

A Free-Space Diffraction BSDF: Supplemental

SHLOMI STEINBERG, University of Waterloo, Canada
 RAVI RAMAMOORTHI* and BENEDIKT BITTERLI, NVIDIA, United States
 ARSHIYA MOLLAZAINALI, University of Waterloo, Canada
 EUGENE D'EON, NVIDIA, New Zealand
 MATT PHARR, NVIDIA, United States

S1 PRELIMINARIES

The problem we tackle is the diffraction of a beam by an aperture consisting of an arbitrary polyhedron. Consider a beam with a smoothly-varying cross-section (a change in intensity over the transverse plane). Let that beam impinge upon a bounded, closed polyhedron, thereby the beam diffracts around the polyhedron's geometry. We assume that the polyhedron is perfectly conductive (it is opaque, and no significant amount of light refracts through it).

The incident beam. We assume the beam is a spatially-modulated plane wave, i.e. has constant phase on the *aperture plane* that is defined as the xy plane ($z = 0$), and its peak amplitude is modulated over that plane. Let the beam's direction of propagation is denoted $+\hat{z}$ (i.e., positive z -axis), and let its *peak amplitude function* be denoted φ . As an example, an incident beam with a common Gaussian modulation of its peak amplitude would take the following form on the aperture plane:

$$\varphi(\vec{q}^\perp) = \frac{E}{\sqrt{\pi}|\Lambda|^{1/4}} e^{-\frac{1}{2}(\vec{q}^\perp)^\top \Lambda^{-1} \vec{q}^\perp}, \quad (\text{S1.1})$$

for some constant $E \in \mathbb{C}$, and $\vec{q} \in \mathbb{R}^2$. The 2×2 matrix Λ describes the spatial spread of light's intensity over the xy -plane, and is typically a function of z (propagation distance), however we assume that it changes slowly enough over z that we let Λ remain constant. The total power contained in the beam is denoted \mathcal{P}_{in} , and for the Gaussian-modulated beam above it becomes

$$\mathcal{P}_{\text{in}} \triangleq \int_{\mathbb{R}^2} d^2\vec{q}^\perp |\varphi(\vec{q}^\perp)|^2 = |E|^2. \quad (\text{S1.2})$$

Note that Eq. (S1.1) is just an example: we may work with other beam profiles, as we will approximate φ piecewise-linearly over the diffracting polyhedron geometry, see Subsection S1.2.

*Also with University of California San Diego.

Authors' addresses: Shlomi Steinberg, p@shlomisteinberg.com, University of Waterloo, Waterloo, Canada; Ravi Ramamoorthi, ravir@cs.ucsd.edu; Benedikt Bitterli, benedikt.bitterli@gmail.com, NVIDIA, San Francisco, United States; Arshiya Mollazainali, arshiya.mollazainali@uwaterloo.ca, University of Waterloo, Waterloo, Canada; Eugene d'Eon, ejdeon@gmail.com, NVIDIA, Wellington, New Zealand; Matt Pharr, matt@pharr.org, NVIDIA, San Francisco, United States.

© 2024 Copyright held by the owner/author(s). Publication rights licensed to ACM. This is the author's version of the work. It is posted here for your personal use. Not for redistribution. The definitive Version of Record was published in . .

The diffracting polyhedron. Assume that the polyhedron is centered around the origin. We ignore faces that face away from the beam, i.e., we ignore faces whose normals lie in the $z \geq 0$ half-space, and we assume that no (front-facing) faces overlap when projected upon the xy -plane (e.g., a sufficient, but not necessary, condition is convexity of the polyhedron). We also accept disjoint (bounded, closed) polyhedra, i.e. a collection of one or more (bounded, closed) polyhedra that do not share faces. The projection of an arbitrary vector $\vec{u} \in \mathbb{R}^3$ onto the xy -plane is denoted as \vec{u}^\perp . We define $z(\vec{u}^\perp)$ to be the z coordinate of \vec{u} (before projection). As we assume no overlapping faces, this function is well defined.

We denote the projection of the diffracting shape, i.e. the collection of all (projected) front-facing faces, onto the xy -plane as $\overline{\mathcal{A}}^\perp$. The bar denotes the complement set. The complement of this projected diffracting shape is then \mathcal{A}^\perp , i.e. $\mathcal{A}^\perp \cup \overline{\mathcal{A}}^\perp = \mathbb{R}^2$ and $\mathcal{A}^\perp \cap \overline{\mathcal{A}}^\perp = \emptyset$.

