Near-field diffraction from amplitude diffraction gratings: theory, simulation and results

Abedin, Kazi Monowar, Rahman, S.M. Mujibur


Event: 14th Conference on Education and Training in Optics and Photonics, ETOP 2017, 2017, Hangzhou, China
Near-field diffraction from amplitude diffraction gratings: 
Theory, simulation and results

Kazi Monowar Abedin*, S.M. Mujibur Rahman
Department of Physics, College of Science, Sultan Qaboos University,
P.O. Box 36, Al-Khoudh, Muscat, P.C. 123, Muscat, OMAN

ABSTRACT

We describe a computer simulation method by which the complete near-field diffract pattern of an amplitude diffraction grating can be generated. The technique uses the method of iterative Fresnel integrals to calculate and generate the diffraction images. Theoretical background as well as the techniques to perform the simulation is described. The program is written in MATLAB, and can be implemented in any ordinary PC. Examples of simulated diffraction images are presented and discussed. The generated images in the far-field where they reduce to Fraunhofer diffraction pattern are also presented for a realistic grating, and compared with the results predicted by the grating equation, which is applicable in the far-field. The method can be used as a tool to teach the complex phenomenon of diffraction in classrooms.

Keywords: Computer simulation, near-field diffraction, Fresnel integrals, amplitude diffraction gratings, MATLAB

1. INTRODUCTION

Diffraction is the characteristic of light to propagate around obstacles in a limited way, and is a well-known phenomenon in optics. Diffraction influences the real-life behavior of optical systems, such as microscopes and telescopes, and limits the fundamental resolution achievable in these devices. Diffraction phenomena can be classified into two types, near-field (Fresnel) diffraction and far-field (Fraunhofer) diffraction. Fraunhofer, or far-field diffraction is simpler to handle mathematically, because both the wavefronts incident on the aperture and on the observation screen are planar. In contrast, in the case of near-field (Fresnel) diffraction, where either the aperture-source distance or aperture-screen distance is finite, the wavefronts are not planar, and the solution becomes more complex, and the analytical form of the diffraction pattern cannot be found even in the simplest cases. Therefore, numerical methods must be used. Solutions can be derived in terms of Fresnel-Kirchhoff, or Rayleigh-Sommerfeld diffraction integrals. The large amount of numerical data in the calculation of the integrals in real situations means that the use of computers and simulation techniques can be used effectively. With the use of modern computers and software, two-dimensional fast-Fourier transform (FFT) methods can be used to calculate diffraction integrals. For example, the diffraction and propagation of waves from any aperture have been simulated in computers by Rudolf et al. using Helmholtz–Kirchhoff diffraction integrals.

In previous work, we introduced new technique to simulate the diffraction field of apertures having rectangular shapes. This technique is called the Iterative Fresnel Integral Method (IFIM). It uses virtual displacement of the aperture, and repeated calculations of Fresnel integrals to construct the diffraction pattern from rectangular-shaped apertures. The technique can be applied to many interesting optical problems of importance. Using the technique, implemented in computer codes such as MATLAB, the diffraction field from any rectangular-shaped aperture in any situation can be calculated in a few minutes. An amplitude diffraction grating is essentially an array of long parallel rectangular apertures separated by opaque regions. Both the width of the apertures and the opaque regions are of the order of wavelength of light. This grating is widely used in spectrometers to separate the light into constituent wavelengths. In all these cases, the wavefront incident on the grating and that emitted from the grating are both rendered plane by a collimating mirror or lens. In that case, the analytical treatment of the diffraction grating becomes much simpler, because the Fraunhofer diffraction regime can be used to calculate the diffraction fields. The situations, where this Fraunhofer condition is not apploable are not considered or used, because of difficulty of treating Fresnel diffraction from a grating.

In the present Proceedings, we describe how the IFIM technique can be used and extended to the non-trivial case of amplitude diffraction gratings, and discuss some possible applications in teaching diffraction phenomena in classrooms.

