$$
$$\frac{1}{4\lambda z}R^2\theta^2 << 1$$

Now making use of our definition of the Fresnel Number (Eq. 1) we find that our requirement is:

$$N_F << \frac{4}{\theta_{max}^2}$$

With these approximations the Fresnel approximation becomes a convolution:

$$U(x,y) = \frac{e^{jkz}e^{\frac{k}{2z}(x^2+y^2)}}{z}\iint_{-\infty}^{\infty} U(\xi,\eta)exp(-j\frac{k}{z}\big[(x-\xi)^2+(y-\eta)^2\big]d\xi d\eta \tag{11}$$

Alternatively, this can be written as a Fourier transform:

$$U(x,y) = \frac{e^{ikz}e^{\frac{k}{2z}(x^2+y^2)}}{i\lambda z}\iint_{-\infty}^{\infty}\Big[U(\xi,\eta)e^{j\frac{k}{2z}(\xi^2+\eta^2)}\Big]e^{i\frac{k}{2z}(x\xi+y\eta)}d\xi d\eta \tag{12}$$

These two representations of the Fresnel Diffraction integral form the basis of their numerical implementations which are discussion in Sec. 4.2. In deriving the Fresnel diffraction integral we note (see Eq. 12) that the linear superposition of spherical waves obtained via the Rayleigh-Sommerfeld diffraction integral has been transformed into a linear superposition of paraboloidal waves.

3.3 Fraunhofer Approximation

The Fraunhofer is an additional approximation made starting from the Fresnel approximation:

$$z >> \frac{k(\xi^2+\eta^2)_{max}}{2} \tag{13}$$

and such is also subject to all the constraints and approximations made thus far. This approximation is equivalent to the condition of Eq. 2, which can be shown as follows:

$$\begin{aligned} \frac{k(\xi^2+\eta^2)_{max}}{2z} &<< \pi \\ \frac{2\pi D_1^2}{2\lambda z} &<< \pi \\ \frac{D_1^2}{\lambda z} &<< 1 \\ \implies N_F' &<< 1 \end{aligned}$$

With this approximation, the diffraction integral can be expressed as a Fourier transform:

$$U(x,y) = \frac{e^{ikz}}{i\lambda z}\int\int_{-\infty}^{\infty} U(\xi,\eta) exp(-i\frac{k}{2z}\big[(x\xi+y\eta)\big]d\xi d\eta$$

Physically, the term in the integral can be interpreted as a linear superposition of plane waves with spatial frequencies $f_x = \frac{x}{\lambda d}$ and $f_y = \frac{y}{\lambda d}$. In the case where a unit plane wave is incident on the aperture, we can see that the far-field intensity distribution (the diffraction pattern) is given by [20, p.122]:

$$I(x,y) = \frac{1}{(\lambda z)^2}|F(\frac{x}{\lambda z}, \frac{y}{\lambda z})|^2$$

4 Numerical Implementations and Considerations

The following subsections discussion some considerations which must be made when evaluating the Rayleigh-Sommerfeld, Fresnel and Fraunhofer diffraction integrals numerically.

4.1 Rayleigh-Sommerfeld Diffraction Integral

The Rayleigh-Sommerfeld integral (Eq. 6) can be computed by the Angular Spectrum method or Direct Integration [2]. Evaluation of the Rayleigh-Sommerfeld diffraction integral has been performed, with excellent agreement with theoretical results (for irradiance distribution) along the z axis , for a circular aperture using Simpson's $\frac{1}{3}$ rule for direct integration [16] . The authors show their results in the form of a diagram (Fig. 2):

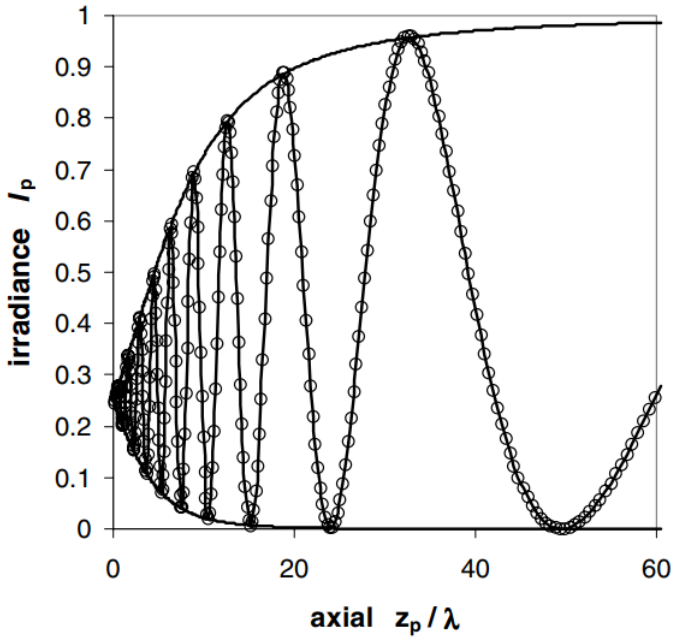

Figure 2: The circles are the results obtained using Simpson's 1/3 rule and the solid lines passing through them are the analytical results.

However, authors make no mention of the required grid spacing for accurate results. Whereas Simpson's rule approximates the integrand by a second order polynomial for each interval [14, pp.94-95] higher order quadratures are possible. Evaluation of the Rayleigh-Sommerfeld diffraction integral, for a circular aperture, using Gauss-Legendre quadrature is explored in [17]. Gauss-Legendre quadrature uses polynomials of order $2N - 1$, where N is the number of abscissas used for integration [6, p.175], to approximate the integrand over each interval. It should be noted however, that the authors make use of the Fresnel and Fraunhoffer approximations in determining their integrand, in Eq. 26 and Eq.31 of [17] respectively and hence it is not a true numerical solution of the Rayleigh-Sommerfeld diffraction integral.

