Diffraction effects on broadband radiation: formulation for computing total irradiance

Eric L. Shirley

I present a formulation for treating diffraction effects on total irradiance in the case of a Planck source; earlier work generally depended on calculating diffraction effects on spectral irradiance followed by summation over spectral components. The formulation is derived and demonstrated for Fraunhofer diffraction by circular apertures, rectangular apertures and slits, and Fresnel diffraction by circular apertures. The prospects for treating other sources and optical systems are also discussed. © 2004 Optical Society of America

OCIS codes: 050.1960, 120.5630.

1. Introduction

Diffraction of electromagnetic radiation at the edges of apertures and lenses leads to losses or gains of flux in radiometry that are not accounted for in geometrical optics. Furthermore, practical radiometry often relies on intrinsically broadband sources: Planck-like sources, such as stars, idealized laboratory blackbodies, other, imperfect blackbody radiators that are frequently encountered in nature and laboratory environments, and synchrotrons. For such sources, assessing diffraction effects on the position-dependent total irradiance of one's detector requires summation over spectral components, each of which is affected differently by diffraction.

Traditionally, diffraction theory is formulated for monochromatic radiation. Thus, diffraction effects have been treated by first considering the spectral irradiance at various wavelengths and subsequently summing this quantity over wavelength. Diffraction effects on spectral irradiance are traditionally computed in some version of the Kirchhoff method, which solves the wave equation approximately by Green's function techniques.[1] This can lead to cumbersome integrals with highly oscillatory integrands. Even after such integration, the diffraction effects on spectral irradiance can oscillate with wavelength λ, so that care might be needed when one is summing over wavelength in order to obtain the total irradiance.

Because of the above difficulties, this work circumvents such a two-step process as follows. The squared Kirchhoff integral is rewritten as the Fourier transform of the autoconvolution of the distribution of total path length that light can travel from source to detector, such that path-length differences are associated with a wavelength-dependent complex phase shift when a wave interferes with itself. The then-trivial integration over wavelength (or angular wave number) weighted by source spectral radiance is carried out, which removes almost all oscillatory behavior of subsequent integrands and so simplifies their integration. Simplified integrals over total path length or path-length differences are then more easily evaluated. In this way this work also naturally treats and includes all spectral components output by a broadband source simultaneously and takes advantage of the common spatial aspects of light propagation that are wavelength invariant.

The present work only considers Planck sources, for which spatial and spectral properties of radiance are decoupled. That is, the spectral radiance $L_\lambda(\lambda, T_s, \mathbf{r}_s, \hat{\Omega})$ (power emitted per unit wavelength per steradian per projected unit area of source) may be expressed as the product of one factor that describes its spatial properties and one factor that describes its spectral properties:

$$L_\lambda(\lambda, T_s, \mathbf{r}_s, \hat{\Omega}) = f_1(\mathbf{r}_s, \hat{\Omega}) f_2(\lambda, T_s). \tag{1}$$

Here λ is the wavelength, T_s is the source temperature (assumed to be the same value everywhere), $\mathbf{r}_s$ is a point on the source, and $\hat{\Omega}$ is the emission direction.

E. L. Shirley (eric.shirley@nist.gov) is with the Optical Technology Division, National Institute of Standards and Technology, 100 Bureau Drive, MS 8441, Gaithersburg, Maryland 20899-8441.

Received 16 April 2003; revised manuscript received 11 February 2004; accepted 18 February 2004.

0003-6935/04/132609-12$15.00/0

For a Lambertian source, the radiance does not depend on $\hat{\Omega}$. Furthermore, for the sources we consider, the radiance does not depend on $\mathbf{r}_s$, except in that $f_1(\mathbf{r}_s, \hat{\Omega})$ has one value everywhere on the area of an extended source and is zero elsewhere. However, if $f_1(\mathbf{r}_s, \hat{\Omega})$ were to vary as a function of $\mathbf{r}_s$ and $\hat{\Omega}$, the salient aspects of this work would remain valid. It remains to be seen whether this work could be adapted for similar analysis involving synchrotron radiation.

This work considers a broad class of optical setups, regarding how diffraction affects the relationship between the source temperature T_s and irradiance at a point $\mathbf{r}_d$ in the detector plane, $E(\mathbf{r}_d, T_s)$, which is the power incident per unit area. The dependence of $E(\mathbf{r}_d, T_s)$ on T_s is interwoven with its dependence on the location of $\mathbf{r}_d$ and on the geometrical aspects of all optics between and including the source and detector, which affect propagation of light because of geometrical-optics effects and diffraction effects. Note that geometrical optics describes the propagation of light in a manner that is independent of wavelength λ, whereas the spectral dependence of diffraction effects precludes separating the dependences of $E(\mathbf{r}_d, T_s)$ on $\mathbf{r}_d$ and on T_s.

The most well-known diffraction problems include Fraunhofer diffraction by rectangular apertures and slits and circular apertures and Fresnel diffraction by circular apertures.[2] In Section 3 of this work the diffraction effects for these well-known problems are analyzed in particular. Concluding remarks and a technical appendix follow.

2. Formulation

For simplicity, all that follows will rely on the scalar Fresnel–Kirchhoff treatment of diffraction in the paraxial (Gaussian optics) approximation.[3] This treatment is limited, because it can fail to describe polarization effects and various aberrations, including the focal shift.[4] However, the analogous treatment of geometrical optics leads to extremely simple approximate expressions for the throughput of optical systems, and the small-wavelength behavior of paraxial Fresnel–Kirchhoff results approaches this limit. Therefore the difference between a paraxial Fresnel–Kirchhoff result and its geometrical-optics counterpart can often provide a reasonable assessment of diffraction effects on the behavior of an optical system.

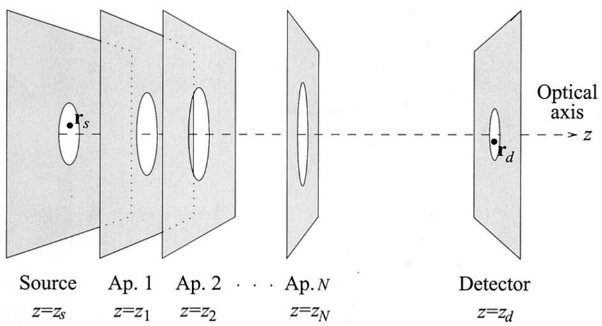

Fig. 1. Canonical optical arrangement considered in this work. Light emitted from an extended source passes by N optical elements before reaching a detector. The optical axis is assumed to be the z axis. Ap., aperture.

Consider the general system shown in Fig. 1. The z axis is the optical axis, with $x = y = 0$. The source, N apertures (or lenses), and the detector are positioned along that axis. Radiation emitted at $\mathbf{r}_s$ can propagate to $\mathbf{r}_d$ by passing by all N intermediate optical elements. If one spectral component originates at $\mathbf{r}_s$ as a spherical wave of the form $u(\mathbf{r}) = u_0 \exp(ik|\mathbf{r} - \mathbf{r}_s|)/|\mathbf{r} - \mathbf{r}_s|$, where $k = 2\pi/\lambda$ is the angular wave number, the approximation used provides a prescription for evaluating the resulting wave field at $\mathbf{r}_d$, $u(k, \mathbf{r}_s, \mathbf{r}_d)$:

$$u(k, \mathbf{r}_s, \mathbf{r}_d) \approx \frac{u_0}{(i\lambda)^N} \int_{\mathrm{Ap1.\,.\,.ApN}} \mathrm{d}^2\mathbf{r}_1 . . . \mathrm{d}^2\mathbf{r}_N \times G(k, \mathbf{r}_s, \mathbf{r}_1) . . . G(k, \mathbf{r}_N, \mathbf{r}_d) \times \exp[ik\delta L(\{\mathbf{r}_\mu\})], \tag{2}$$

where

$$G(k, \mathbf{r}_\mu, \mathbf{r}_\nu) = \frac{\exp(ik|\mathbf{r}_\mu - \mathbf{r}_\nu|)}{|\mathbf{r}_\mu - \mathbf{r}_\nu|} \approx \left(\frac{1}{z_\nu - z_\mu}\right) \exp\left\{ik\left[z_\nu - z_\mu + \frac{(x_\nu - x_\mu)^2 + (y_\nu - y_\mu)^2}{2(z_\nu - z_\mu)}\right]\right\} \tag{3}$$

is the k-dependent free-space propagator or Green's function for the wave field between two points. Here a point's Cartesian coordinates are denoted $\mathbf{r}_\mu = x_\mu\hat{\mathbf{x}} + y_\mu\hat{\mathbf{y}} + z_\mu\hat{\mathbf{z}}$. Also, we always assume $z_\nu > z_\mu$ and that the difference in z coordinates is much larger than the differences in x or y coordinates. The factor

$$\exp[ik\delta L(\{\mathbf{r}_\mu\})] = \exp\left[ik \sum_{\mu=1}^{N} \left(-\frac{x_\mu^2 + y_\mu^2}{2f_\mu}\right)\right] \tag{4}$$

can be used to introduce a finite focal length f_μ that can convert an aperture into a focusing optic. Correspondingly, an aperture effectively has $|f_\mu| = \infty$. Integration of $\mathbf{r}_\mu$ is understood to run over the area of optical element μ. Also, note that the units of $u(k, \mathbf{r}_s, \mathbf{r}_d)$ and u_0 differ, because the units of u_0 have an additional factor of length.