* abedin@squ.edu.om; phone +968-2414-1446; fax:+ 968-2441-4228
2. INTRODUCTION TO THE ITERATIVE FRESNEL INTEGRALS METHOD

2.1 Theory of a Fresnel diffraction for $N$ rectangular apertures

Let us consider the case of $N$ rectangular parallel apertures (which is the basis of an amplitude diffraction grating) being illuminated by a light source with a wavelength $\lambda$ at a distance $p$ away. The observation screen is located at a distance $q$ away. The width of each aperture is $a$ and the opaque regions between apertures has a width $b$, the inter-aperture separation being $(a+b)$. Each aperture has a height of $c$ in the $z$-direction. The two edges of the central aperture of the $N$-aperture system (called aperture $0$) are then located at $y_0'= -a/2$ and $y_0=a/2$ respectively, and the edges of the aperture for the next aperture to the right (called aperture +1) are located at $y_1'=a/2+b$ and $y_1=3a/2+b$ respectively. The next aperture to the right (called aperture +2) are located at $y_2=3a/2+2b$ and $y_2'=5a/2+2b$. Finally the edges of the first left aperture (aperture -1) are located at $y_{-1}=-a/2-b$ and $y_{-1}'=-3a/2-b$, and, the edges of the aperture -n will be located at $y_{-n}=(2n-1)a/2+nb$ and $y_{-n}'=(2n+1)a/2+nb$. As shown in Fig. 1 (shown for the case of $N=5$), the coordinate systems on the aperture and on the image planes are chosen to be centered on the optical axis passing through the center of the central aperture (aperture 0) and normal to it, and are denoted by $(y, z)$ and $(Y, Z)$ axes, respectively. The Huygens–Fresnel principle is then invoked to calculate the total electric field at any given point of the image plane $(Y, Z)$ by summing up all the contributions of all the elementary wavelets emitted by different area elements inside each of the rectangular apertures. The net complex electric field at the point $P$ (located at the origin of the image plane or screen) due to waves emitted from the central aperture is given by

$$E_{Po} = \int_{aperture} \frac{\epsilon_0 k(\theta)}{p,q} \exp\left[j(k(p+q)-\theta_0)\right] dS$$

where $E_{Po}$ is the electric field of the source, $k$ is the wavenumber ($=2\pi/\lambda_0$) and $K(\theta)$ is the obliquity factor, which accounts for secondary Huygens wavelets emitted from area elements $dS$ in an oblique direction. By integrating the contributions of the Huygens wavelets over the entire area of the central aperture, it can be shown\(^{19}\) that the total electric field at $P$ is,

$$E_{Po} = \left[C(u) + jS(u)\right]_{u_0}^{u_v} \left[C(v) + jS(v)\right]_{v_0}^{v_v}$$

where $E_o$ is the unobstructed electric field at $P$ (i.e. the electric field that would have existed if the entire aperture were absent), and $C(u)$ and $S(u)$ are the Fresnel cosine and sine integrals, being defined by,

$$C(w) = \int_0^w \cos(\pi w^2/2) dw$$
$$S(w) = \int_0^w \sin(\pi w^2/2) dw$$

Here $w$ represents either of the two dimensionless variables $u$ or $v$,

$$u = y\sqrt{2(p_0 + q_0)/\lambda p_0 q_0}, \quad v = z\sqrt{2(p_0 + q_0)/\lambda p_0 q_0}$$

---

Figure 1. Geometric configuration for Fresnel diffraction for $N=5$. $a$ is the individual aperture width and $(a+b)$ is the center-to-center aperture separation. The element at $(y,z)$ is shown inside aperture 1, but it can be anywhere in any of the apertures. (Reproduced from Abedin et al. Optik Vol. 126. pp. 3743-3751, Copyright (2015) with permission from Elsevier)
The variables \(u\) and \(v\) are proportional to Cartesian coordinates \(x\) and \(y\). The limits \(u_0\) and \(u_0'\) in Eq. (2) are the values of \(u\) corresponding to two left and right edges of the central aperture, i.e. for \(y_0=-a/2\) and \(y_0=a/2\), respectively. Similarly, the limits \(v'\) and \(v\) are the values of the variable \(v\) corresponding to the lower and the upper edges of this aperture i.e. for \(z=-c\) and \(z'+=+c\) respectively.