The Monte Carlo technique is a well established method for estimating integrals (see [6], [14]) and the theory of Monte Carlo integration can be applied to evaluating the Rayleigh-Sommerfeld diffraction integral. To this authors surprise, there is a lack of literature exploring this technique in obtaining approximate results to the Rayleigh-Sommerfeld diffraction integral. The theoretical underpinning of the Monte Carlo technique is the use of the mean value theorem [14, p.105]:

$$\int_a^b f(x)dx = (b-a)\frac{1}{N}\langle f\rangle$$

Which allows an integral to be approximated as:

$$\int_a^b f(x)dx \approx (b-a)\big[\frac{1}{N}\sum_{i=1}^{N} f(x_i) \pm \sigma_N\big]$$

The standard deviation on the integrand is [6]:

$$\sigma_N = \sqrt{\frac{\frac{1}{N}\sum_{i=1}^{N} f(x_i)^2 - (\frac{1}{N}\sum_{i=1}^{N} f(x_i))^2}{N-1}}$$

Evaluating the Rayleigh-Sommerfeld formula using this method requires that the plane of observation be discretized ,and this resolution will limit the accuracy with which the field can be evaluated between data points. Another option is to sample both the input aperture in the observation plane randomly. However in either case the amplitude distribution in the observation plane will not be sufficient for perfect reconstruction of the exact result if the sampling requirements are not met. From [18], when using nonuniform sampling, as is the case with Monte Carlo because the points are chosen randomly, as long as the average distance between points is above or equal to to twice the band-limit of the function under consideration then that function can be reconstructed exactly from the samples [3].

[3]See *Lemma for 1-D Signals* in Chapter 4 of [18] and by *Proposition 1* the author believes this holds for 2-D signals provided the random samples do not exclusively sample the zeros of the integrand.

As a first step in exploring the feasibility of Monte Carlo techniques for evaluating the Rayleigh-Sommerfeld diffraction integral, a MATLAB program was developed for a square aperture. To do this, we start by writing Eq. 6 as:

$$U(P_o) = \frac{z}{i\lambda} \iint_{\Sigma} U(P_1) \frac{e^{ikr_{01}}}{r_{01}^2} d\xi d\eta$$

$$U(x,y) \approx A \frac{z}{i\lambda N^2} \sum_{i=1}^{N} \sum_{k=1}^{N} \frac{e^{ik\left[(x-\xi)^2+(y-\eta)^2+z^2\right]^{0.5}}}{\left[(x-\xi)^2+(y-\eta)^2+z^2\right]}$$

Where N is the number of samples in each dimension of the planar input aperture of area A. Shown below are simulation results for a unit amplitude plane wave with an optical wavelength of 1 μm incident upon a square diffracting aperture of size 2×10^{-3} m $N = 50000$, observed on an aperture of size 0.004 m with grid spacing 2×10^{-5} m in each dimension, at distance of $z = 1$ m:

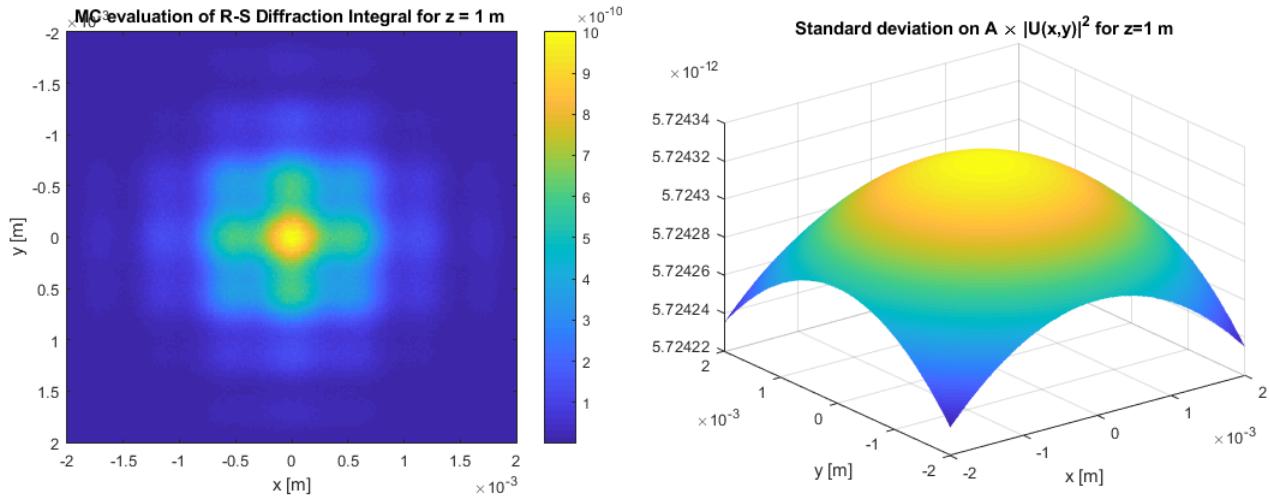

Figure 3: Left: Result of Monte Carlo integration of R-S diffraction integral. Right: Standard deviation of intensity at each grid point in the observation plane, it is approximately flat across the observation region.