Let us now introduce the abbreviations

$$L(\{\mathbf{r}_\mu\}) = z_d - z_s + \frac{(x_1 - x_s)^2 + (y_1 - y_s)^2}{2(z_1 - z_s)} + . . . + \frac{(x_d - x_N)^2 + (y_d - y_N)^2}{2(z_d - z_N)} + \delta L(\{\mathbf{r}_\mu\}), \tag{5}$$

which (approximately) is a path length from $\mathbf{r}_s$ to $\mathbf{r}_d$ sampled in Eq. (2), and

$$\Delta = (z_1 - z_s)(z_2 - z_1)\ldots(z_d - z_N). \tag{6}$$

We also define

$$f(l, \mathbf{r}_s, \mathbf{r}_d) = \int_{\mathrm{Ap}1} \mathrm{d}^2\mathbf{r}_1 \ldots \int_{\mathrm{Ap}N} \mathrm{d}^2\mathbf{r}_N \delta[l - L(|\mathbf{r}_\mu\})], \tag{7}$$

which describes the frequency of occurrence of a path length equal to l in Eq. (2). This gives

$$u(k, \mathbf{r}_s, \mathbf{r}_d) = \frac{u_0}{\Delta(i\lambda)^N} \int_{-\infty}^{\infty} \mathrm{d}l f(l, \mathbf{r}_s, \mathbf{r}_d)\exp(ikl), \tag{8}$$

which involves the Fourier transform of $f(l, \mathbf{r}_s, \mathbf{r}_d)$ with respect to l. Squaring yields

$$|u(k, \mathbf{r}_s, \mathbf{r}_d)|^2 = \frac{|u_0|^2}{\Delta^2\lambda^{2N}} \int_{-\infty}^{\infty} \mathrm{d}l \int_{-\infty}^{\infty} \mathrm{d}l' f(l, \mathbf{r}_s, \mathbf{r}_d) \times f(l', \mathbf{r}_s, \mathbf{r}_d)\exp[ik(l - l')], \tag{9}$$

which involves the Fourier transform of the autoconvolution of $f(l, \mathbf{r}_s, \mathbf{r}_d)$, all with respect to l. Finally, the ratio $|u(k, \mathbf{r}_s, \mathbf{r}_d)/u_0|^2$ is given by the function

$$\begin{aligned} T(k, \mathbf{r}_s, \mathbf{r}_d) &= |u(k, \mathbf{r}_s, \mathbf{r}_d)/u_0|^2 \\ &= \frac{k^{2N}}{(2\pi)^{2N}\Delta^2} \int_{-\infty}^{\infty} \mathrm{d}l \int_{-\infty}^{\infty} \mathrm{d}l' f(l, \mathbf{r}_s, \mathbf{r}_d) \\ &\quad \times f(l', \mathbf{r}_s, \mathbf{r}_d)\exp[ik(l - l')] \\ &= \frac{k^{2N-2}}{(2\pi)^{2N}\Delta^2} \int_{-\infty}^{\infty} \mathrm{d}l \int_{-\infty}^{\infty} \mathrm{d}l' \frac{\mathrm{d}f(l, \mathbf{r}_s, \mathbf{r}_d)}{\mathrm{d}l} \\ &\quad \times \frac{\mathrm{d}f(l', \mathbf{r}_s, \mathbf{r}_d)}{\mathrm{d}l'} \exp[ik(l - l')]. \end{aligned} \tag{10}$$

In many cases there is a limit,

$$T_0(\mathbf{r}_s, \mathbf{r}_d) = \lim_{k\to\infty} T(k, \mathbf{r}_s, \mathbf{r}_d)|_{\mathrm{Illum}}. \tag{11}$$

If this limit exists, it appears to be the geometrical-optics counterpart of $T(k, \mathbf{r}_s, \mathbf{r}_d)$ in the illuminated region of the $z = z_d$ plane. As an example, consider the case of one optical element between the source and detector. In such a case, we may abbreviate $z_1 - z_s = d_s$ and $z_d - z_1 = d_d$. If the optical element is nonfocusing, we have $T_0(\mathbf{r}_s, \mathbf{r}_d) = (d_d + d_s)^{-2}$. If the optical element is a lens with focal length f, we have

$$T_0(\mathbf{r}_s, \mathbf{r}_d) = \frac{1}{d_s^2 d_d^2}\left(\frac{1}{d_s} + \frac{1}{d_d} - \frac{1}{f}\right)^{-2}. \tag{12}$$

At $|f| = \infty$, we have the previous result, and we have the result $T_0 = d_s^{-2}$ for $d_d \to 0$, independent of f. Otherwise, $T_0(\mathbf{r}_s, \mathbf{r}_d)$ varies in size as the area of the illuminated region varies inversely as a function of d_d. Near the focal plane, $T_0(\mathbf{r}_s, \mathbf{r}_d)$ diverges as the illuminated region shrinks to a point, so that a $k \to \infty$ limit for $T(k, \mathbf{r}_s, \mathbf{r}_d)$ does not exist.

The total and spectral irradiance at $\mathbf{r}_d$ can be expressed as the sum of contributions from each area element of the extended source, $\mathrm{d}A_s$, according to

$$\begin{aligned} E(\mathbf{r}_d, T_s) &= \int_{\mathrm{Source}} \mathrm{d}^2\mathbf{r}_s \frac{\mathrm{d}E(\mathbf{r}_d, T_s)}{\mathrm{d}A_s} \\ &= \int_{\mathrm{Source}} \mathrm{d}^2\mathbf{r}_s \int_0^{\infty} \mathrm{d}\lambda \frac{\mathrm{d}E_\lambda(\lambda, \mathbf{r}_d, T_s)}{\mathrm{d}A_s}. \end{aligned} \tag{13}$$

For an extended-area incoherent source, there is an incremental spectral irradiance $\mathrm{d}E_\lambda(\lambda, \mathbf{r}_d, T_s)$ related to a source area element, $\mathrm{d}A_s$. This equals the irradiance commensurate with a fictitious point source times the effective density of point sources per unit area, ρ. If $|u_0|^2$ is the spectral power emitted per point source per steradian, one has $L_\lambda(\lambda, T_s) = \rho|u_0|^2$. The incremental irradiance is therefore $\mathrm{d}E_\lambda(\lambda, \mathbf{r}_d, T_s) = \rho|u(k, \mathbf{r}_s, \mathbf{r}_d)|^2\mathrm{d}A_s$, giving

$$\begin{aligned} \frac{\mathrm{d}E_\lambda(\lambda, \mathbf{r}_d, T_s)}{\mathrm{d}A_s} &= \rho T(k, \mathbf{r}_s, \mathbf{r}_d)|u_0|^2 \\ &= T(k, \mathbf{r}_s, \mathbf{r}_d)L_\lambda(\lambda, T_s). \end{aligned} \tag{14}$$

Hence $T(k, \mathbf{r}_s, \mathbf{r}_d)$ is the ratio of $\mathrm{d}E_\lambda(\lambda, \mathbf{r}_d, T_s)/\mathrm{d}A_s$ to the source spectral radiance. Therefore we have

$$\begin{aligned} \frac{\mathrm{d}E(\mathbf{r}_d, T_s)}{\mathrm{d}A_s} &= \int_0^{\infty} \mathrm{d}\lambda T(k, \mathbf{r}_s, \mathbf{r}_d)L_\lambda(\lambda, T_s) \\ &= \frac{\epsilon c_1}{\pi}\int_0^{\infty} \frac{\mathrm{d}\lambda}{\lambda^5} \frac{T(k, \mathbf{r}_s, \mathbf{r}_d)}{\exp[c_2/(\lambda T_s)] - 1}, \end{aligned} \tag{15}$$

where c_1 and c_2 are the first and second radiation constants, respectively. The factor ϵ is the source emissivity, which is unity for an ideal blackbody and which we shall assume to be wavelength independent. It is more useful to express this result by using the angular wave number k, yielding

$$\frac{\mathrm{d}E(\mathbf{r}_d, T_s)}{\mathrm{d}A_s} = \frac{\epsilon c_1}{16\pi^5}\int_0^{\infty} \mathrm{d}k k^3 \frac{T(k, \mathbf{r}_s, \mathbf{r}_d)}{\exp(\beta k) - 1}, \tag{16}$$

where I have introduced

$$\beta = \frac{c_2}{2\pi T_s} = \frac{\hbar c}{k_B T_s}. \tag{17}$$

Without diffraction effects, the irradiance expected from geometrical optics is given by

$$\frac{\mathrm{d}E_0(\mathbf{r}_d, T_s)}{\mathrm{d}A_s} = \left[\frac{3\zeta(4)\epsilon c_1}{8\pi^5\beta^4}\right]T_0(\mathbf{r}_s, \mathbf{r}_d). \tag{18}$$