For the rest of these Proceedings, for simplicity, we consider plane-wave illumination of the aperture. This will effectively move the source \(S\) to infinity, which can be achieved by placing \(S\) in the focal point of a positive (convex) lens and allowing the collimated light from the lens to fall on the apertures. For the simple case of plane-wave illumination, \(p_0\) is effectively infinity and, equations (4) can be simplified to,

\[
\begin{align*}
    u &= \sqrt{2/\lambda} q_u, \\
    v &= z \sqrt{2/\lambda} q_v.
\end{align*}
\]

Upon illumination from the source, the total electric field at the center \(P\) of the image plane consists of the contributions from all of the \((2n+1)\) apertures. The electric field contribution from the aperture \(I\) is given, in analogy to Eq. (2), by

\[
E_{P_I} = \frac{E}{2} \left[ C(u) + j S(u) \right]_{u_0}^{u_0'} + \left[ C(v) + j S(v) \right]_{v_0}^{v_0'} - \left[ C(v) + j S(v) \right]_{v'0}^{v'0'}
\]

where \(u_0\) and \(u_0'\) are the values of \(u\) corresponding to two edges of aperture \(I\), i.e. for \(y_0=a/2+b\) and \(y_0=-a/2+b\), respectively. Similarly, the limits \(v\) and \(v'\) are the values of \(v\) for \(z=-c\) and \(z'=+c\) respectively, as before. The electric field contribution by the aperture \(n\) is given by

\[
E_{P_n} = \frac{E}{2} \left[ C(u) + j S(u) \right]_{u_0}^{u_n} + \left[ C(v) + j S(v) \right]_{v_0}^{v_n} - \left[ C(v) + j S(v) \right]_{v_n'}^{v_0'} + \left[ C(v) + j S(v) \right]_{v_0'}^{v_n'} - \left[ C(v) + j S(v) \right]_{v_0'}^{v_n'}
\]

where \(u_n\) and \(u_n'\) are the values of \(u\) corresponding to two edges of aperture \(n\), i.e. for \(y_n=(2n-1)a/2+nb\) and \(y_n'=(2n+1)a/2+nb\), respectively. The electric field contributions by the apertures on the other side, (e.g., for the \(-n\) apertures denoted by \(E_{P_j}\) and \(E_{P_{-n}}\)) are given by equations similar to Eqs. (6) and (7), where all the limits are correspondingly negative. The net complex electric field at \(P\) contributed by all the \(N=(2n+1)\) apertures is given by the arithmetic sum of all complex amplitudes \(E_{P_0} \ldots E_{P_I} E_{P_{-1}} E_{P_{-2}} \ldots E_{P_{-n}}\); i.e., by

\[
E_N = E_{P_0} + E_{P_1} + E_{P_2} + E_{P_{-1}} + E_{P_{-2}} + \ldots + E_{P_{-n}}
\]

where \(E_N\) denotes the intensity of the unobstructed wave due to \(E_u\), i.e. \(I_0 = E_u^2\).

To calculate the intensity at any off-axis point \(P'\) on the image plane, one can fix the SOP line and instead of moving the point \(P\), one can move the entire aperture system by small amounts in the \(yz\) plane, so that the relative positions of the aperture edges in the new position and \(P\) remains unchanged. For example, to find the intensity a point \(P'\) \(1\) mm below \(P\), one can keep the screen undisturbed and move the apertures \(1\) mm upwards, and find the intensity, at illumination, lens and allowing the collimated light from the lens to fall on the apertures. For the simple case of plane-wave illumination, \(p_0\) is effectively infinity and, equations (4) can be simplified to,

\[
\begin{align*}
    u &= \sqrt{2/\lambda} q_u, \\
    v &= z \sqrt{2/\lambda} q_v.
\end{align*}
\]

For the summation notation for the Fresnel sines and cosines, we can write the total complex electric field at \(P\) as,

\[
E_N = \frac{E}{2} \left[ \sum_{j=-n}^{j=n} \left[ C(u_j) - C(u_j) \right] + j \sum_{j=-n}^{j=n} \left[ S(u_j) - S(u_j) \right] \right] \times \left[ C(v') - C(v') \right] + j \left[ S(v') - S(v') \right].
\]