For this problem the Fresnel Numbers are $N_F = 16$ and $N'_F = 4$, from which we deduce that we are in the very near field. At a distance of 500 m $N'_F = 8 \times 10^{-}3$, which implies that we are in the far-field. The diffraction pattern at this distance can then be compared with analytical results of the Fraunhofer approximation. The exact solution in the Fraunhoffer approximation was plotted at the same observation points used in the MC evaluation of the R-S integral, those results are shown alongside the results obtained for the R-S integral for the same problem:

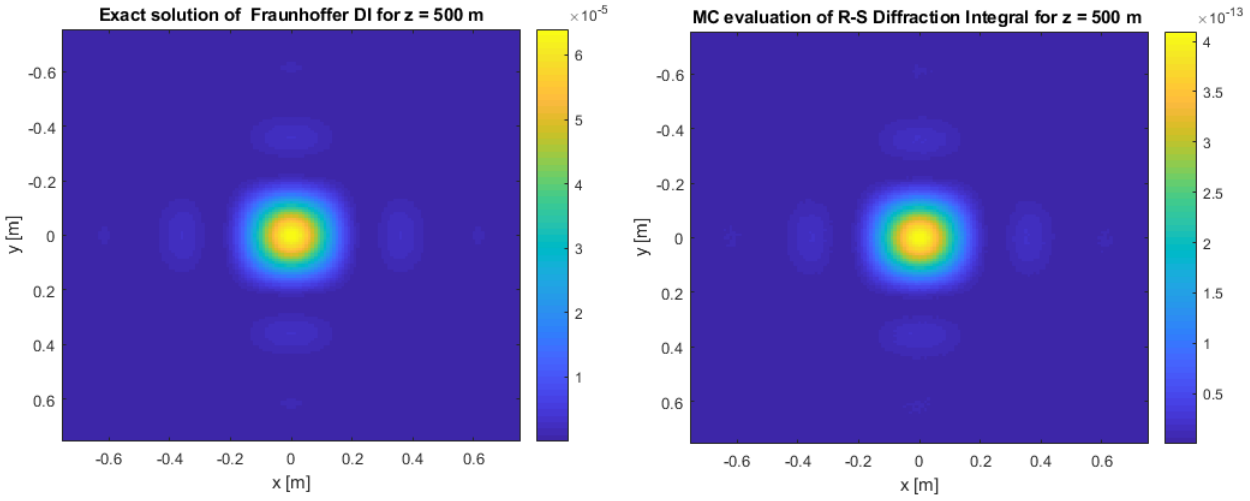

Figure 4: Left: Exact solution of Franhofer diffraction integral at z=500 plotted at the grid points used for the MC simulation. Right: Result of Monte Carlo integration of R-S diffraction integral

Qualitatively the agreement is excellent, however examining the colorbars to the right of each diffraction pattern it is obvious that the amplitude is incorrect by many orders of magnitude. The code used for this particular problem along with the defintions of the parameters can be found in Appendix A.

4.2 Fresnel Diffraction

The literature discusses many methods of evaluating the Fresnel diffraction integral and these originate from the different ways of expressing Eq. 11. The one-step propagation discretizes the diffraction integral in the form of Eq. 12, in short-hand:

$$U(x,y) = \frac{e^{ikz}e^{\frac{k}{2z}(x^2+y^2)}}{i\lambda z}\mathcal{F}\{U(\xi,\eta)e^{i\frac{k}{2z}(\xi^2+\eta^2)}\} \tag{14}$$

One of the main challenges of implementing the diffraction integral in this form is sampling of the complex exponential in the fourier transform in Eq. 14 ([4, p.121]). The complex exponential is chirped and highly oscillatory for small propagation distances which makes adequate sampling impractical in some situations [9]. In an x-y plane, transverse to the propagation direction the local spatial frequency, in $\frac{rad}{m}$, is given by the gradient of the phase ϕ of the argument of the complex exponential:

$$\omega_{local} = \boldsymbol{\nabla}\phi \tag{15}$$

$$= \frac{\partial\phi}{\partial x}\frac{k}{2z}(x^2+y^2)\hat{x} + \frac{\partial\phi}{\partial y}(\frac{k}{2z})(x^2+y^2)\hat{y} \tag{16}$$

$$= 2\pi\frac{x}{\lambda z}\hat{x} + 2\pi\frac{y}{\lambda z}\hat{y} \tag{17}$$

When determining constraints in computing the DFT in Eq. 14, we can restrict ourselves to an analysis in the x dimension due to the separability of the impulse response function [11]. Assuming identical grid spacing and size in ξ and η, for a grid spacing of δ_1 with N grid points in each transverse dimension we obtain a Nyquist frequency of $\frac{1}{2\delta_1}$ and a frequency sample spacing of $\delta_{fx} = \frac{1}{N\delta_1}$ then referring to Eq. 12, we note that $f_x = \frac{x}{\lambda z}$ [9], resulting in an observation plane grid spacing of [4, p.90]:

$$\delta_{fx} = \frac{\delta x}{\lambda z} \tag{18}$$

$$\delta x = \lambda z\delta_{fx} \tag{19}$$

$$= \frac{\lambda z}{N\delta_1} \tag{20}$$

These relations, along with the constraints imposed by Eq. 5 and the approximation made in the binomial expansion (Eq. 8) must be taken into consideration when using the one-step Fresnel diffraction. Each criterion imposes a lower limit on the propagation distance and they dictate the ultimate lower limit [9] [4]. As an example, consider a diffraction problem where for an incident plane wave with wavelength $\lambda = 0.6328$ μm with a grid spacing of 50 μm and a 500×500 point sampling grid, resulting a grid dimension of 25 mm in the aperture plane. The highest frequency that can be sampled by such a grid (the Nyquist frequency) is $f_N = \frac{1}{2}\frac{1}{50\times10^{-6}}\ \frac{Cycles}{m}$. The local frequency (see Eq. 17) will reach a maximum value of $\frac{25\times10^{-3}}{0.6328\times10^{-6}z}$ $\frac{Cycles}{m}$. Equating the Nyquist frequency with the maximum local frequency and solving for z we obtain a minimum propagation distance of just over 76 m. Noting then that Eq. 5 is satisfied, computing Eq. 10 yields the condition $z >> 2.9998$ m for an observation plane size of 2×0.4829 m (obtained via the relation $x_{2max} = v_{xmax}\lambda z$ to ensure all frequency content is contained in the plane). It is also important to note that when using the one-step propagation method, the grid spacing in the observation plane is fixed by the grid spacing in the input plane. This limitation can be overcome by introducing the two-step propagation method which is discussed briefly in the following subsection.