Here $\zeta(z) = \sum_{n=1}^{\infty} n^{-z}$ is the Riemann ζ function. When diffraction effects are taken into account, we have

$$\begin{aligned}\frac{\mathrm{d}E(\mathbf{r}_d, T_s)}{\mathrm{d}A_s} &= \frac{\epsilon c_1}{16\pi^5[(2\pi)^{2N}\Delta^2]} \\ &\quad\times \int_{-\infty}^{\infty} \mathrm{d}l \int_{-\infty}^{\infty} \mathrm{d}l' f(l, \mathbf{r}_s, \mathbf{r}_d) f(l', \mathbf{r}_s, \mathbf{r}_d) \\ &\quad\times \int_0^{\infty} \mathrm{d}k k^{3+2N} \frac{\exp[ik(l-l')]}{\exp(\beta k) - 1} \\ &= \frac{(3+2N)!\epsilon c_1}{16\pi^5[(2\pi)^{2N}\Delta^2]} \\ &\quad\times \int_{-\infty}^{\infty} \mathrm{d}l \int_{-\infty}^{\infty} \mathrm{d}l' f(l, \mathbf{r}_s, \mathbf{r}_d) f(l', \mathbf{r}_s, \mathbf{r}_d) \\ &\quad\times \sum_{n=1}^{\infty} \frac{1}{[n\beta - i(l-l')]^{2N+4}}. \end{aligned} \tag{19}$$

Integration by parts with respect to l and l' also gives

$$\begin{aligned}\frac{\mathrm{d}E(\mathbf{r}_d, T_s)}{\mathrm{d}A_s} &= \frac{\epsilon c_1}{16\pi^5[(2\pi)^{2N}\Delta^2]} \\ &\quad\times \int_{-\infty}^{\infty} \mathrm{d}l \int_{-\infty}^{\infty} \mathrm{d}l' \frac{\mathrm{d}f(l, \mathbf{r}_s, \mathbf{r}_d)}{\mathrm{d}l} \\ &\quad\times \frac{\mathrm{d}f(l', \mathbf{r}_s, \mathbf{r}_d)}{\mathrm{d}l'} \int_0^{\infty} \mathrm{d}k k^{1+2N} \\ &\quad\times \frac{\exp[ik(l-l')]}{\exp(\beta k) - 1} \\ &= \frac{(1+2N)!\epsilon c_1}{16\pi^5[(2\pi)^{2N}\Delta^2]} \\ &\quad\times \int_{-\infty}^{\infty} \mathrm{d}l \int_{-\infty}^{\infty} \mathrm{d}l' \frac{\mathrm{d}f(l, \mathbf{r}_s, \mathbf{r}_d)}{\mathrm{d}l} \\ &\quad\times \frac{\mathrm{d}f(l', \mathbf{r}_s, \mathbf{r}_d)}{\mathrm{d}l'} \\ &\quad\times \sum_{n=1}^{\infty} \frac{1}{[n\beta - i(l-l')]^{2N+2}}. \end{aligned} \tag{20}$$

These integrations are all symmetric under exchange of l and l', so only the real parts of integrals survive. Also, results are unaffected if the function $f(l, \mathbf{r}_s, \mathbf{r}_d)$ is translated with respect to l.

The last result can be especially useful, because $f(l, \mathbf{r}_s, \mathbf{r}_d)$ can have the property

$$\frac{\mathrm{d}}{\mathrm{d}l}\left[\frac{f(l, \mathbf{r}_s, \mathbf{r}_d)}{(l-l_0)^{N-1}}\right] = g\delta(l - l_0) + b(l, \mathbf{r}_s, \mathbf{r}_d), \tag{21}$$

where $b(l, \mathbf{r}_s, \mathbf{r}_d)$ is zero except for certain ranges of l. In the case of $N = 1$, an assumption to be made for the remainder of this section, some rearrangement gives

$$\begin{aligned}\frac{\mathrm{d}E(\mathbf{r}_d, T_s)}{\mathrm{d}A_s} &= \frac{3\epsilon c_1}{64\pi^7 d_s^2 d_d^2} \int_{-\infty}^{\infty} \mathrm{d}l S(\beta, l) \\ &\quad\times \int_{-\infty}^{\infty} \mathrm{d}s \left[\frac{\mathrm{d}f(s, \mathbf{r}_s, \mathbf{r}_d)}{\mathrm{d}s}\right] \\ &\quad\times \left[\frac{\mathrm{d}f(s+l, \mathbf{r}_s, \mathbf{r}_d)}{\mathrm{d}s}\right], \end{aligned} \tag{22}$$

where the function $S(\beta, l)$ is easy to evaluate and is discussed in Appendix A.

This manner of expressing the last result extracts all spectral aspects (for a thermal source, appearing here by means of the parameter β), from the innermost integral, which becomes a purely geometric entity. Such a separation of spectral and spatial aspects of the flow of radiation is highly advantageous, because the spatial aspects do not depend on wavelength and hence can be treated simultaneously for all spectral components. This is the main result of this work. If the integration over s is performed once for all l, the total irradiance may be deduced by at most a single integration over l. Because the integrands are not highly oscillatory, numerical integration can be practical as an alternative to analytic integration.

Typically, the parameter g has one value in the illuminated region and is zero otherwise in the case of Fresnel diffraction, and its contributions to $\mathrm{d}f(l, \mathbf{r}_s, \mathbf{r}_d)/\mathrm{d}l$ are related to the geometrical wave in the boundary-diffraction-wave formulation of Kirchhoff diffraction theory.[5,6] In that case the parameter l_0 is an extremal value of l. Contributions to $\mathrm{d}E(\mathbf{r}_d, T_s)/\mathrm{d}A_s$ arising from the product of two geometrical-wave terms are the same as what one would have according to geometrical optics, and typically having a nonzero g is synonymous with the existence of the limit, $T_0(\mathbf{r}_s, \mathbf{r}_d)$. This is clear from Eq. (10), at least for the case of $N = 1$, where the factor outside of the integral has a $k \to \infty$ limit and contributions from $b(l, \mathbf{r}_s, \mathbf{r}_d)$ and/or $b(l', \mathbf{r}_s, \mathbf{r}_d)$ in the integrand do not contribute upon integration in the $k \to \infty$ limit if such functions are well behaved and, in particular, do not have δ-function-like contributions. For other, pathological cases, closer examination is necessary to interpret the present assertions.

Contributions to $\mathrm{d}f(l, \mathbf{r}_s, \mathbf{r}_d)/\mathrm{d}l$ from $b(l, \mathbf{r}_s, \mathbf{r}_d)$ are related to the boundary-diffraction wave. Their contributions to $\mathrm{d}E(\mathbf{r}_d, T_s)/\mathrm{d}A_s$ are synonymous with diffraction effects, and this permits a clear identification of diffraction effects even at the start of a calculation. In particular, in the case of $N = 1$ we have

$$\frac{dE(\mathbf{r}_d, T_s)}{dA_s} = \frac{3\epsilon c_1 g^2 S(\beta, 0)}{64\pi^7 d_s^2 d_d^2} + \frac{3\epsilon c_1 g}{32\pi^7 d_s^2 d_d^2} \times \int_{-\infty}^{\infty} dl S(\beta, l) b(l_0 + l, \mathbf{r}_s, \mathbf{r}_d) + \frac{3\epsilon c_1}{64\pi^7 d_s^2 d_d^2} \int_{-\infty}^{\infty} dl S(\beta, l) \times \int_{-\infty}^{\infty} ds b(s, \mathbf{r}_s, \mathbf{r}_d) b(s + l, \mathbf{r}_s, \mathbf{r}_d) = \frac{dE_0}{dA_s} [I_G + I_X(\beta, \mathbf{r}_s, \mathbf{r}_d) + I_B(\beta, \mathbf{r}_s, \mathbf{r}_d)]. \tag{23}$$

The first term is the one that would arise in geometrical optics, while the two remaining terms are direct consequences of diffraction and are expressed in terms of the functions $I_X(\beta, \mathbf{r}_s, \mathbf{r}_d)$ and $I_B(\beta, \mathbf{r}_s, \mathbf{r}_d)$ that have been introduced. Because I_X and I_B are always functions of the same arguments, their arguments are suppressed in much of what follows. Also, even symmetry with respect to l of the inner integrand contributing to I_B permits one to integrate over $l > 0$ only and double the result. Furthermore, in the integral contributing to I_X, $b(l_0 + l, \mathbf{r}_s, \mathbf{r}_d)$ may be identically zero for all $l < 0$ or all $l > 0$. I_G is one in the illuminated region and zero otherwise, and henceforth dE_0/dA_s is assumed to denote the geometrical-optics value in the illuminated region. Although one might anticipate that analogous results such as those just discussed can also be true for $N > 1$, such an assertion remains to be investigated.

3. Application

Let us now seek to apply the above formulation to assess diffraction effects on irradiance at the detector in the $N = 1$ case. Without additional loss of generality, in what follows the area element of the source that illuminates a diffracting aperture is assumed to be centered on the optical axis.