The summation from \(-n\) to \(+n\) for both the Fresnel cosine and sine integrals (for the \(u\) variable only), in effect, carries out the summation of the complex electric field contributions from aperture \(-n\) to aperture \(n\), a total of \(N=(2n+1)\) apertures in the system. No such summation is involved for the \(v\) variable, since the two edges (upper and lower) of the apertures are involved in this \(z\)-direction. The net intensity at \(P\) is proportional to \((EE^*)\), i.e.,

\[
I_P = \frac{E}{2} (E_N E_N^*)
\]

where \(I_0\) denotes the intensity of the unobstructed wave due to \(E_u\), i.e. \(I_0 = E_u^2\).
These equations are the basis of calculation of the complete intensity distribution of the Fresnel diffraction pattern from an \( N \)-aperture, as will be explained next.

### 2.2 Simulation strategy and methodology

The flow chart of the algorithm is shown in fig. 2. For calculation of the electric field distribution using Eq. (10) at all the points (pixels) on the image plane, a large number of the Fresnel cosine and sine integrals will need to be evaluated quickly and efficiently. This is accomplished in the MATLAB program using the special functions \texttt{mfun(‘FresnelC’, \( i:s:f \))} and \texttt{mfun(‘FresnelS’, \( i:s:f \))}. For example, the MATLAB statement \( R=mfun(‘FresnelC’, i:s:f) \) generates an array \( R \) of the Fresnel cosine integrals with arguments starting from \( i \) and ending in \( f \), with a step size \( s \).

As shown in fig. 3 for \( N=3 \) (three apertures), the aperture was displaced (virtually) by an amount \( W \), and the extreme limits of the \( 2N \) edges of the displaced aperture were determined and compared to the corresponding limits for the undisplaced aperture to find the range of \( u \) and \( v \) values. In a more general \( N \)-aperture system, when the aperture is moved by \( W \) to the left, the ranges for the \( 2N \) edges of the \( N=(2n+1) \) apertures will be established as:

\[
\begin{align*}
\text{for } y_{1} &: (2n-1)a/2 + nb \text{ to } W+(2n-1)a/2 + nb \\
\text{for } y_{2} &: (a/2+b) \text{ to } (W+a/2+b) \\
\text{for } y_{0} &: (a/2) \text{ to } (W-a/2) \\
\text{for } y_{1} &: -(a/2-b) \text{ to } (W-a/2-b) \\
\text{for } y_{2} &: -(2n-1)a/2 - nb \text{ to } W-(2n-1)a/2 - nb
\end{align*}
\]

Array of Fresnel cosine and sine integrals need to be calculated corresponding to these input ranges by the \texttt{mfun} statements, with a given step size \( s \). By using these Fresnel arrays, the electric field or intensity dependence in the \( y \) (or \( u \)) direction can be calculated (first factor in Eq. (9)). Following a similar procedure for the entire aperture, i.e., by moving the aperture by \( W \) downwards, the \( z' \) and \( z \) ranges are determined: \( c \) to \( (W+c) \) for \( z' \) and \( -c \) to \( (W-c) \) for \( z \). Four arrays of Fresnel integrals are to be evaluated for these two ranges, giving numerical values for the calculation of the second factor in Eq. (9). Further details of the computation method and explanation of the MATLAB program is found in Ref. (8) and Ref. (5).
3. SIMULATION RESULTS

3.1 General

To use the program, the values of the required inputs, i.e. aperture parameters a and b, height c, image area W, step size s, the aperture-image distance q₀ (all in mm), the wavelength λ (in nm), the exposure and the number N of the slits (apertures) are input through a GUI (fig. 4a). The exposure factor ex is used to adjust the apparent visual intensity in the generated image. The computer generated Fresnel image due to N-slits for wavelength λ=500 nm, q₀=400mm, image area W=5mm, aperture width a=0.1 mm, b=0.1 mm (with aperture separation 0.2 mm), c=4mm, s=0.01mm and N=31 is shown in fig. 4b. Since (a+b)=0.2mm, this system is equivalent to an amplitude grating N=31 with 5 lines/mm. The generated diffraction image resembles that of a grating in the near-field, with Fresnel-like characteristics exhibited in the top and bottom edges. Some interference effects can be seen in the left and right edges of the image.