[4] There may be a mistake in the calculation of lower limit computed in [9, Eq. 5] because the condition on the binomial approximation is actually on z^3, not z^2

Returning now to the discussion of the frequency content of Eq. 14, the sampling condition required in order to avoid folding is the sum of the maximum frequency contained in the chirp function, the input wave and the Fourier transform of the aperture (ex: $\Pi(x)$). This can be seen by multiplying the Fourier series of each function. In one dimension,:

$$\begin{aligned} f_1(x)f_2(x)f_3(x) &= \Big\{ \sum_{n1=1}^{\infty} (c_{n1}e^{in_1x}\Big\} \times \Big\{ \sum_{n2=1}^{\infty} (c_{n2}e^{in_2x}\Big\} \times \Big\{ \sum_{n3=1}^{\infty} (c_{n3}e^{in_3x}\Big\} \\ &= c_{11}e^{in_1x}c_{12}e^{in_2x}c_{13}e^{in_1x} + c_{21}e^{in_1x}c_{12}e^{in_2x}c_{13}e^{in_1x} + ... \\ &+ c_{N1}e^{iN_1x}c_{N2}e^{iN_2x}c_{N3}e^{iN_3x} \end{aligned}$$

Expanding the last term, which represents the multiplication of the maximum frequency content of all three functions we obtain:

$$c_{N1}e^{N_1x}c_{N2}e^{iN_2x}c_{N3}e^{iN_3x} = c_{N1}c_{N2}c_{N3} \times e^{i(N_1+N_2+N_3)x}$$

We thus see that the cutoff frequency of the product of three functions will be equal to the sum of the maximum frequency contained in each function. Alternatively, one can consider the effect of a convolution in the frequency domain to obtain this result by considering the distances at which the spectrums will start to overlap. Unfortunately, in many cases the functions will not be bandlimited. For example, consider a square aperture of size $L = \delta N = 50 \times 10^{-6}$ m $\times 500 = 0.0250$ m, the Fourier transform of a function which is 0 within the aperture and outside the aperture is $Lsinc(Lf_x)$, which is not band-limited. Hence, the propagated field in the observation plane over a region D_2 cannot represent the full extent of the propagated field but instead defines the region over which we want to accurately represent the propagated field.

4.2.1 One-Step Propagation

Following the work of [7], the Fourier transform in Eq. 14 can be decomposed into it's plane wave spectrum with the spatial frequencies mapped to direction cosines. Since we have defined the maximum spatial frequency of interest in the defininion of the grid spacing ($f_{max} = \frac{1}{2\delta_1}$) we see that the maximum angular content that can affect the observed field is $\alpha_{max} = \frac{\lambda}{2\delta_1}$ and $\beta_{max} = \frac{\lambda}{2\delta_1}$ (assuming identical grid spacing in ξ and η)where α and β are the direction cosines of the angular spectrum. The local spatial frequencies of these wavefronts correspond to rays propagating from a point on D_1 to a point on D_2 and with a finite and equal number of grid points on D_1 and D_2 only a finite number of rays can be propagated. Figure 5 illustrates these constraints:

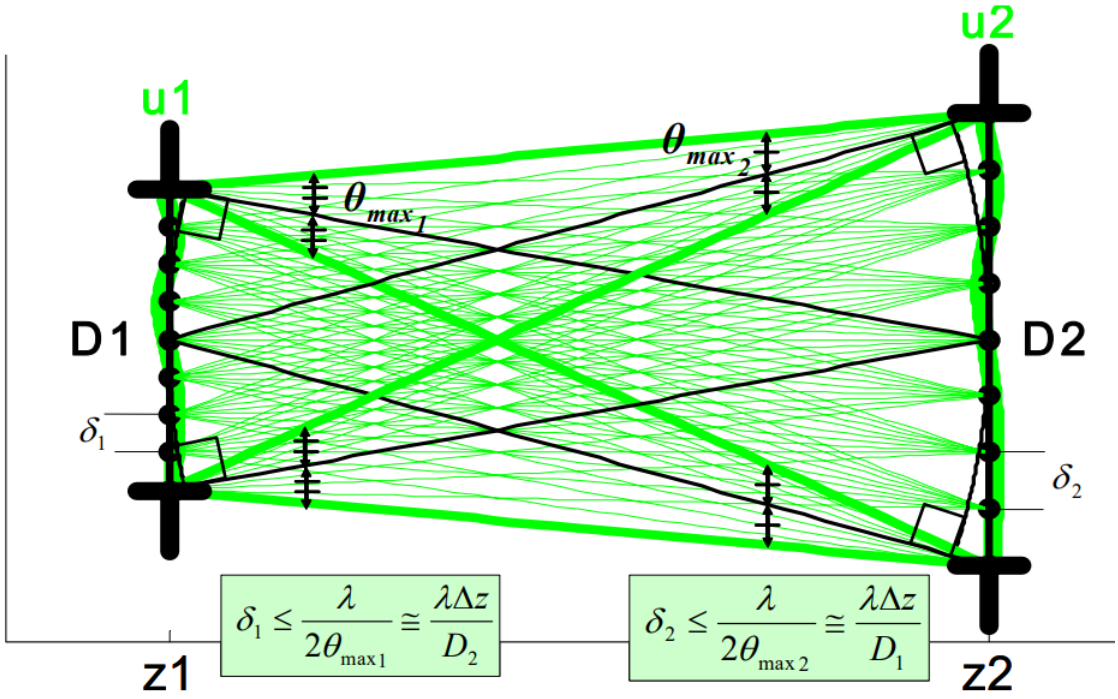

Figure 5: Propagation diagram showing sampling constraints [7].

