An analytically simple and computationally efficient Gaussian beam mode-decomposition approach to classical diffraction theory

Zhizhao Xiao,1,∗ R. Nicholas Lanning,1 Mi Zhang,2 Irina Novikova,2 Eugeni E. Mikhailov,2 and Jonathan P. Dowling1,3,4,5
1 Hearne Institute for Theoretical Physics, and Department of Physics & Astronomy, Louisiana State University, Baton Rouge, Louisiana 70803, USA
2 Department of Physics, College of William & Mary, Williamsburg, Virginia 23187, USA
3 NYU-ECNU Institute of Physics at NYU Shanghai, Shanghai 200062, China
4 CAS-Alibaba Quantum Computing Laboratory, USTC, Shanghai 201315, China
5 National Institute of Information and Communications Technology, 4-2-1, Nukui-Kitamachi, Koganei, Tokyo 184-8795, Japan

We present a method of Gaussian-beam-mode decomposition to calculate classical diffraction of optical beams by apertures. This method offers a entirely different approach to examine the classic problem. Although our method is based on a very straightforward setup, it is surprisingly effective. We validate our method by comparing its results with those of Kirchhoff’s full diffraction formula. Not only does our method have a simple and organized analytical framework, it also offers significant computational advantage.

Introduction.—Traditional scalar diffraction theory for optical beams is based on the Huygens-Fresnel principle and Kirchhoff’s diffraction formula. There have been numerous studies [1–4] on the subject. In spite of some recent developments [5–11], the approach to classical diffraction is still based on these same principles. Scenarios under conditions such as a special field depth or a special symmetry in the aperture and source field have been studied [12–22], but the applications are not as general as Kirchhoff’s diffraction formula, and accuracy in the calculations needs improvement.

To simplify the calculation, the diffracted field is divided into the near field (known as Fresnel diffraction) or the far field (known as Fraunhofer diffraction), so that certain approximations can be applied. This arrangement is far from ideal, in the sense that we trade one problem with a complicated calculation for two problems, which require separate, simpler calculations. Additionally, there is no simple solution for the intermediate field between the near and the far field, and there is no clear picture describing how the near or the far field would transition into the other.

We propose a different approach for examining the diffraction problem, which is based on Gaussian-beam mode decomposition. While previous studies [23–25] have been carried out on special cases where certain symmetry or other restrictions on the optical beams or apertures are assumed, here we investigate the general case. Our method, briefly discussed in Ref. [26], takes advantage of the fact that spatial modes of a Gaussian beam, Laguerre-Gaussian (LG) modes and Hermite-Gaussian (HG) modes, each form a complete orthonormal basis in any given plane perpendicular to the beam axis. As a consequence, any source amplitude diffracted through an aperture can be expressed as a linear superposition of LG or HG modes. From now on we mainly focus on the LG modes, but for HG modes similar results can be derived.

Gaussian-beam-mode decomposition method.—Gaussian-beam modes are the solutions that satisfy the free-space Maxwell’s equations within the paraxial approximation. Specifically they are called Laguerre-Gaussian (LG) modes in cylindrical coordinates and Hermite-Gaussian (HG) modes in Cartesian coordinates. The exact mathematical expression of LG modes is [27].

\[
\begin{align*}
\psi_{l,p}(r,\phi,z) &= \frac{C_{l,p}^{\text{LG}}}{w(z)} \left( \frac{r}{w(z)} \right)^{|l|} \exp \left( -\frac{r^2}{w^2(z)} \right) \\
&\times L_{|l|}^{l} \left( \frac{2r^2}{w^2(z)} \right) \exp \left( -ikR(z) \right) \exp (-ikz) \exp (il\phi) \exp [i(2lp + |l| + 1)\zeta(z)],
\end{align*}
\]