To find the effect of a change of parameters on the diffracted images, we generated a series of images. In the first series, the aperture separation (a+b) was made successively smaller, while keeping the aperture width (a=0.5mm) constant. A series of diffraction images were simulated as shown in figs 5(a)-(d) for a value of N=7. In fig. 5(a), in the region between the apertures, no significant interference could be seen between the light from each of the aperture diffraction patterns. But, in figs 5(b) and (c), some distinct interference between diffracted light from the individual apertures was observed as the separation of the slits is further reduced. As the separation was decreased to zero [fig. 5(d) for b=0], a typical single aperture Fresnel diffraction image was obtained for an aperture size 3.5mmX3mm.

In the next series, for N=7, the width a of the apertures was made successively narrower while their separation (a+b) was kept constant at 1.5 mm [fig. 6]. From figs 6(a) and (b), as the apertures become narrower, the light is diffracted over...
a wider region, and some interference between diffracted light can be detected. In figs. 6(c) and (d), at smaller aperture widths, strong mutual interference between diffracted light has generated fringes which look like typical Young-like fringes. Light from each aperture is seen to be diffracted over a wider portion of the image area. In figs 6(c) and (d), fringe density is the same, since this depends only on aperture separation (constant at 1.5mm).

3.2 Comparison with N-slit far-field diffraction

If the aperture-screen distance \( q_0 \) is increased, it is well-known in optics that a gradual transition from the Fresnel (near-field) regime to Fraunhofer regime (far-field) should occur. As a general rule of thumb, if \( D \) is the lateral dimension of an aperture, then Fresnel diffraction occurs if \( D \) satisfies the following relation:

\[
q_0 < \frac{D^2}{\lambda}.
\]

Otherwise, Fraunhofer diffraction should be observed. In the next simulations, we use a practical diffraction grating with \( a=0.001 \) and \( b=0.003 \) (with aperture separation \( a+b=0.004 \) mm, i.e. 250 lines/mm). We selected the following parameters for an 11-slit system, \( \lambda=500 \) nm, \( c=2 \) mm, image area size \( W \) variable, \( s=0.01 \) mm or \( 0.05 \) mm, as appropriate. We then increased the aperture screen distance \( q_0 \) in steps from 1 mm to 100 mm while keeping other parameters constant [fig. 7]. A transition from the Fresnel regime to Fraunhofer regime is clearly observed, being consistent with the predictions of Eq. (11). For example, for a 11-slit system with \( (a+b)=0.004 \) mm, if we take the total lateral dimension of the slit system to be \( D=0.04 \) mm, then \( D^2/\lambda \) is about 3.2 mm. We expect the diffraction to be Fresnel-like if \( q_0 \) is less than 3.2 mm. If, \( q_0 > (D^2/\lambda)=3.2 \) mm, and we expect the diffraction to be Fraunhofer-like. In between these extremes, a transition from Fresnel to Fraunhofer regime should occur [as in figs 7(c) and 7(d)]. The usual mathematical analysis of an \( N \)-slit system\(^1\) deals with the \( N \)-slit pattern only in the far-field Fraunhofer regime. The far-field diffraction pattern of an \( N \)-aperture system in the image plane \((Y, Z)\) can be exactly calculated by an analytical formula:\(^2\)

\[
I(Y, Z) = I_0 \left( \frac{\sin \beta}{\beta^2} \right) \left( \frac{\sin \gamma}{\gamma^2} \right) \left( \frac{\sin \alpha}{\alpha^2} \right) \left( \frac{\sin \gamma}{\gamma^2} \right),
\]

\( \alpha, \beta, \gamma \)

Figure 5. Computer-simulated Fresnel diffraction images for decreasing (center-to-center) aperture separation \((a+b)\), while keeping the aperture width \((a=0.5\text{mm})\) constant. for \( N=7 \). Other parameters are: \( \lambda=500 \) nm, \( q_0=400 \) mm, \( W=8 \) mm, \( s=0.01 \) mm and \( c=3 \) mm. Fig. 5(d) corresponds to \( b=0 \) (single aperture). (Reproduced from Abedin et.al. Optik Vol.126. pp. 3743-3751, Copyright (2015) with permission from Elsevier)
Figure 7. Computer simulated Fresnel diffraction images for a 7-aperture system. Aperture-screen distance $q_0$ is increased in steps, while keeping other parameters constant, with $a=0.001$ mm, $b=0.003$ mm, $c=2$ mm, $\lambda=500$ nm. (a) and (b) are in the Fresnel regime, while (c) and (f) are in the Fraunhofer regime, according to Eq. (12).