S1.1 Fraunhofer Diffraction

We consider diffraction in the Fraunhofer (far-field) region, where the projected polyhedron acts as the diffracting aperture. This implies that the characteristic size of the diffracting polyhedron (before projection) is small compared to the distance from the origin (where diffraction happens) to the observation plane located at $z = R$. In the Fraunhofer region, diffraction can be formulated as a Fourier transform [Born and Wolf 1999]:

$$\psi(\vec{\xi}) \triangleq C \int_{\mathcal{A}^\perp} d^2\vec{q}^\perp \varphi(\vec{q}^\perp) e^{-ik\vec{\xi} \cdot \vec{q}^\perp}, \quad (\text{S1.3})$$

up to some constant C , with $\mathcal{A}^\perp \subset \mathbb{R}^2$ being the (assumed to be bounded) diffracting aperture, $\vec{\xi} = \frac{1}{R}[x', y']^\top$, where x', y' are positions on the observation plane, and $k = \frac{2\pi}{\lambda}$ is the wavenumber, with λ being wavelength. In the Fraunhofer region $R \gg x', R \gg y'$, therefore $\frac{x'}{R} = \tan \theta_x \approx \sin \theta_x$, where θ_x is the scattering direction from the scattering region at the origin, and similarly $\frac{y'}{R} \approx \sin \theta_y$.

The constant C arises from energy-conservation constraints [Born and Wolf 1999], where the total diffracted power must equal the total power incident upon the aperture:

$$\mathcal{P}_{\mathcal{A}^\perp} \triangleq \underbrace{\int_{\mathbb{R}^2} d^2\vec{\xi} R^2 |\psi(\vec{\xi})|^2}_{\text{total diffracted power that reaches the observation plane}} = \underbrace{\int_{\mathcal{A}^\perp} d^2\vec{q}^\perp |\varphi(\vec{q}^\perp)|^2}_{\text{total power incident upon the aperture}}. \quad (\text{S1.4})$$

Similarly to Born and Wolf [1999], we make above the approximation that $R^2 d^2\vec{\xi}$ is the differential area on the observation plane, because $|\psi(\vec{\xi})|^2$ is non-negligible only for $\vec{\xi} \ll R$.

To compute the integral in Eq. (S1.4), we apply the Plancherel theorem (we denote $\tilde{\psi}$ as the Fourier transform of ψ), and use Eq. (S1.3) to compute the the Fourier transform $\tilde{\psi}$:

$$\begin{aligned} \int_{\mathbb{R}^2} d^2 \vec{\xi} R^2 \left| \psi \left(\frac{\vec{\xi}}{R} \right) \right|^2 &= R^2 \int_{\mathbb{R}^2} d^2 \vec{\xi} \left| \tilde{\psi} \left(\frac{\vec{\xi}}{R} \right) \right|^2 \\ &= \left(\frac{2\pi R |C|}{k} \right)^2 \int_{\mathcal{A}^\perp} d^2 \vec{q}^\perp \left| \varphi \left(\vec{q}^\perp \right) \right|^2 \\ &= \left(\frac{2\pi R |C|}{k} \right)^2 \mathcal{P}_{\mathcal{A}^\perp}. \end{aligned} \quad (\text{S1.5})$$

Therefore, by substituting Eq. (S1.4) into the above, we deduce

$$C = \frac{k}{2\pi R}, \quad (\text{S1.6})$$

up to an irrelevant phase term.

S1.2 The Diffracting Aperture as a Triangular Mesh

We assume that the (projected) diffracting polyhedron \mathcal{A}^\perp consists of a given triangular mesh, that is $\mathcal{A}^\perp = \cup_t T_t$, where T_t are two-dimensional triangles on the aperture plane, and such that the triangles do not overlap, viz. $T_t \cap T_l = \emptyset$ for $t \neq l$. We also assume that each edge in the triangular mesh is either completely shared by exactly two distinct, adjacent triangles, or is an edge of the polyhedron \mathcal{A}^\perp (what is known as a *triangle strip* in computer graphics). Let a triangle T_t be defined by its vertices $\vec{u}_{t,1}, \vec{u}_{t,2}, \vec{u}_{t,3} \in \mathbb{R}^2$. We assume a non-degenerate triangle. We denote its edge vectors (that connects the edge's vertices) and outward-facing edge normals as

$$\vec{l}_{t,j} \triangleq \vec{u}_{t,j-1} - \vec{u}_{t,j+1} \quad \text{and} \quad \hat{m}_{t,j} \triangleq \frac{\vec{l}_{t,j} \times \hat{n}}{|\vec{l}_{t,j} \times \hat{n}|} \quad (\text{S1.7})$$

respectively, where the vertex indices should be understood in the modulo sense, i.e., $\vec{u}_{t,4} \equiv \vec{u}_{t,1}$ and so on, and \hat{n} is the triangle's face normal (for a counterclockwise-winded triangle, it is the back-face normal, i.e. $\hat{n} \equiv \hat{z}$). Note that we index the vectors above such that $\vec{l}_{t,j}$ and $\hat{m}_{t,j}$ are the edge and edge normal opposite of vertex $\vec{u}_{t,j}$. We also write the edge lengths as $l_{t,j} = |\vec{l}_{t,j}|$. We may drop the “ t ” index from the subscripts when it can be deduced from context.