(1)

where \( r, \phi \) and \( z \) are cylindrical coordinates; \( l \) and \( p \) are the azimuthal and radial indices, which are integers; \( p \geq 0 \); \( C_{l,p}^{\text{LG}} = \sqrt{\frac{2}{\pi l!} \frac{p!}{(l+p)!}} \) is a normalization constant; \( L_{|l|}^{l} \) is the associated Laguerre polynomial; \( \lambda \) is the wavelength; \( k = 2\pi/\lambda \) is the wave number; \( w(z) = w_0 \sqrt{1 + \left( \frac{z}{z_R} \right)^2} \) is the beam waist; \( w_0 \) is the beam waist at the beam focus; \( z_R = \frac{\lambda^2}{4\pi} \) is the Rayleigh range; \( R(z) = |1 + \left( \frac{z}{z_R} \right)^2| \) is the radius of curvature; \( \zeta(z) = \arctan \left( \frac{z}{z_R} \right) \) is the Gouy phase. Along the beam axis the beam waist will become wider or narrower, while the shapes of the intensity profiles remain similar. Note that we have chosen the \( z \) axis to be the beam axis and \( z = 0 \) to be the beam focus.

Also, the three parameters (the beam focus position, the wave number \( k \) and the beam waist at the beam focus \( w_0 \)) determine the entire LG mode orthonormal basis. Changing any one of the three parameters will give a different basis. We always have the freedom of choosing a particular set of the three parameters, in other words...
The underlying principle of the classical diffraction is its quantum version \[26\], where \( B_{l,p} \) are interpreted as coherent state amplitudes.

We now consider a general aperture, which is not necessarily perpendicular to the beam axis nor flat. We model the diffraction problem as follows. Let \( S_A \) be the surface of the aperture where the optical beam is absorbed and \( S_f \) be the surface of the opening of the aperture, which is illuminated by an amplitude distribution \( u_I(r,\phi,z) \). \( S_A \cup S_f \) forms a continuous, entire surface stretching to infinity in all transverse directions. Let us arrange the diffracted field amplitude coefficients \( B_{l,p} \) into a column vector:

\[
B = \begin{bmatrix} B_{l_1,p_1} & B_{l_2,p_2} & \ldots & B_{l_3,p_3} & B_{l_4,p_4} & \ldots \end{bmatrix}^T, \tag{4}
\]

where \( l_i, p_i \) are row indexes.

Together with the Gaussian modes, the diffracted field amplitude is determined by \( B \) which can be derived to be given by

\[
M = M^{-1}C, \tag{5}
\]

where \( M \) is a matrix whose elements are given by

\[
M_{l,p;l',p'} = \int_{S_A \cup S_f} u^*_I(r,\phi,z) \times u_I(r',\phi',z)dS, \tag{6}
\]

where \( l, p \) are row indexes, \( l', p' \) are column indexes and \( dS \) is the infinitesimal surface projected on the transverse plane. The vector \( C \) is a column vector whose elements are given by

\[
C_{l,p} = \int_{S_f} u^*_I(r,\phi,z) \times u_I(r,\phi,z)dS. \tag{7}
\]

Let us suppose an optical beam with a superposition of Gaussian modes propagates through free space. This amounts to a fictitious surface of \( S_f \), which is illuminated by the superposition of Gaussian modes themselves. Since the free space has no absorption, \( S_A = \emptyset \).

With Eqs. (5–7), we can verify that regardless how we...
construct the shape of \( S_l \), so long it stretches to infinity in every perpendicular direction, the coefficients of the Gaussian modes remain the same, as they are supposed to. From this seemingly trivial scenario, we can see that all Eqs. (5–7) are needed for a general aperture.

The most common situation is that the aperture is confined in a flat plane perpendicular to the beam axis. In this case, \( M \) and its inverse \( M^{-1} \) simplify to the identity matrix, and \( B \) simplifies to \( C \), and mode coefficients can be calculated from Eq. (7) alone. This gives us a new and intuitive way to interpret Eq. (5): \( M^{-1} \) can be seen as the correction to \( C \), due to the fact that orthonormal condition of Gaussian modes needs to be corrected when an aperture surface is neither flat nor perpendicular to the beam axis.

We have shown that our method offers a different way of calculating the diffracted field amplitude by making use of Eqs. (2, 5) instead of Kirchhoff’s diffraction formula. Our method is developed based on the principle that the source beam and the diffracted beam should have the same boundary field amplitude at the aperture. This is different from Kirchhoff’s diffraction formula, which is based on Huygens-Fresnel principle.