A. Fraunhofer Diffraction

In the case of Fraunhofer diffraction, it is convenient to express $f(l, \mathbf{r}_s, \mathbf{r}_d)$ and all related quantities in terms of the direction cosines from the center of the diffracting aperture to $\mathbf{r}_d$: $\theta_x = x_d/d_d$ and $\theta_y = y_d/d_d$. The path length traveled by light from the source through the aperture to the detector plane is given by

$$L(\mathbf{r}_s, \mathbf{r}_1, \mathbf{r}_d) \approx d_s + d_d - \theta_x x_1 - \theta_y y_1 \approx l_0 - \theta_x x_1 - \theta_y y_1. \tag{24}$$

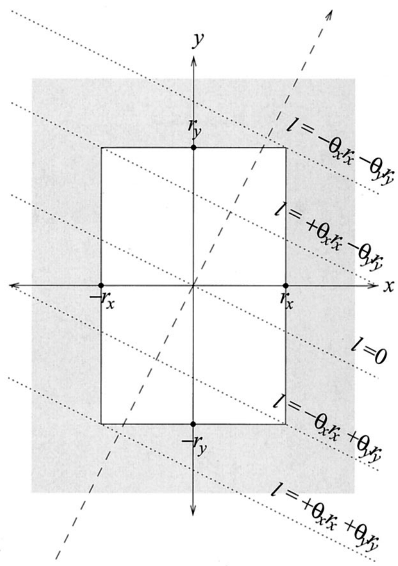

Fig. 2. Rectangular aperture with several geometrical parameters indicated.

Let us henceforth reset the constant l_0 to zero, as we are allowed to do. Using the identity

$$\int_{-\infty}^{\infty} dl f(l, \mathbf{r}_s, \mathbf{r}_d) = A_{\mathrm{Ap1}}, \tag{25}$$

where A_{Ap1} is the area of the diffracting aperture, inspection of Eq. (19) already gives this general result for irradiance on the optical axis:

$$\frac{dE(\mathbf{r}_d, T_s)}{dA_s} = \Xi_0 = \frac{15\zeta(6)\epsilon c_1 A_{\mathrm{Ap1}}^2}{8\pi^7 d_s^2 d_d^2 \beta^6}. \tag{26}$$

B. Fraunhofer Diffraction by a Rectangular Aperture

Figure 2 depicts a rectangular aperture with width $2r_x$ along the x direction and height $2r_y$ along the y direction. The dashed line indicates the direction corresponding to a possible pair of values of direction cosines (θ_x, θ_y). Dotted lines are constant-l contours in the $z = z_1$ (aperture) plane. For convenience, l is defined to be zero at the center of the aperture. Other values of l are indicated on each contour. It is helpful to introduce two parameters, $l_\pm = |(|\theta_x r_x| \pm |\theta_y r_y|)|$. From inspection, we see that $f(l, \mathbf{r}_s, \mathbf{r}_d)$ is zero for $l < -l_+$, rises steadily for $-l_+ < l < -l_-$, is constant for $-l_- < l < l_-$, decreases steadily in the same manner for $l_- < l < l_+$, and is zero for $l > l_+$. Using the sum rule of Eq. (25), we may deduce $f(l, \mathbf{r}_s, \mathbf{r}_d) = A_{\mathrm{Ap1}}/(l_- + l_+)$ for $-l_- < l < l_-$, giving $df(l, \mathbf{r}_s, \mathbf{r}_d)/dl = \pm A_{\mathrm{Ap1}}/(l_+^2 - l_-^2) \equiv \pm B$ for $-l_+ < l < -l_-$ and $l_- < l < l_+$, respectively.

Using Eq. (20), we have

$$\frac{\mathrm{d}E(\mathbf{r}_d, T_s)}{\mathrm{d}A_s} = \frac{3\epsilon c_1 B^2}{32\pi^7 d_s^2 d_d^2}\left(\int_{-l_+}^{-l_-}\mathrm{d}l\int_{-l_+}^{-l_-}\mathrm{d}l' - \int_{-l_+}^{-l_-}\mathrm{d}l\int_{+l_-}^{+l_+}\mathrm{d}l' - \int_{+l_-}^{+l_+}\mathrm{d}l\int_{-l_+}^{-l_-}\mathrm{d}l' + \int_{+l_-}^{+l_+}\mathrm{d}l\int_{+l_-}^{+l_+}\mathrm{d}l'\right)$$
$$\times\sum_{n=1}^{\infty}\frac{1}{[n\beta - i(l-l')]^4}$$
$$= \frac{\epsilon c_1 B^2}{64\pi^7 d_s^2 d_d^2}\left(\left\{\sum_{n=1}^{\infty}\frac{1}{[n\beta - i(l-l')]^2}\Bigg|_{l=-l_+}^{l=-l_-}\right\}\Bigg|_{l'=-l_+}^{l'=-l_-} - \ldots\right). \tag{27}$$

The term shown corresponds to the first limits of integration shown, and terms not shown follow accordingly. Gathering all terms and simplifying, including using

$$\frac{1}{(a-ib)^2} + \frac{1}{(a+ib)^2} = \frac{(a+ib)^2 + (a-ib)^2}{(a^2+b^2)^2} = \frac{2(a^2-b^2)}{(a^2+b^2)^2} = \frac{2}{a^2+b^2} - \frac{4b^2}{(a^2+b^2)^2}, \tag{28}$$

we arrive at

$$\frac{\mathrm{d}E(\mathbf{r}_d, T_s)}{\mathrm{d}A_s} = \frac{\epsilon c_1 B^2}{16\pi^7 d_s^2 d_d^2 \beta^2}[\zeta(2) - 4\pi^2 \times g(2\pi(l_+ - l_-)/\beta) - 4\pi^2 \times g(2\pi(l_+ + l_-)/\beta) + 2\pi^2 g(4\pi l_+/\beta) + 2\pi^2 g(4\pi l_-/\beta)]. \tag{29}$$

Here we have introduced $g(z) = S_1(z) - 2z^2 S_2(z)$; $S_1(z)$ and $S_2(z)$ and a means to evaluate these functions are described in Appendix A. A contour plot of this result as a function of $(\theta_x^2 r_x^2 + \theta_y^2 r_y^2)^{1/2}/\beta$ is shown in Fig. 3 for $\beta = 0.01$ mm, $r_x = 1$ mm, $r_y = 2$ mm. It is normalized to give a result of unity for $\theta_x = \theta_y = 0$. Curves are shown for $(\theta_y r_y)/(\theta_x r_x) = \tan\phi$, with $\tan\phi = 0$, $\tan\phi = 0.5$, and $\tan\phi = 1$.

If we have $\theta_y = 0$, then we have $l_+ = l_- = \bar{l}$, and the above results simplify to

$$\frac{\mathrm{d}f(l, \mathbf{r}_s, \mathbf{r}_d)}{\mathrm{d}l} = \left(\frac{A_{\mathrm{Ap1}}}{2\bar{l}}\right)\delta(l+\bar{l}) - \left(\frac{A_{\mathrm{Ap1}}}{2\bar{l}}\right)\delta(l-\bar{l}), \tag{30}$$

$$\frac{\mathrm{d}E(\mathbf{r}_d, T_s)}{\mathrm{d}A_s} = \frac{3\epsilon c_1 A_{\mathrm{Ap1}}^2}{128\pi^7 d_s^2 d_d^2 \bar{l}^2}\sum_{n=1}^{\infty}\left[\frac{2}{(n\beta)^4} - \frac{1}{(n\beta - 2i\bar{l})^4} - \frac{1}{(n\beta + 2i\bar{l})^4}\right] = \frac{3\epsilon c_1 A_{\mathrm{Ap1}}^2}{128\pi^7 d_s^2 d_d^2 \bar{l}^2 \beta^4}[2\zeta(4) - S(1, 2\bar{l}/\beta)]. \tag{31}$$

In evaluating such results, care must be taken for $l_+ \approx l_-$ in Eq. (29) and small $\bar{l}$ in Eq. (31), because the finiteness of $\mathrm{d}E(\mathbf{r}_d, T_s)/\mathrm{d}A_s$ can depend on mutual cancellation of two or more divergent terms.

C. Fraunhofer Diffraction by a Long, Narrow Slit

Diffraction by a slit is closely related to diffraction by a rectangle. When the slit is very long, the Fraunhofer approximation is only appropriate along the narrow direction, whereas the paraxial Fresnel approximation may be used along the long direction. One therefore has, for a slit parallel to the y axis,

$$L(\mathbf{r}_s, \mathbf{r}_1, \mathbf{r}_d) \approx l_0 - \theta_x x_1 + \left(\frac{1}{2d_s} + \frac{1}{2d_d}\right)y_1^2, \tag{32}$$

and we again immediately reset l_0 to zero. Here, terms linear in y_1 may be dropped when one takes the limit of an infinitely long slit, as we shall do. In place of Eq. (8), we shall use

$$u(k, \mathbf{r}_s, \mathbf{r}_d) = \frac{u_0}{i\lambda d_s d_d}\int_{-\infty}^{\infty}\mathrm{d}l f_1(l, \mathbf{r}_s, \mathbf{r}_d)\exp(ikl) \times \int_{-\infty}^{\infty}\mathrm{d}y_1 \exp\left[ik\left(\frac{1}{2d_s} + \frac{1}{2d_d}\right)y_1^2\right], \tag{33}$$