The thick green lines on the diagram are lines drawn from the extremal point on one aperture to the extremal points on the other and determine the maximum angle a given ray can make with the reference normals and still be observed within the aperture D_2. From this geometry, we can determine that the geometrical constrain on α is [4, p. 118]:

$$\alpha_{max} = \frac{D1\frac{\delta_2}{\delta_1} + D_2}{2\Delta z}$$

Using this constraint, along with the Nyquist criterion for α_{max}, we find that in order to adequately sample the spatial bandwidth that affects the observation region [4, pp.118 -119], we must meet the following constraint:

$$\frac{D1\frac{\delta_2}{\delta_1} + D_2}{2\Delta z} \le \frac{\lambda}{2\delta_1}$$

Because the DFT assumes a periodic input we must introduce a guard-band in the observation plane such that the rays from "adjacent" apertures do not affect do not affect the observation region D_2. This is discussed in [8] and is referred to as Full-Aperture-Illumination Theory. It be visualized as follows:

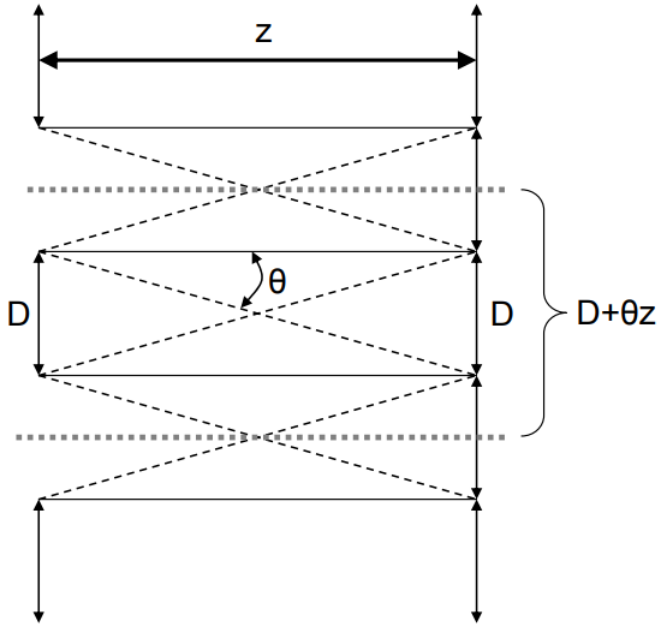

Figure 1 - Variable definition for choosing mesh parameters for a simple propagation

Figure 6: Symmetrical propagation diagram showing imposed guard band [8].

We see that the propagated rays from adjacent "apertures" will intersect the observation region. In the case shown above where $D_1 = D_2 = D$ we see that the mesh in the observation plane should be extended to include points in the region $\frac{D}{2} \pm \theta z$. For a non-symmetrical propagation diagram, this can constraint can be written as:

$$\begin{aligned} D_{grid} &= D_1 \frac{\delta_2}{\delta_1} + 2\alpha_{max} \\ &= D_1 \frac{\delta_2}{\delta_1} + 2\frac{\lambda z}{\delta_1} \end{aligned}$$

In doing so, we sacrifice the accuracy of the results in the region $x \in \pm[\frac{D_2}{2}, D_1 \frac{\delta_2}{2\delta_1} + \frac{\lambda z}{\delta_1}]$ in order to avoid aliasing in the observing aperture and we find that the number of required grid points in the observation plane is $N = \frac{D_{grid}}{\delta_2}$ [4, p.119-120].

With these approximations in mind and extending the the theory to two dimensions, Fresnel diffraction can be performed. However smearing of the spectrum is inevitable due to the fact that we must consider an aperture of finite support and hence the spectrum will not be of finite support. When performing numerical computations it is good practice [5] to repeat the computations for multiple grid spacings and to check the results for convergence up to a user-defined/problem-specific tolerance.

[5]This thinking is embedded into the RK45 algorithm and this author has seen papers (ex: "Simple Patterns in a Complex System", by Pearson), which use FEM to solve problems, make statements on the convergence of their results.

4.2.2 Two-Step Propagation

In a two-step propagation an intermediate plane is placed between the aperture plane and the observation plance as a means of introducing a scaling factor between the aperture grid spacing and the observation grid spacing ([4],[7]) and can be derived most easily by expressing the diffraction integral in operator form [6] The drawback of this method is that it requires two DFTs, making it more computationally expensive than the first method. Using this method, we find that the grid spacing in the observation plane is given by $\delta_2 = |\frac{\Delta z_2}{\Delta z_1}|\delta_1$ ([4, p.94], [7]).