Kirchhoff’s diffraction formula takes the form of

\[
\begin{align*}
  u(r, \phi, z) &= \frac{i}{2\lambda} \int_{S_l} u_l(r', \phi', z') e^{i kr} \left( \cos \chi_1 + \cos \chi_2 \right) dS', \\
  (8)
\end{align*}
\]

where \( \rho \) is the distance between the position of the diffracted field \((r, \phi, z)\) and the position of illumination field \((r', \phi', z')\), \( \chi_1 \) is the angle between aperture opening surface normal vector at \((r', \phi', z')\) and beam axis \(z\), and \( \chi_2 \) is the angle between aperture opening surface normal and vector from \((r', \phi', z')\) to \((r, \phi, z)\). (a) Our method has a neat and clear mathematical structure—Eqs. (2, 5). It also has a simpler and more efficient computational formalism. The coefficients of the Gaussian modes \( B_{l,p} \) are computed via Eq. (5) that are integrals. Once they are calculated, they can be used again and again to calculate the diffracted field at any position, using Eq. (2). Calculating the amplitude in different positions in space simply means repeatedly doing the summations. This can be done very efficiently. (b) For Kirchhoff’s diffraction formula, near- and far-field approximations generally result in different outcomes, causing difficulties if both near and far field need to be examined. On the other hand, our method offers one unified and already simple computing framework, which applies to all near, far, and intermediate fields. (c) For our method, to yield completely accurate results, all (infinite) orders of Gaussian modes in Eqs. (2, 5) must be accounted for, which means \( l \in (-\infty, +\infty) \) and \( p \in (0, +\infty) \). But the coefficients for higher-order modes usually drop off rapidly. According to Eq. (5), every mode coefficient is calculated individually, meaning the calculation of each coefficient is independent of the others. If higher accuracy is needed, we can simply calculate additional higher-order coefficients on demand and add the amplitude to the existing lower-order mode amplitudes. In other words, we can make use of the lower-accuracy result to obtain the higher accuracy result.

Demonstration of the validity.—We now compare our method to Kirchhoff’s diffraction formula. Let us examine a simple case in which a circular aperture with radius \( a \) is placed at the \( z = z_1 \) plane and centered on the beam axis, and the aperture is illuminated by a plane wave traveling along the beam axis with normalized amplitude and wavelength \( \lambda \). The surface in Eq. (5) now has the form of \( S_l = \{ r < a; z = z_1 \} \).

Kirchhoff’s diffraction formula (in cylindrical coordinates) takes the form of [27]

\[
\begin{align*}
  u(r, z) &= i2\pi Ne^{-i\pi N(r/a)^2} \\
  &\quad \int_0^a r_0 u_{l,0} e^{-i\pi N(r/a)^2} \frac{2\pi Nrr_0}{a^2} J_0(\frac{2\pi Nrr_0}{a^2})dr_0, \\
  (9)
\end{align*}
\]

where the source field (plane wave) amplitude \( u_l(r, \phi, z_1) \) is set to be uniformly \( u_1 \). The diffracted field amplitude at plane \( z \) and radial distance \( r \) is \( u(r, z) \). \( N \equiv \frac{a^2}{(z-z_1)^2} \) is the Fresnel number and \( J_0 \) is the zeroth order Bessel function. The region where the Fresnel number \( N << 1 \) is considered as far-field diffraction, and Kirchhoff’s diffraction formula simplifies to,

\[
\begin{align*}
  u(r, z) &= i\pi Ne^{-i\pi N(r/a)} \frac{2J_1(2\pi N r/a)}{2\pi N r/a}, \\
  (10)
\end{align*}
\]

where \( J_1 \) is the first order Bessel function. This result is the well-known Airy pattern. The region where the Fresnel number \( N >> 1 \) is considered near-field diffraction. We can see that there are a very limited number of situations where we can simplify Kirchhoff’s diffraction formula, as there is no simplification even for this uniformed source field and highly symmetrical aperture in the intermediate field.