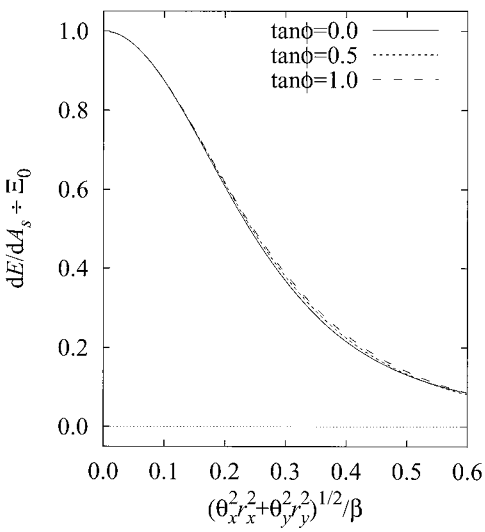

Fig. 3. $(\mathrm{d}E/\mathrm{d}A_s)/\Xi_0$ for case of Fraunhofer diffraction by a rectangular aperture along three lines in the (θ_x, θ_y) plane as discussed in the text.

where we have

$$f_1(l, \mathbf{r}_s, \mathbf{r}_d) = \int_{-r_x}^{+r_x} \mathrm{d}x_1 \delta[l - (-\theta_x x_1)]$$
$$= \Theta(|\theta_x r_x| - l)\Theta(l + |\theta_x r_x|)/|\theta_x|. \quad (34)$$

This leads to

$$T(k, \mathbf{r}_s, \mathbf{r}_d) = \frac{k}{2\pi d_s d_d(d_s + d_d)} \times \int_{-\infty}^{\infty} \mathrm{d}l \int_{-\infty}^{\infty} \mathrm{d}l' f_1(l, \mathbf{r}_s, \mathbf{r}_d) f_1(l', \mathbf{r}_s, \mathbf{r}_d) \times \exp[ik(l - l')]. \quad (35)$$

Using Eq. (16), we arrive at

$$\frac{\mathrm{d}E(\mathbf{r}_d, T_s)}{\mathrm{d}A_s} = \frac{\epsilon c_1}{32\pi^6 d_s d_d(d_s + d_d)} \times \int_{-\infty}^{\infty} \mathrm{d}l \int_{-\infty}^{\infty} \mathrm{d}l' f_1(l, \mathbf{r}_s, \mathbf{r}_d) f_1(l', \mathbf{r}_s, \mathbf{r}_d) \times \int_0^{\infty} \frac{\mathrm{d}k k^4 \exp[ik(l - l')]}{\exp(\beta k) - 1}$$
$$= \frac{3\epsilon c_1}{4\pi^6 d_s d_d(d_s + d_d)} \times \int_{-\infty}^{\infty} \mathrm{d}l \int_{-\infty}^{\infty} \mathrm{d}l' f_1(l, \mathbf{r}_s, \mathbf{r}_d) f_1(l', \mathbf{r}_s, \mathbf{r}_d) \times \sum_{n=1}^{\infty} [n\beta - i(l - l')]^{-5}$$
$$= \frac{\epsilon c_1}{16\pi^6 d_s d_d(d_s + d_d)} \times \int_{-\infty}^{\infty} \mathrm{d}l \int_{-\infty}^{\infty} \mathrm{d}l' \frac{\mathrm{d}f_1(l, \mathbf{r}_s, \mathbf{r}_d)}{\mathrm{d}l} \times \frac{\mathrm{d}f_1(l', \mathbf{r}_s, \mathbf{r}_d)}{\mathrm{d}l'} \sum_{n=1}^{\infty} [n\beta - i(l - l')]^{-3}. \quad (36)$$

The derivatives are given by

$$\frac{\mathrm{d}f_1(l, \mathbf{r}_s, \mathbf{r}_d)}{\mathrm{d}l} = [\delta(l + |\theta_x r_x|) - \delta(l - |\theta_x r_x|)]/|\theta_x|. \quad (37)$$

Continuing analysis in the same way that was used for treating Fraunhofer diffraction by a rectangular aperture, we obtain

$$\frac{\mathrm{d}E(\mathbf{r}_d, T_s)}{\mathrm{d}A_s} = \frac{\epsilon c_1 h(z)}{16\pi^6 d_s d_d(d_s + d_d)\theta_x^2 \beta^3}, \quad (38)$$

with

$$h(z) = 2\zeta(3) - \sum_{n=1}^{\infty} \frac{1}{(n - iz)^3} - \sum_{n=1}^{\infty} \frac{1}{(n + iz)^3} \quad (39)$$

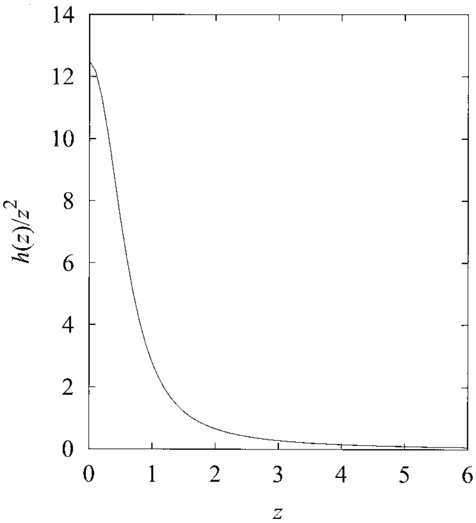

Fig. 4. Universal function $h(z)$ relevant for the case of Fraunhofer diffraction by a rectangular slit, plotted as $h(z)/z^2$ versus z.

and $z = 2\theta_x r_x/\beta$. The function $h(z)$, universal for all infinitely long, narrow slits, is plotted divided by z^2 in Fig. 4, and $h(z)$ is tabulated for several values in Table 1. Note that we have $h(0) = 0$, $h(z) \approx 12\zeta(5)z^2$ for small z, and $h(\infty) = 2\zeta(3)$.

D. Fraunhofer Diffraction by a Circular Aperture

For a circular aperture of radius R, the irradiance in the detector plane depends only on $\theta = (\theta_x^2 + \theta_y^2)^{1/2}$. Noting that we have

$$L(\mathbf{r}_d, \mathbf{r}_1, \mathbf{r}_s) = l_0 - x_1\theta_x - y_1\theta_y, \quad (40)$$

and $x_1^2 + y_1^2 < R^2$ on the aperture, we reset l_0 to zero and can easily obtain

$$f(l, \mathbf{r}_s, \mathbf{r}_d) = \frac{2}{\theta^2}(R^2\theta^2 - l^2)^{1/2} \quad (41)$$

for $|l| < R\theta$, as well as $f(l, \mathbf{r}_s, \mathbf{r}_d) = 0$ otherwise. This gives, when appropriate,

$$b(l, \mathbf{r}_s, \mathbf{r}_d) = -\frac{2l}{\theta^2(R^2\theta^2 - l^2)^{1/2}}. \quad (42)$$

Table 1. Function $h(z)$ Versus z

z	$h(z)$	z	$h(z)$	z	$h(z)$
0.0	0.0000	1.2	2.7871	5.5	2.4366
0.1	0.1215	1.4	2.7509	6.0	2.4315
0.2	0.4527	1.6	2.7029	6.5	2.4275
0.3	0.9105	1.8	2.6577	7.0	2.4243
0.4	1.3986	2.0	2.6191	7.5	2.4217
0.5	1.8407	2.5	2.5505	8.0	2.4196
0.6	2.1953	3.0	2.5088	10.0	2.4141
0.7	2.4522	3.5	2.4823	15.0	2.4085
0.8	2.6216	4.0	2.4646	20.0	2.4066
0.9	2.7219	4.5	2.4523	100.0	2.4042
1.0	2.7727	5.0	2.4433	∞	2.4041

Application of Eq. (23) by use of $b(s - l/2, \mathbf{r}_s, \mathbf{r}_d)$ and $b(s + l/2, \mathbf{r}_s, \mathbf{r}_d)$ gives

$$\begin{aligned}\frac{\mathrm{d}E}{\mathrm{d}A_s} &= \frac{3\epsilon c_1}{8\pi^7 d_s^2 d_d^2 \theta^4}\int_0^{2R\theta} \mathrm{d}l S(\beta, l)\int_{-R\theta+l/2}^{R\theta-l/2} \frac{\mathrm{d}s(s-l/2)(s+l/2)}{[(R\theta - s - l/2)(R\theta + s + l/2)(R\theta - s + l/2)(R\theta + s - l/2)]^{1/2}} \\ &= \frac{3\epsilon c_1}{8\pi^7 d_s^2 d_d^2 \theta^4}\int_0^{2R\theta} \frac{\mathrm{d}l S(\beta, l)}{R\theta - l/2}\int_{-1}^{1} \frac{\mathrm{d}x[(R\theta - l/2)^2 x^2 - l^2/4]}{((1-x^2)\{[(R\theta + l/2)/(R\theta - l/2)]^2 - x^2\})^{1/2}} \\ &= \frac{3\epsilon c_1}{8\pi^7 d_s^2 d_d^2 \theta^4}\int_0^{2R\theta} \frac{\mathrm{d}l S(\beta, l)}{R\theta + l/2}\int_{-1}^{1} \frac{\mathrm{d}x[(R\theta - l/2)^2 x^2 - l^2/4]}{((1-x^2)\{1 - [(R\theta - l/2)/(R\theta + l/2)]^2 x^2\})^{1/2}} \\ &= \frac{3\epsilon c_1}{8\pi^7 d_s^2 d_d^2 \theta^4}\int_0^{2R\theta} \frac{\mathrm{d}l S(\beta, l)}{2R\theta + l}\int_{0}^{1} \frac{\mathrm{d}x[(2R\theta - l)^2 x^2 - l^2]}{((1-x^2)\{1 - [(2R\theta - l)/(2R\theta + l)]^2 x^2\})^{1/2}} \\ &= \frac{3\epsilon c_1 R^2}{2\pi^7 d_s^2 d_d^2 \theta^2}\int_0^{1} \frac{\mathrm{d}y S(\beta, 2R\theta y)}{1+y}\int_{0}^{1} \frac{\mathrm{d}x[(1-y)^2 x^2 - y^2]}{((1-x^2)\{1 - [(1-y)/(1+y)]^2 x^2\})^{1/2}}. \qquad (43)\end{aligned}$$