4.3 Fraunhofer Diffraction Integral

We compute the digital spectrum at values: $f_x = f_y = \frac{-1}{2\delta} : \frac{1}{N\delta} : \frac{1}{2\delta}$ Where the spectrum is obtained via the DFT:

$$F(l,p) = \delta^2 \sum_{m=-\frac{N}{2}}^{\frac{N}{2}} \sum_{n=-\frac{N}{2}}^{\frac{N}{2}} h(m\delta, n\delta) e^{-i2\pi(\frac{lm}{N}+\frac{pn}{N})}$$

On the observation plane, the spatial frequencies are mapped to the following x-y coordinates:

$$x = f_x \lambda z$$
$$y = f_y \lambda z$$

From which the spectrum of the aperture could theoretically be obtained. This authors' implementation of the DFT is based on a shifted indexing and identical grid spacing and number of samples in each dimension was adequate for this study. The DFT thus has the form:

$$F(l,p) = \delta^2 \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} h((m-N/2)\delta, (n-N/2)\delta) e^{-i2\pi(\frac{l(m-N/2)}{N}+\frac{p(n-N/2)}{N})} \tag{21}$$

$$= \delta^2 \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} h((m-N/2)\delta, (n-N/2)\delta) e^{-i2\pi(\frac{l}{N}(m-N/2)+\frac{p}{N}(n-N/2))} \tag{22}$$

$$\tag{23}$$

Which allows us to avoid shifting the function around in the spatial domain prior to computing the DFT. The resulting spectrum still has to be rearanged however, because the second half of the output corresponds the the negative frequency content. The MATLAB implementation of the DFT (Eq. 21) is uses a succession of matrix multiplications to evaluate the DFT as follows:

$$\delta^2 \begin{bmatrix} e^{-i2\pi\frac{v}{N}(0-N/2)} & e^{-i2\pi\frac{v}{N}(1-N/2)} & \dots & e^{-i2\pi\frac{v}{N}(N-1-N/2))} \end{bmatrix} \begin{bmatrix} f(1,1) & f(1,2) & \dots & f(1,N) \\ f(2,1) & f(2,2) & \dots & f(2,N) \\ \vdots & \vdots & \vdots & \vdots \\ f(N,1) & f(N,2) & \dots & f(N,N) \end{bmatrix} \begin{bmatrix} e^{-i2\pi\frac{w}{N}(0-N/2)} \\ e^{-i2\pi\frac{w}{N}(1-N/2)} \\ \vdots \\ e^{-i2\pi\frac{w}{N}(N-1-N/2))} \end{bmatrix}$$

Where $f(1,1)$ is $f(-wN/2, -wN/2)$ etc.

[6] See [10] for definitions and examples.

The code for which can be found in Appendix B. The DFT based Fraunhofer diffraction integral was applied to the same scenario discussed in the text immediately preceeding Fig. 4. A grid spacing of 10^{-4} m was used, providing a spectrum in the range $[-\frac{1}{2} \times 10^4, \frac{1}{2} \times 10^4]$.

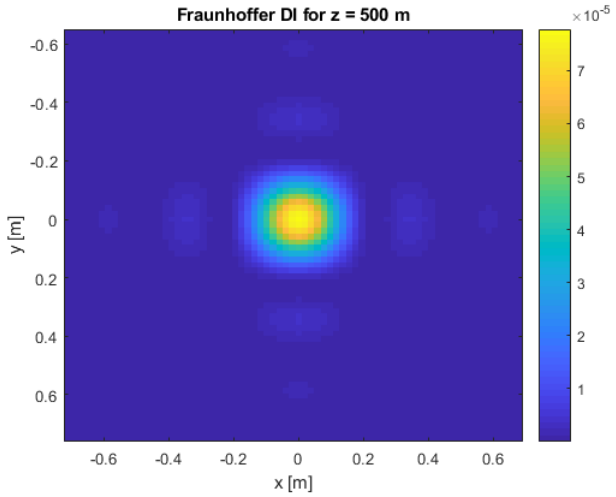

Figure 7: Fraunhofer diffraction pattern for a square 10^{-3} m aperture at 500 m.

Comparing the above figure with the exact results from Fig. 4 we see good agreement between the two. As an additional measure to verify the accuracy of the results, a grid with half the resolution was chosen and the resulting diffraction pattern was compared with the original. The resulting diffraction pattern closely matched the original and hence the effect of aliasing could be assumed to be small. Furthermore, the fact that the exact result is known and that the resulting sinc function has a very small value at frequencies on the order of 10^4 Hz, the upper frequency limit of the DFT (with frequencies mapped to spatial coordinates) shown in Fig. 7 (which was enlarged and cropped to show detail).

5 Conclusion

The stipulation that a wavefront consists of a linear superposition of spherical waves, which serves as the foundation of the theory in this paper, provides an accurate formalism for examining diffraction problems and it is surprising that [1, p.52] claims that the Huygen's-Fresnel principle is simply a mathematical construct that allows us to solve diffraction problems.

When performing numerical simulations of diffraction, in order to guarantee the validity of the results it is important to take into consideration the relevant constraints and approximations laid out in this paper. Nevertheless, these constraints and approximations should not be seen as a burden but knowledge that can allow for the use of the most effective tool for a given problem. Future work would be envisioned to involve refining the Monte-Carlo method for evaluating the Rayleigh-Sommerfeld diffraction integral; by increasing fidelity and by using more advanced sampling algorithms.

References

[1] Goodman, Joseph W., "Introduction to Fourier Optics", Third Edition, Roberts & Company

[2] Fabin Shen and Anbo Wang, "Fast-Fourier-transform based numerical integration method for the Rayleigh-Sommerfeld diffraction formula", 20 February 2006 Vol. 45, No. 6 APPLIED OPTICS

[3] Ersoy, Okhan K., "Diffraction, Fourier Optics and Imaging", John Wiley and Sons, Inc, 2007

[4] Jason D. Schmidt, "Numerical Simulation of Optical Wave Propagation with Examples in MATLAB", 2010, DOI: 10.1117/3.866274

[5] Bracewell, Ronald N. "The Fourier Transform and Its Applications". New York :McGraw-Hill, 2000.