For all non-zero \( l \), \( B_{l,p} = 0 \), since the source-field amplitude and the aperture are both cylindrically symmetric. Therefore only \( l = 0 \) modes contribute to diffracted field amplitude. Furthermore, in the numerical simulation we need to set a maximum \( p \) index \( p_{\text{max}} \) and all modes exceeding \( p_{\text{max}} \) will be ignored. Therefore the result is reduced to,

\[
\begin{align*}
  u_{\text{eff}}(r, \phi, z) &= \sum_{p=0}^{p_{\text{max}}} B_{0,p} \times u_{0,p}(r, \phi, z). \\
  (11)
\end{align*}
\]

Since the diffracted field has cylindrical symmetry, we need to only examine the intensity in terms of radius in near, intermediate and far field regions, as shown
in Fig. 2(a)–2(d). In all these regions we can see that our method fits very well with Kirchhoff’s diffraction formula. Notice in the far-field, far-axis region (shown in Fig. 2(d)) the result from our method with cutoff \( p_{\text{max}} = 80 \) is deviating from the Kirchhoff’s formula result (Airy rings). The further we go to the far-axis region (larger radius) the more inaccurate the \( p_{\text{max}} = 80 \) result becomes. This inaccuracy can be easily improved by including more higher-order LG modes, as the \( p_{\text{max}} = 300 \) result shows. A similar phenomenon is observed in the far-axis region of near field and intermediate field as well: the further from the beam axis, the more modes are needed to be accurate. The explanation of this phenomenon is two fold: (a) lower order LG or HG modes concentrate a large part of their intensity near the beam axis, while the higher-order modes have a much more spread-out intensity distribution in both near-axis and far-axis. Therefore, lower-order modes are the dominating factor in the near-axis region, while higher-order modes must be included in the far-axis. (b) In general, the far-axis intensity is lower than near-axis, as one can see from the ranges of vertical axis of Fig. 2(c) and Fig. 2(d). Therefore inaccuracy is more apparent in far-axis than in near-axis, making the far-axis result more sensitive to error.

Though the Kirchhoff integral simplification to the Airy pattern is accurate in the far field (Fig. 2(c)), it performs poorly in the intermediate field (Fig. 2(b)). It performs even worse in the near field (which is not shown in Fig. 2(a) because it is too inaccurate). This provides evidence for our point that our method has the same range of application as the full Kirchhoff’s formula.

The essence of our method is to represent the diffracted field in a LG (or HG) basis. If all LG modes are included, the method gives a completely accurate result. However, for practical reasons, only limited LG modes \( l = 0, p \leq p_{\text{max}} \) are included, and \( u_{\text{eff}}(r, \phi, z) \) is employed to approximate \( u(r, \phi, z) \). To examine the closeness between the amplitude distributions of \( u_{\text{eff}}(r, \phi, z) \) and \( u(r, \phi, z) \) across any entire perpendicular diffraction plane \( z = z_1 \), we introduce an amplitude spatial fidelity defined as

\[
F(u_{\text{eff}}, u) = \frac{\int_{S:z_1} u_{\text{eff}}^* u \, dS}{\int_{S:z_1} u_{\text{eff}}^2 \, dS \times \int_{S:z_1} u^2 \, dS},
\]

which is closer to one if \( u_{\text{eff}}(r, \phi, z) \) resembles \( u(r, \phi, z) \) more. This fidelity does not depend on the position of \( z_1 \), therefore we can examine the plane immediately after the aperture (\( z_1 = z_1^\text{eff} \)). With our method, the calculation of amplitude spatial fidelity is conveniently simplified to:

\[
F(u_{\text{eff}}, u) = \sum_{p=0}^{p_{\text{max}}} |B_0,p|^2 \int_{0 \leq \rho < a} |u_r|^2 2\pi r \, dr.
\]

The fidelity cannot be so conveniently calculated with Kirchhoff’s diffraction formula.