Successive steps feature various substitutions and rearrangements, with key changes of variables being $x = s/(R\theta - l/2)$ and $y = l/(2R\theta)$. Using

$$\frac{(1-y)^2 x^2 - y^2}{1+y} = \left(\frac{1+2y}{1+y}\right) - (1+y) \times \left[1 - \left(\frac{1-y}{1+y}\right)^2 x^2\right], \qquad (44)$$

we obtain

$$\frac{\mathrm{d}E}{\mathrm{d}A_s} = \frac{3\epsilon c_1 R^2}{2\pi^7 d_s^2 d_d^2 \theta^2}\int_0^1 \mathrm{d}y S(\beta, 2R\theta y) \times \left[\left(\frac{1+2y}{1+y}\right) K\left(\left(\frac{1-y}{1+y}\right)^2\right) - (1+y) E\left(\left(\frac{1-y}{1+y}\right)^2\right)\right], \qquad (45)$$

where complete elliptic integrals of the first and second kind are defined by

$$K(u) = \int_0^1 \frac{\mathrm{d}t}{[(1-t^2)(1-ut^2)]^{1/2}}, \qquad (46)$$

$$E(u) = \int_0^1 \mathrm{d}t \left(\frac{1-ut^2}{1-t^2}\right)^{1/2}, \qquad (47)$$

with the indicated convention regarding the meaning of their argument. Introducing $\alpha = 2R\theta/\beta$, and from transformation formulas for elliptic integrals, one finds

$$\begin{aligned}\frac{\mathrm{d}E}{\mathrm{d}A_s} &= \frac{3\epsilon c_1 R^2}{4\pi^7 d_s^2 d_d^2 \theta^2 \beta^4}\int_0^1 \mathrm{d}y S(1, \alpha y)[K(1-y^2) - 2E(1-y^2)] \\ &= \frac{3\epsilon c_1 A_{\mathrm{Ap1}}^2}{\pi^9 d_s^2 d_d^2 \alpha^2 \beta^6}\int_0^1 \mathrm{d}y S(1, \alpha y)[K(1-y^2) - 2E(1-y^2)]. \qquad (48)\end{aligned}$$

Because one may derive

$$A_m = \int_0^1 \mathrm{d}y y^{2m} K(1-y^2) = \frac{\pi}{4}\left[\frac{\Gamma(m+1/2)}{\Gamma(m+1)}\right]^2, \qquad (49)$$

$$B_m = \int_0^1 \mathrm{d}y y^{2m} E(1-y^2) = \frac{\pi}{4}\left[\frac{\Gamma(m+1/2)\Gamma(m+3/2)}{\Gamma(m+1)\Gamma(m+2)}\right], \qquad (50)$$

for small α we have

$$\frac{\mathrm{d}E}{\mathrm{d}A_s} = \frac{3\epsilon c_1 A_{\mathrm{Ap1}}^2}{\pi^9 d_s^2 d_d^2 \alpha^2 \beta^6}\sum_{m=0}^{\infty} [(2\pi)^{2m} s_m (A_m - 2B_m)]\alpha^{2m}, \qquad (51)$$

which is valid for $\alpha < 1$. The $\{s_m\}$ coefficients are found in Appendix A. The factor $(2\pi)^{2m}$ arises because the argument z in Appendix A is equal to $2\pi\alpha y$ in this case. For somewhat larger α, direct numerical integration of Eq. (48) may be done. For very large α, the asymptotic behavior of $\eta(\alpha) = (\mathrm{d}E/\mathrm{d}A_s)/$

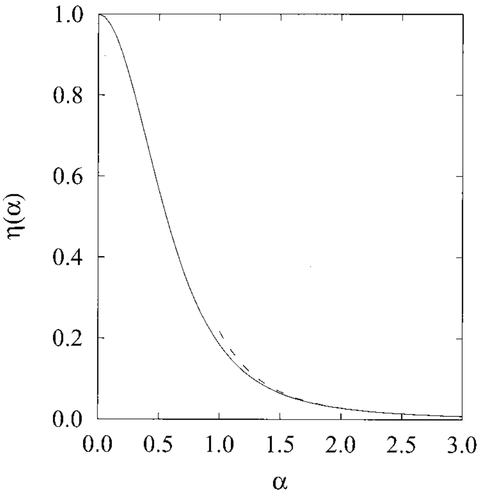

Fig. 5. Universal behavior of $\eta(\alpha)$ for the case of Fraunhofer diffraction by a circular aperture, as a function of angular parameter α (solid curve), and asymptotic result (dashed curve).

Ξ_0, determined by using properties of hypergeometric functions, is helpful:

$$\eta(\alpha) \sim \frac{504\zeta(3)}{\pi^7\alpha^3} + \frac{378 \log_e\alpha}{\pi^7\alpha^5} + \frac{63(6\gamma + 12 \log_e 2 - 11)}{\pi^7\alpha^5} + \cdots. \quad (52)$$

The function $\eta(\alpha)$ is plotted in Fig. 5 and tabulated in Table 2. Note that we have $\eta(0) = 1$.

E. Fresnel Diffraction

In the case of Fresnel diffraction, it is expedient to relate dE/dA_s to the value of dE_0/dA_s in the illuminated region, as is done in Eq. (23). Let us continue to assess the contribution to irradiance by an area element centered on the optical axis. This area element is assumed to be at $\mathbf{r}_s = (0, 0, -d_s)$, and we shall consider the irradiance at $\mathbf{r}_d = (x_d, y_d, d_d)$. A line segment between $\mathbf{r}_s$ and $\mathbf{r}_d$ intersects the $z = 0$ plane at $(x_i, y_i, 0)$, with $x_i = d_s x_d/(d_s + d_d)$ and $y_i = d_s y_d/(d_s + d_d)$. A total path length is reckoned according to

$$\begin{aligned} L(\mathbf{r}_s, \mathbf{r}_1, \mathbf{r}_d) &\approx d_s + d_d + \frac{x_d^2 + y_d^2}{2(d_s + d_d)} \\ &\quad + [(x_1 - x_i)^2 + (y_1 - y_i)^2]\left(\frac{1}{2d_s} + \frac{1}{2d_d}\right) \\ &= l_0 + C[(x_1 - x_i)^2 + (y_1 - y_i)^2], \end{aligned} \quad (53)$$

where a parameter C has been introduced.

F. Fresnel Diffraction by a Circular Aperture

A circular aperture of radius R is assumed to be centered around the z axis. In this case x_1 and y_1 should sample the aperture. By symmetry, we may introduce $\tau = (x_i^2 + y_i^2)^{1/2}$, and deal solely with τ.

For $\mathbf{r}_d$ in the illuminated region, where we have $\tau < R$, l_0 may denote the smallest value taken by $L(\mathbf{r}_s, \mathbf{r}_1, \mathbf{r}_d)$, and $L(\mathbf{r}_s, \mathbf{r}_1, \mathbf{r}_d)$ has increasingly larger values for $\mathbf{r}_1$ on increasingly larger circular contours centered around $\mathbf{r}_i$. Freely setting l_0 to zero, we have

$$L(\mathbf{r}_s, \mathbf{r}_1, \mathbf{r}_d) = C[(x_1 - x_i)^2 + (y_1 - y_i)^2] = C\rho^2, \quad (54)$$

where we have introduced a new variable, ρ. An entire circular contour is sampled by $\mathbf{r}_1$ for $\rho < R - \tau$. For $R - \tau < \rho < R + \tau$, the sampled fraction of the contour is θ/π, where we have

$$\theta = \cos^{-1}\left(\frac{\rho^2 + \tau^2 - R^2}{2\tau\rho}\right). \quad (55)$$