[6] Devries, Hasbun. "A First Course in Computational Chysics. Jones and Bartlett", Second Edition.

[7] Steve Coy, "Choosing mesh spacings and mesh dimensions for wave optics simulation", Proc. SPIE 5894, Advanced Wavefront Control: Methods, Devices, and Applications III, 589405 (18 August 2005); doi: 10.1117/12.619994; https://doi.org/10.1117/12.619994

[8] J. D. Mansell, R. Praus, and S. Coy. "Determining Wave-Optics Mesh Parameters for Complex Optical Systems", SPIE Proc. Vol. 6675-17 (2007)

[9] Sypek, Maciej. "Light Propagation in the Fresnel Region. New Numerical Approach." Optics Communications 116.1-3 (1995): 43-8. Web. 31 Mar. 2018

[10] M. Nazarathy and J. Shamir, "Fourier optics described by operator algebra," J. Opt. Soc. Am. 70, 150-159 (1980)

[11] David G. Voelz and Michael C. Roggemann, "Digital simulation of scalar optical diffraction: revisiting chirp function sampling criteria and consequences," Appl. Opt. 48, 6132-6142 (2009)

[12] Wolfe, Chris, "Multiphysics:THE FUTURE of SIMULATION", ANSYS ADVANTAGE Volume VIII, Issue 2, 2014

[13] Lyndon Evans and Philip Bryant 2008 JINST 3 S08001

[14] Landau, Rubin H, Mejia M. J. Paez, and Cristian C. Bordeianu. "Computational Physics: Problem Solving with Computers". Weinheim: Wiley-VCH, 2007. Print.

[15] Benov, Dobriyan M. "The Manhattan Project, the First Electronic Computer and the Monte Carlo Method." Monte Carlo Methods and Applications 22.1 (2016): 73-9.

[16] Cooper, I. J., C. J. R. Sheppard, and M. Sharma. "Numerical Integration of Diffraction Integrals for a Circular Aperture." Optik - International Journal for Light and Electron Optics 113.7 (2002): 293-8. Web. 2 Apr. 2018

[17] Evans, W., and A. Torre. "On Numerical Integration with High-Order Quadratures: With Application to the Rayleigh-Sommerfeld Integral." Applied Physics B 109.2 (2012): 295-309.

[18] Marvasti, Farokh. "Nonuniform Sampling Theory and Practice", Springer, Boston, MA

[19] Grella, R. "Fresnel Propagation and Diffraction and Paraxial Wave Equation", Journal of Optics 13.6 (1982): 367-74.

[20] Saleh, Bahaa E. A., Teich, Malvin Carl. "Fundamentals of photonics", Wiley-Interscience, Second Edition