In the usual analysis of far-filed Fraunhofer diffraction from an $N$-slit system, it is shown that the principal maxima of the diffraction pattern occurs when the $[\sin N\gamma\sin \gamma]^2$ factor have maximum values, i.e. when,

$$(a+b) \sin \theta = m\lambda .$$  \hspace{1cm} (14)

Here, $\theta$ is the diffraction angle and $m$ is the diffraction order. For $m=0$, we have the zeroth-order, for $m=1$, the first order, and so on. In addition to these principal maxima, several much weaker secondary maxima, and several minima of zero intensity are predicted to occur with equal separation between the principal maxima. If $N$ is the number of grating slits (apertures), then the number of secondary maxima is equal to $(N-2)$, and the number of minima is equal to $(N-1).$ To make comparisons, we reproduce fig. 7(f) next, with the intensity values multiplied by 10 to bring out the faint details. We observe 7 principal maxima, with $m=0$ (at $Y=0$), with $m=1$ (at $Y=12.5$ mm), $m=2$ (at $Y=25.03$ mm), $m=3$ (at $Y=37.59$ mm), with $m=-1$ (at $Y=-12.5$ mm), $m=-2$ (at $Y=-25.0$ mm), and $m=-3$ (at $Y=-37.5$ mm). In addition, 9 (i.e., $N-2$) secondary maxima can be discerned between the principal maxima at the zeroth order and the first and the second orders, along with 10 minima of zero intensity, interspersed between them. The values of $\sin \theta$ can be estimated by finding the ratios ($Y/R$, with $R=100$ mm) from the computer-simulated image (Fig. 8). These calculated values are: 0.125 (for $m=+1$ and $m=-1$), 0.2503 (for $m=+2$ and $m=-2$), and 0.375 (for $m=+3$ and $m=-3$) From Eq. (14), on the other hand, for $\lambda=500$ nm, $(a+b)= 0.004$ mm, and the following values of $\sin \theta$ are calculated: 0.125 (for $m=+1$ and $m=-1$) 0.250 (for $m=+2$ and $m=-2$), and 0.375 (for $m=+3$ and $m=-3$). We therefore conclude: the computer-simulated results of the diffraction field, in the far-filed Fraunhofer limit, precisely agree, both qualitatively and quantitatively, with the results from the exact Fraunhofer theory, as represented by Eqs. (12) and (14). This agreement gives strong support to the validity and correctness of the present simulation methods.
4. DISCUSSIONS

In this Proceedings, we describe a technique to simulate the near-field diffraction image of a $N$-aperture system or, equivalently, an amplitude diffraction grating, using the iterative Fresnel Integrals method. The theoretical background and the implementation of the algorithm are described. Some examples of computer-simulated images as well as the practical case of a 250 line/mm amplitude diffraction grating are presented in the results. For the simulations, an Intel core-i7 PC having 4GB of RAM and running at 3.4 GHz was used. The simulation time for a typical 2000X2000 pixel image for $N=11$ is less than 45 seconds. The time depends on the computer configuration used, or, on the image size. The expected transition from Fresnel to Fraunhofer regions can be clearly observed under the appropriate conditions. For example, when the aperture-screen distance $q_0$ is increased sufficiently into the far-field regime, the predicted principal maxima in the Fraunhofer limit are observed at the expected positions for different diffraction orders, and the predicted number of secondary maxima and minima agrees with the far-field Fraunhofer theory. The simulation techniques and the resulting virtual experiments provide a simple and easy-to-use method to study the complex phenomenon of diffraction from amplitude diffraction gratings. It can be used to teach diffraction phenomena from amplitude gratings in a classroom setting, where the virtual experiments can be performed in real-time in front of the students.

REFERENCES