We have the freedom of choosing the Gaussian mode basis. However, some Gaussian mode bases work better than others. There is no black-or-white rule regarding the choice of basis, so we only give some general guidelines here. To be able to better depict the variation of the diffracted field across various transverse positions, we should choose a basis with a small waist. In addition, in order to better depict the variation of the diffracted field along the beam axis, especially in the near field, we should choose a basis with a smaller Rayleigh range, which also requires a smaller waist. However, the smaller the waist we choose, the more modes we will generally need to include. To illustrate this, we can examine \( F(u_{\text{eff}}, u) \) versus the number of modes included. As the number of modes increases in one particular basis, \( F(u_{\text{eff}}, u) \) converges to one, typically in two distinct regimes: sharp converging and long tail regimes, shown in Fig. 2(e). Once a basis is chosen, sufficient modes need to be included to at least cover the sharp converging regime, otherwise a significant part of the amplitude will be unaccounted for. For this to happen, we have found the number of modes needed is inversely proportional to the
square of waist size. The area one Gaussian mode can cover is proportional to the square of waist size. If we use a smaller waist size, to fit in the same illuminated area, the number of modes needed must increase.

On one hand, \(F(u_{\text{eff}}, u)\) is useful for assessment of the accuracy of \(u_{\text{eff}}(r, \phi, z)\) globally across any perpendicular diffraction plane. On the other hand, to assess the accuracy of our effective diffraction field intensity calculation locally in any specific position, we can use the following error estimation,

\[
\text{Err}(r, \phi, z; p_{\text{max}}) = \left| \int_{0<r<a} |u_I|^2 2\pi r dr - \sum_{p=0}^{p_{\text{max}}} |B_{0,p}|^2 \right| \times \max\{|u_{0,p}(r, \phi, z)|^2\}, \quad p > p_{\text{max}},
\]

which works well in the limit of \(F(u_{\text{eff}}, u) \rightarrow 1\). If, in a certain region, the error becomes comparable to the intensity, it suggests that the modes currently included are no longer adequate. The error bars in Fig. 2(d), calculated with Eq. (13), show that the error bars provide effective warning that 80 modes are no longer sufficient in this far axis region, and by increasing \(p_{\text{max}}\) from 80 to 300 in this case, we can reduce the error and achieve greater accuracy.

**Application to an Example with No Symmetry.**—As mentioned previously, our method can be applied to any source field and any aperture. In the absence of symmetry, Kirchhoff’s formula generally cannot be simplified and the power of our method becomes even more apparent. In this example, we are using a off-centered \(l = 2, p = 2\) LG mode as the source field (shown in Fig. 1(a)) and a rectangular aperture centered on the beam axis. In this setup, there is no symmetry that is helpful for simplification of Kirchhoff’s formula. We again compare the intensity patterns generated by our method and by Kirchhoff’s formula, shown in Figs. 1(b)–1(f). Our method is validated again, as its results matches the results from Kirchhoff’s formula. In each of the 2D plots, there are 100 \(\times\) 100 data points. Due to the computational advantages of our method, generating its plots takes only a small fraction of time as the plots with Kirchhoff’s formula. Should we need to generate more data points, the superiority of our method becomes more apparent.

**Conclusion.**—We have developed the Gaussian-mode decomposition method to calculate classical diffraction of optical beams through apertures of arbitrary shape. We have explained the setup of this method and demonstrate the effectiveness and efficiency of our method. We validate our method by comparing its result with that of Kirchhoff’s diffraction formula. In addition, we have established a proper way of accessing accuracy. We have shown that not only our method has a simple and organized analytical framework, it also offers significant computational advantage over the traditional Kirchhoff’s diffraction formula.

The authors would like to acknowledge support from the the Air Force Office for Scientific Research, the Army Research Office, the Defense Advanced Projects Agency, the National Science Foundation, and the Northrop Grumman Corporation.

---

* xiaozhihao@hotmail.com

[1] John David Jackson. *Classical electrodynamics*. John Wiley & Sons, 2007.

[2] Max Born and Emil Wolf. *Principles of optics: electromagnetic theory of propagation, interference and diffraction of light*. Elsevier, 2013.

[3] Christoffel Jacob Bouwkamp. *Diffraction theory*. *Rep. Prog. Phys.*, 17(1):35, 1954.

[4] Arnold Sommerfeld. *Lectures on theoretical physics: Optics*. Academic Press, 1954.