A Code Listing for Monte Carlo integration of R-S Diffraction Integral

```
clear
% Half the length of the observation plane
%l=0.002;
l=0.75;
%grid spacing in observation plane
dl=1e-02;

%Aperture half-lengths
lx=1e-3;
ly=1e-3;

%optical wavelength
lambda=10^-6;
k=2*pi/lambda;

%propagation distance (such that r >> lambda)
z=500;

%Compute fresnel numbers of the problem
N_F= (2*l)^2/(z*lambda);
N_Fp = (2*lx)^2/(z*lambda);
%Number of sources on the aperture plane which contribute to the a given grid point on the
%observation plane
N=50000;
%initialize x dimension counter
i=0;
%initialize y dimension counter
m=0;
U(1:length(-l:dl:l),1:length(-l:dl:l))=0;
A=2*lx*2*ly;
%loop over the observation grid and perform uniform sampling to evaluate
%the RS diffraction integral at each point
for Xo=-l:dl:l
    i=i+1;
    for Yo=-l:dl:l
        m=m+1;
        %generate uniformly distributed points in aperture plane
        X = (2*lx).*rand(N,1)-lx;
        Y = (2*ly).*rand(N,1)-ly;
        %total number of points generated (N^2)
        N2=length(X)*length(Y);
        %array of contributions from each point in the aperture plane
        ps=(1*z/(1i*lambda))*A*(1./((X-Xo).^2+(Y-Yo).^2+z^2)).*exp(1i*k*((X-Xo).^2+(Y-Yo).^2+
            z^2).^0.5);
        %sum up the contributions
        U(i,m)=(k/(1i*2*pi))*A*sum(sum(ps));
        %take the average
        U(i,m)=U(i,m)/N2;
        %standard deviation on the intensity distrivution at each point on the observation
            plane
        sig(i,m)=A*(((1/N2)*sum(sum(abs(ps).^4)) -abs(U(i,m)).^4)/(N2-1))^0.5;
    end
    m=0;
end

figure
imagesc([-l:dl:l],[-l:dl:l],abs(U).^2)
colorbar
xlabel('x [m]')
ylabel('y [m]')
```

```
title('MC evaluation of R-S Diffraction Integral for z = 500 m ')
 [X,Y]=meshgrid([-l:dl:l],[-l:dl:l]);
figure
surf(X,Y,sig)
shading interp
xlabel('x [m]')
ylabel('y [m]')
title('Standard deviation on A \times |U(x,y)|^2 for z=500 m')
%
% figure
% surf(X,Y,abs(U).^2)
% shading interp
% xlabel('x [m]')
% ylabel('y [m]')
% title('MC evaluation of R-S Diffraction Integral for z = 1 m ')


%Compare with exact solution from [p.76], Goodman
I=@(x,y,A,lambda,z,w) (A.^2./(lambda.^2.*z.^2)).*sinc(2*w*x./(lambda.*z)).^2.*sinc(2*w*y./(
    lambda.*z)).^2;
x=[-l:dl:l];
y=x;
[X,Y]=meshgrid(x,y);
AnalyticSquare=I(X,Y,A,lambda,z,lx);
RelErr=abs(abs(U).^2./(max(max(abs(U).^2)))- AnalyticSquare./(max(max(AnalyticSquare)))./(
    AnalyticSquare./(max(max(AnalyticSquare))));
figure
%surf(X,Y,RelErr)
imagesc(x,y,AnalyticSquare)
colorbar
xlabel('x [m]')
ylabel('y [m]')
title('Exact solution of  Fraunhoffer DI for z = 500 m')
%title('Relative Error between Exact Fraunhoffer and MC R-S for z = 500 m')

% Half the length of the observation plane
%l=0.002;
l=0.75;
%grid spacing in observation plane
dl=1e-02;

%Aperture half-lengths
lx=1e-3;
ly=1e-3;

%optical wavelength
lambda=10^-6;
k=2*pi/lambda;

%propagation distance (such that r >> lambda)
z=500;

%Compute fresnel numbers of the problem
N_F= (2*l)^2/(z*lambda);
N_Fp = (2*lx)^2/(z*lambda);
%Number of sources on the aperture plane which contribute to the a given grid point on the
%observation plane
N=50000;
%initialize x dimension counter
i=0;
%initialize y dimension counter
m=0;
U(1:length(-l:dl:l),1:length(-l:dl:l))=0;
A=2*lx*2*ly;
```

```
%loop over the observation grid and perform uniform sampling to evaluate
%the RS diffraction integral at each point
for Xo=-l:dl:l
    i=i+1;
    for Yo=-l:dl:l
        m=m+1;
        %generate uniformly distributed points in aperture plane
        X = (2*lx).*rand(N,1)-lx;
        Y = (2*ly).*rand(N,1)-ly;
        %total number of points generated (N^2)
        N2=length(X)*length(Y);
        %array of contributions from each point in the aperture plane
        ps=(1*z/(1i*lambda))*A*(1./((X-Xo).^2+(Y-Yo).^2+z^2)).*exp(1i*k*((X-Xo).^2+(Y-Yo).^2+
            z^2).^0.5);
        %sum up the contributions
        U(i,m)=(k/(1i*2*pi))*A*sum(sum(ps));
        %take the average
        U(i,m)=U(i,m)/N2;
        %standard deviation on the intensity distrivution at each point on the observation
            plane
        sig(i,m)=A*(((1/N2)*sum(sum(abs(ps).^4)) -abs(U(i,m)).^4)/(N2-1))^0.5;
    end
    m=0;
end

figure
imagesc([-l:dl:l],[-l:dl:l],abs(U).^2)
colorbar
xlabel('x [m]')
ylabel('y [m]')
title('MC evaluation of R-S Diffraction Integral for z = 500 m ')
 [X,Y]=meshgrid([-l:dl:l],[-l:dl:l]);
figure
surf(X,Y,sig)
shading interp
xlabel('x [m]')
ylabel('y [m]')
title('Standard deviation on A \times |U(x,y)|^2 for z=500 m')
%
% figure
% surf(X,Y,abs(U).^2)
% shading interp
% xlabel('x [m]')
% ylabel('y [m]')
% title('MC evaluation of R-S Diffraction Integral for z = 1 m ')


%Compare with exact solution from [p.76], Bracewell
I=@(x,y,A,lambda,z,w) (A.^2./(lambda.^2.*z.^2)).*sinc(2*w*x./(lambda.*z)).^2.*sinc(2*w*y./(
    lambda.*z)).^2;
x=[-l:dl:l];
y=x;
[X,Y]=meshgrid(x,y);
AnalyticSquare=I(X,Y,A,lambda,z,lx);
RelErr=abs(abs(U).^2./(max(max(abs(U).^2)))- AnalyticSquare./(max(max(AnalyticSquare)))./(
    AnalyticSquare./(max(max(AnalyticSquare))));
figure
%surf(X,Y,RelErr)
imagesc(x,y,AnalyticSquare)
colorbar
xlabel('x [m]')
ylabel('y [m]')
title('Exact solution of  Fraunhoffer DI for z = 500 m')
%title('Relative Error between Exact Fraunhoffer and MC R-S for z = 500 m')
```

B Code Listing for DFT implementation of Fraunhofer Diffraction Integral

```
clear all
N=256;
%Fraunhoffer
t=zeros(N,N);
dN = 10;
t(N/2-dN:N/2+dN,N/2-dN:N/2+dN)=1;

lambda=10^-6;

z=500;

l=10^-3;
tau=l/(dN);
temp=DFT3(t,N);

f=[-(1/(2*tau)):1/(N*tau):(1/(2*tau))-1/(N*tau)]

figure
imagesc(f*lambda*z,f*lambda*z,(1/(lambda^2*z^2))*abs(tau^2*fftshift(temp)).^2)
colorbar
xlabel('x [m]')
ylabel('y [m]')
title('Fraunhoffer DI for z = 500 m')

function [F] = DFT4(f,N)
%Computes DFT of function f sampled at x, with a total of N
%sampled

for v=0:N-1
    for w=0:N-1
        F(v+1,w+1)=exp(-1i*2*pi*(v/N).*(-(N/2)+[0:N-1]))*(f)*exp(-1i*2*pi*(w/N).*((-N/2)+[0:N-1]))';
    end
end
end
```