Hence, for $l < C(R - \tau)^2$, we have

$$f(l, \mathbf{r}_s, \mathbf{r}_d) = \pi/C, \quad (56)$$

whereas for $C(R - \tau)^2 < l < C(R + \tau)^2$ we have

$$f(l, \mathbf{r}_s, \mathbf{r}_d) = \theta/C. \quad (57)$$

As l increases over this range, θ decreases from π to zero. For $l > C(R + \tau)^2$, we have

$$f(l, \mathbf{r}_s, \mathbf{r}_d) = 0. \quad (58)$$

Table 2. Function $\eta(\alpha) = (dE/dA_s)/\Xi_0$ Versus α

α	$\eta(\alpha)$	α	$\eta(\alpha)$	α	$\eta(\alpha)$	α	$\eta(\alpha)$	α	$\eta(\alpha)$
0.0	1.000000	2.0	0.027678	4.0	0.003313	6.0	0.000959	8.0	0.000400
0.2	0.903648	2.2	0.020738	4.2	0.002853	6.2	0.000868	8.2	0.000371
0.4	0.683744	2.4	0.015904	4.4	0.002474	6.4	0.000788	8.4	0.000345
0.6	0.459616	2.6	0.012446	4.6	0.002159	6.6	0.000718	8.6	0.000321
0.8	0.292239	2.8	0.009913	4.8	0.001895	6.8	0.000655	8.8	0.000300
1.0	0.184166	3.0	0.008019	5.0	0.001673	7.0	0.000600	9.0	0.000280
1.2	0.118206	3.2	0.006576	5.2	0.001484	7.2	0.000551	9.2	0.000262
1.4	0.078239	3.4	0.005458	5.4	0.001323	7.4	0.000507	9.4	0.000246
1.6	0.053589	3.6	0.004579	5.6	0.001184	7.6	0.000468	9.6	0.000230
1.8	0.037937	3.8	0.003878	5.8	0.001064	7.8	0.000432	9.8	0.000216

For $\mathbf{r}_d$ in the shadow region, where we have $\tau > R$, we have

$$f(l, \mathbf{r}_s, \mathbf{r}_d) = 0 \tag{59}$$

for $l < C(\tau - R)^2$ or $l > C(\tau + R)^2$, whereas we have

$$f(l, \mathbf{r}_s, \mathbf{r}_d) = \theta/C \tag{60}$$

for $C(\tau - R)^2 < l < C(R + \tau)^2$. Here θ is given by the same expression as in the illuminated region. As l increases over this range, θ starts at zero, increases to some maximum value, and then decreases to zero.

When relevant, we have $g = \pi/C$. Also, whatever the case, for $C|\tau - R|^2 < l < C(\tau + R)^2$, we may use the chain rule to obtain

$$b(l) = \frac{1}{C}\frac{\mathrm{d}\theta}{\mathrm{d}l} = \frac{1}{C}\frac{\mathrm{d}\theta}{\mathrm{d}\rho}\frac{\mathrm{d}\rho}{\mathrm{d}l} = \frac{1}{2\rho C^2}\frac{\mathrm{d}\theta}{\mathrm{d}\rho}. \tag{61}$$

Differentiating θ with respect to ρ and simplifying the result leads to

$$b(l) = \frac{A - Bl}{l[(l - l_1)(l_2 - l)]^{1/2}}, \tag{62}$$

where we have

$$A = \frac{\tau^2 - R^2}{2}, \qquad B = \frac{1}{2C},$$
$$l_1 = C|\tau - R|^2, \qquad l_2 = C(\tau + R)^2. \tag{63}$$

As expected, we have $I_G = 1$ in the illuminated region and $I_G = 0$ otherwise. Next, we have

$$I_X = \frac{\beta^4 C}{\pi\zeta(4)}\int_{l_1}^{l_2}\frac{\mathrm{d}l S(\beta, l)(A - Bl)}{l[(l - l_1)(l_2 - l)]^{1/2}}. \tag{64}$$

The integration may be carried out numerically with relative ease, particularly with a change of the variable of integration to θ, where we have $l = l_1 + (l_2 - l_1)(1 - \cos\theta)/2$, and use of Gauss–Legendre quadrature. Finally, we have

$$I_B = \frac{\beta^4 C^2}{\pi^2\zeta(4)}\int_0^{l_2 - l_1}\mathrm{d}l S(\beta, l)\int_{l_1}^{l_2 - l}\mathrm{d}s\,\frac{(A - Bs)(A - Bs - Bl)}{s(s + l)[(s - l_1)(s + l - l_1)(l_2 - s)(l_2 - s - l)]^{1/2}}. \tag{65}$$

We may introduce $a' = (l_2 - l_1 - l)/2$ and $b' = (l_2 + l_1 - l)/2$, so that if we have $s = a'x + b'$, we have $x = -1$ at $s = l_1$ and $x = +1$ at $s = l_2 - l$. This leads to

$$\begin{aligned} I_B &= \frac{\beta^4 C^2}{\pi^2\zeta(4)a'}\int_0^{l_2 - l_1}\mathrm{d}l S(\beta, l) \times \int_{-1}^{+1}\mathrm{d}x\,\frac{(A/a' - Bx - Bb'/a')(A/a' - Bx - Bb'/a' - Bl/a')}{(x + b'/a')(x + b'/a' + l/a')\{(1 - x^2)[(1 + l/a')^2 - x^2]\}^{1/2}} \\ &= \frac{\beta^4 B^2 C^2\kappa}{\pi^2\zeta(4)a'}\int_0^{l_2 - l_1}\mathrm{d}l S(\beta, l)\int_{-1}^{+1}\frac{\mathrm{d}x(n_1 - x)(n_2 - x)}{(x + d_1)(x + d_2)[(1 - x^2)(1 - \kappa^2 x^2)]^{1/2}}, \end{aligned} \tag{66}$$

where we have introduced

$$n_1 = \frac{A}{Ba'} - \frac{b'}{a'}, \qquad n_2 = \frac{A}{Ba'} - \frac{b'}{a'} - \frac{l}{a'},$$
$$d_1 = \frac{b'}{a'}, \qquad d_2 = \frac{b'}{a'} + \frac{l}{a'}, \qquad \kappa = \frac{1}{1 + l/a'}. \tag{67}$$

Likewise, we have

$$\frac{(n_1 - x)(n_2 - x)}{(x + d_1)(x + d_2)} = 1 + \frac{1}{d_2 - d_1}\left(\frac{\xi_1}{x + d_1} - \frac{\xi_2}{x + d_2}\right), \tag{68}$$

where we have introduced

$$\xi_1 = (d_1 + n_1)(d_1 + n_2),$$
$$\xi_2 = (d_2 + n_1)(d_2 + n_2). \tag{69}$$

However, the replacement

$$1 + \frac{1}{d_2 - d_1}\left(\frac{\xi_1}{x + d_1} - \frac{\xi_2}{x + d_2}\right) \rightarrow 1 + \frac{1}{d_2 - d_1}\left(\frac{d_1\xi_1}{d_1^2 - x^2} - \frac{d_2\xi_2}{d_2^2 - x^2}\right) \tag{70}$$

retains only the part of this expression that is even with respect to x, whereas its odd part does not contribute to results. After such a replacement, we may truncate the integration range for x to being from zero to one, if a compensating prefactor of two is affixed to the result. Hence, we have

$$I_B = \frac{2\beta^4 B^2 C^2 \kappa}{\pi^2 \zeta(4) a'} \int_0^{l_2 - l_1} \mathrm{d}l S(\beta, l) \left[\int_0^1 \frac{\mathrm{d}x}{[(1-x^2)(1-\kappa^2 x^2)]^{1/2}} + \left(\frac{\xi_1/d_1}{d_2 - d_1} \right) \int_0^1 \frac{\mathrm{d}x}{(1 - d_1^{-2} x^2)[(1-x^2)(1-\kappa^2 x^2)]^{1/2}} - \left(\frac{\xi_2/d_2}{d_2 - d_1} \right) \int_0^1 \frac{\mathrm{d}x}{(1 - d_2^{-2} x^2)[(1-x^2)(1-\kappa^2 x^2)]^{1/2}} \right]$$

$$= \frac{2\beta^4 B^2 C^2 \kappa}{\pi^2 \zeta(4) a'} \int_0^{l_2 - l_1} \mathrm{d}l S(\beta, l) \left[K(\kappa^2) + \left(\frac{\xi_1/d_1}{d_2 - d_1} \right) \Pi(d_1^{-2}, \kappa^2) - \left(\frac{\xi_2/d_2}{d_2 - d_1} \right) \Pi(d_2^{-2}, \kappa^2) \right], \tag{71}$$

where

$$\Pi(d^2, \kappa^2) = \int_0^1 \frac{\mathrm{d}t}{(1 - d^2 t^2)[(1-t^2)(1-\kappa^2 t^2)]^{1/2}} \tag{72}$$

is a complete elliptic integral of the third kind (with the indicated convention for the meaning of its arguments). The remaining single integration over l may be done numerically with relative ease, with the same change of variables and quadrature scheme as for I_X.

Figure 6 shows results for $(\mathrm{d}E/\mathrm{d}A_s)/(\mathrm{d}E_0/\mathrm{d}A_s)$ for $\beta = 0.1$ mm, $d_s = d_d = 100$ mm, and $R = 10$ mm (solid line).

4. Conclusion

This work has presented a means by which one can compute diffraction effects on total irradiance for a broadband source. While much of the work may be adapted to a variety of sources, in the immediate context the methodology developed has been done so with Planck sources in mind. I have successfully demonstrated the capacity for treating Fraunhofer diffraction by rectangular and circular apertures and slits, and Fresnel diffraction by circular apertures. Furthermore, a general mathematical framework and many functions and means to evaluate them have been presented that will also be useful in other contexts. Two topics of current interest to examine in due course are treating multiple diffraction and diffraction of synchrotron radiation.