[5] DA Holmes, JE Korka, and PV Avizionis. Parametric study of apertured focused gaussian beams. *Applied optics*, 11(3):565–574, 1972.

[6] Virendra N Mahajan. Uniform versus gaussian beams: a comparison of the effects of diffraction, obscuration, and aberrations. *JOSA A*, 3(4):470–485, 1986.

[7] V Nourrit, P Chancour, et al. Propagation and diffraction of truncated gaussian beams. *JOSA A*, 18(3):546–556, 2001.

[8] Emmanuel M Drège, Neal G Skinner, and Dale M Byrne. Analytical far-field divergence angle of a truncated gaussian beam. *Applied optics*, 39(27):4918–4925, 2000.

[9] Leroy D Dickson. Characteristics of a propagating gaussian beam. *Applied Optics*, 9(8):1854–1861, 1970.

[10] Richard G Schell and George Tiras. Irradiance from an aperture with a truncated-gaussian field distribution. *J. Opt. Soc. A.*, 61(1):31–35, 1971.

[11] JP Campbell and LG DeShazer. Near fields of truncated-gaussian apertures. *J. Opt. Soc. A.*, 59(11):1427–1429, 1969.

[12] Charles Campbell. Fresnel diffraction of gaussian laser beams by circular apertures. *Optical Engineering*, 26(3):263270, 1987.

[13] Gadi Lenz. Far-field diffraction of truncated higher-order laguerre-gaussian beams. *Optics communications*, 123(4-6):423–429, 1996.

[14] Hakan Urey. Spot size, depth-of-focus, and diffraction ring intensity formulas for truncated gaussian beams. *Applied optics*, 43(3):620–625, 2004.

[15] Kazumasa Tanaka, Nobuhiro Saga, and Hiromu Mizokami. Field spread of a diffracted gaussian beam through a circular aperture. *Applied optics*, 24(8):1102–1106, 1985.

[16] Baida Lü, Bin Zhang, and Bangwei Cai. Focusing of a gaussian schell-model beam through a circular lens. *Journal of Modern Optics*, 42(2):289–298, 1995.

[17] DS Burch. Fresnel diffraction by a circular aperture. *American Journal of Physics*, 53:255–260, 1985.

[18] CJR Sheppard and M Hrynycz. Diffraction by a circular aperture: a generalization of fresnel diffraction theory. *JOSA A*, 9(2):274–281, 1992.

[19] O Mata-Mendez and Frnando Chavez-Rivas. Diffraction of gaussian and hermite–gaussian beams by finite grat-
ings. *JOSA A*, 18(3):537–545, 2001.

[20] PL Overfelt and CS Kenney. Comparison of the propagation characteristics of bessel, bessel–gauss, and gaussian beams diffracted by a circular aperture. *JOSA A*, 8(5):732–745, 1991.

[21] G Oluremi Olaofe. Diffraction by gaussian apertures. *J. Opt. Soc. A.*, 60(12):1654–1657, 1970.

[22] P Belland and JP Crenn. Changes in the characteristics of a gaussian beam weakly diffracted by a circular aperture. *Appl. Opt.*, 21(3):522–527, 1982.

[23] Raimondas Čiegis, Galina Šilko, and Aleksandr Dement’ev. A tool for modeling optical beam propagation. *Informatica*, 13(2):149–162, 2002.

[24] Neil Trappe, J Anthony Murphy, and Stafford Withington. The gaussian beam mode analysis of classical phase aberrations in diffraction-limited optical systems. *European journal of physics*, 24(4):403, 2003.

[25] Daniel Brown, RJE Smith, and Andreas Freise. Fast simulation of gaussian-mode scattering for precision interferometry. *Journal of Optics*, 18(2):025604, 2016.

[26] Zhihao Xiao, R. Nicholas Lanning, Mi Zhang, Irina Novikova, Eugeniy E. Mikhailov, and Jonathan P. Dowling. Why a hole is like a beam splitter: A general diffraction theory for multimode quantum states of light. *Phys. Rev. A*, 96:023829, Aug 2017.

[27] A. Siegman. *Lasers*. University Science Books, Sausalito CA, 1986.