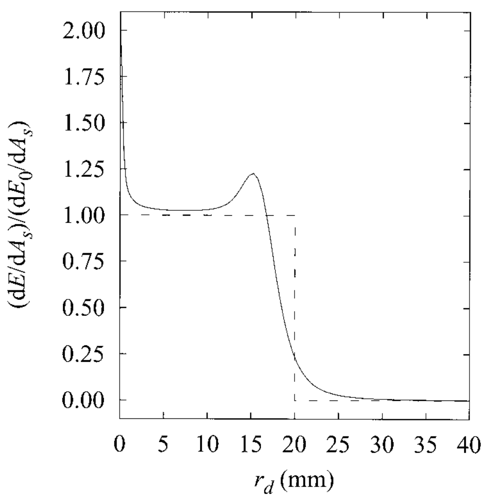

Fig. 6. Diffraction effects on total irradiance because of a circular aperture for a geometry discussed in the text versus distance from the optical axis, r_d. The solid curve shows results of numerical calculations, and the dashed line shows what is expected from geometrical optics.

Appendix A: Evaluating $S(\beta, l)$ and $S_m(z)$

The sum

$$S(\beta, l) = \sum_{n=1}^{\infty} [(n\beta + il)^{-4} + (n\beta - il)^{-4}] \tag{A1}$$

and related functions arise frequently in this work. Note that one may write

$$S(\beta, l) = l^{-4} S(\beta/l, 1) = \beta^{-4} S(1, l/\beta), \tag{A2}$$

and each manner of expressing $S(\beta, l)$ may be helpful in some context. Using

$$\begin{aligned}(b + il)^{-4} + (b - il)^{-4} &= \frac{(b - il)^4 + (b + il)^4}{(b^2 + l^2)^4} \\ &= \frac{2(b^4 - 6b^2 l^2 + l^4)}{(b^2 + l^2)^4} \\ &= \frac{2[(b^2 + l^2)^2 - 8l^2(b^2 + l^2) + 8l^4]}{(b^2 + l^2)^4} \\ &= 2/(b^2 + l^2)^2 - \frac{16 l^2}{(b^2 + l^2)^3} + \frac{16 l^4}{(b^2 + l^2)^4},\end{aligned} \tag{A3}$$

we find

$$S(\beta, l) = 2 \sum_{n=1}^{\infty} [(n\beta)^2 + l^2]^{-2} - 16 l^2 \sum_{n=1}^{\infty} [(n\beta)^2 + l^2]^{-3} + 16 l^4 \sum_{n=1}^{\infty} [(n\beta)^2 + l^2]^{-4}. \tag{A4}$$

Defining $z = 2\pi l/\beta$, we have

$$\begin{aligned}S(\beta, l) &= 2(2\pi/\beta)^4 S_2(z) - 16 l^2 (2\pi/\beta)^6 S_3(z) + 16 l^4 (2\pi/\beta)^8 S_4(z) \\ &= 2(2\pi/\beta)^4 [S_2(z) - 8z^2 S_3(z) + 8z^4 S_4(z)],\end{aligned} \tag{A5}$$

with

$$S_m(z) = \sum_{n=1}^{\infty} (4\pi^2 n^2 + z^2)^{-m}. \tag{A6}$$

From Knopp,[7] we have

$$S_1(z) = \frac{1}{2z}\left[\frac{1}{1-\exp(-z)} - \frac{1}{z} - \frac{1}{2}\right], \qquad \text{(A7)}$$

and $S_2(z)$ to $S_4(z)$ may be found by successively using

$$S_{m+1}(z) = -\frac{1}{2mz}\frac{\mathrm{d}}{\mathrm{d}z}S_m(z). \qquad \text{(A8)}$$

Letting $f = 1/[1-\exp(-z)] = \exp(z)/[\exp(z)-1]$, we may observe

$$\begin{aligned} f' &= \mathrm{d}f/\mathrm{d}z \\ &= \exp(z)/[\exp(z)-1] - \exp(2z)/[\exp(z)-1]^2 \\ &= f - f^2, \\ f'' &= \mathrm{d}^2f/\mathrm{d}z^2 = (f-f^2)(1-2f) = f - 3f^2 + 2f^3, \\ f''' &= \mathrm{d}^3f/\mathrm{d}z^3 \\ &= (f-f^2)(1-6f+6f^2) \\ &= f - 7f^2 + 12f^3 - 6f^4. \end{aligned} \qquad \text{(A9)}$$

Next we can arrive at the following expressions by successive differentiation, while we use the above results for f and its derivatives:

$$\begin{aligned} S_2(z) &= -\frac{1}{4z^2}f' + \frac{1}{4z^3}f - \frac{1}{2z^4} - \frac{1}{8z^3}, \\ S_3(z) &= \frac{1}{16z^3}f'' - \frac{3}{16z^4}f' + \frac{3}{16z^5}f - \frac{1}{2z^6} - \frac{3}{32z^5}, \\ S_4(z) &= -\frac{1}{96z^4}f''' + \frac{1}{16z^5}f'' - \frac{5}{32z^6}f' + \frac{5}{32z^7}f - \frac{1}{2z^8} \\ &\quad - \frac{5}{64z^7}. \end{aligned} \qquad \text{(A10)}$$

Combining these to obtain $S(\beta, l)$, we find

$$S(\beta, l) = -2(2\pi/\beta)^4\left[\frac{1}{2z^4} + \frac{f'''}{12}\right] = -l^{-4} + O(e^{-z}). \qquad \text{(A11)}$$

This form is suitable except when z is much smaller than 1. Using

$$(4\pi^2n^2 + z^2)^{-m} = \frac{1}{(4\pi^2)^m}\sum_{k=0}^{\infty} \times \frac{(m-1+k)![-z^2/(4\pi^2)]^k}{k!(m-1)!n^{2m+2k}}, \qquad \text{(A12)}$$

we also have

$$\begin{aligned} S_m(z) &= \frac{1}{(4\pi^2)^m}\sum_{k=0}^{\infty} \times \frac{(m-1+k)!\zeta(2m+2k)[-z^2/(4\pi^2)]^k}{k!(m-1)!} \\ &= \sum_{k=0}^{\infty} s_{m,k}z^{2k}, \end{aligned} \qquad \text{(A13)}$$

where $\zeta(N) = \sum_{n=1}^{\infty} n^{-N}$ is the Riemann zeta function, and we have introduced coefficients $\{s_{m,k}\}$. This latter expression for $S_m(z)$ is helpful for small z, where it converges after very few terms. To lowest order, we have

$$\begin{aligned} S_1(z) &= \frac{1}{24} - \frac{z^2}{1440} + \frac{z^4}{60480} - \ldots, \\ S_2(z) &= \frac{1}{1440} - \frac{z^2}{30240} + \frac{z^4}{806400} - \ldots, \\ S_3(z) &= \frac{1}{60480} - \frac{z^2}{806400} + \frac{z^4}{15966720} - \ldots, \\ S_4(z) &= \frac{1}{2419200} - \frac{z^2}{23950080} + \frac{691z^4}{261534873600} \\ &\quad - \ldots, \\ S(\beta, l) &= \frac{32\pi^4}{\beta^4}\left[\frac{1}{1440} - \frac{z^2}{6048} + \frac{z^4}{69120} - \ldots,\right], \end{aligned} \qquad \text{(A14)}$$

or

$$\begin{aligned} S(\beta, l) &= \frac{32\pi^4}{\beta^4}\Bigg[s_{2,0} + z^2(s_{2,1} - 8s_{3,0}) \\ &\quad + \sum_{k=2}^{\infty} z^{2k}(s_{2,k} - 8s_{3,k-1} + 8s_{4,k-2})\Bigg] \\ &= \sum_{k=0}^{\infty} s_k z^{2k}, \end{aligned} \qquad \text{(A15)}$$

where we have introduced coefficients $\{s_k\}$. All of these series converge for $|z| < 2\pi$, i.e., $|l/\beta| < 1$.

References

1. See, for example, J. D. Jackson, *Classical Electrodynamics* (Wiley, New York, 1975), pp. 427ff.
2. All of these examples are discussed in M. Born and E. Wolf, *Principles of Optics*, 7th (expanded) ed. (Cambridge U. Press, Cambridge, UK, 1999), Chap. 8.
3. See, for example, J. W. Goodman, *Introduction to Fourier Optics*, 2nd ed. (McGraw-Hill, New York, 1996), pp. 66–67.
4. See, for example, W. T. Welford, *Aberrations of the Symmetrical Optical System* (Academic, New York, 1974), p. 96.
5. K. Miyamoto and E. Wolf, "Generalization of the Maggi-Rubinowicz theory of the boundary diffraction wave—part I," J. Opt. Soc. Am. **52,** 612–625 (1962).
6. K. Miyamoto and E. Wolf, "Generalization of the Maggi-Rubinowicz theory of the boundary diffraction wave—part II," J. Opt. Soc. Am. **52,** 626–637 (1962).
7. K. Knopp, *Theory and Application of Infinite Series*, 2nd English ed. (Blackie and Son, London, 1946), p. 419.