Barycentric coordinates. We denote a triangle's *barycentric coordinates* as $[\gamma_1, \gamma_2, \gamma_3]^\top$, defined w.r.t. Cartesian coordinates x, y via the linear map:

$$\mathbf{B}_T \begin{pmatrix} \gamma_1 \\ \gamma_2 \\ \gamma_3 \end{pmatrix} = \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}, \quad \text{where} \quad \mathbf{B}_T \triangleq \begin{pmatrix} \vec{u}_1 & \vec{u}_2 & \vec{u}_3 \\ 1 & 1 & 1 \end{pmatrix}, \quad (\text{S1.8})$$

i.e., the first two rows of the matrix \mathbf{B}_T consist of the x and y Cartesian components of triangle's vertices \vec{u}_j , and 1 in the entire bottom row. It is important to remember that the barycentric coordinates γ_j are linearly dependant via $\sum_j \gamma_j = 1$. Let the *signed area* of the triangle be

$$S \triangleq \frac{1}{2} |\mathbf{B}_T|. \quad (\text{S1.9})$$

The signed area is positive if and only if the triangle's winding is counterclockwise.

Of interest is integration in barycentric coordinates:

$$\int_T d^2 \vec{q} f(\vec{q}) = 2|S| \int_0^1 d\gamma_1 \int_0^{1-\gamma_1} d\gamma_2 f(\gamma_1, \gamma_2, 1 - \gamma_1 - \gamma_2). \quad (\text{S1.10})$$

We also consider the gradient (w.r.t. the Cartesian coordinate system) of a scalar function defined in barycentric coordinates. Note that

$$\nabla_{\gamma_1, \gamma_2, \gamma_3} f = \begin{pmatrix} \partial f / \partial \gamma_1 \\ \partial f / \partial \gamma_2 \\ \partial f / \partial \gamma_3 \end{pmatrix} = \mathbf{B}_T^\top \begin{pmatrix} \partial f / \partial x \\ \partial f / \partial y \\ 0 \end{pmatrix} = \mathbf{B}_T^\top \begin{pmatrix} \nabla_{x,y} f \\ 0 \end{pmatrix} \quad (\text{S1.11})$$

by the chain rule, where the subscripts on the gradients indicate which coordinate system they are taken with respect to. Invert \mathbf{B}_T , express it in terms of edge lengths l_j and edge normals \hat{m}_j , and substitute it into Eq. (S1.11), solving the linear system, and leading to (equivalent) expressions for the gradient:

$$\begin{aligned} \nabla_{x,y} f &= -\frac{l_1 \hat{m}_1}{2S} \frac{\partial f}{\partial \gamma_1} - \frac{l_2 \hat{m}_2}{2S} \frac{\partial f}{\partial \gamma_2} \\ &= -\frac{l_2 \hat{m}_2}{2S} \frac{\partial f}{\partial \gamma_2} - \frac{l_3 \hat{m}_3}{2S} \frac{\partial f}{\partial \gamma_3} \\ &= -\frac{l_3 \hat{m}_3}{2S} \frac{\partial f}{\partial \gamma_3} - \frac{l_1 \hat{m}_1}{2S} \frac{\partial f}{\partial \gamma_1}, \end{aligned} \quad (\text{S1.12})$$

where we simplify using the facts that $\gamma_3 = 1 - \gamma_1 - \gamma_2$ and $-\vec{l}_3 = \vec{l}_1 + \vec{l}_2$, and hence $-l_3 \hat{m}_3 = l_1 \hat{m}_1 + l_2 \hat{m}_2$, by definition Eq. (S1.7) (i.e. the scaled edge normals can be understood to form a triangle congruent to T). Kim et al. [2021] derived similar results.

Piecewise-linear approximation. We now piecewise-linearly approximate (PLA) the peak amplitude function φ of the incident beam by linearly interpolating it over the triangular mesh that composes the diffracting polyhedron \mathcal{A}^\perp :

$$\varphi_{\text{pl}}(\vec{q}^\perp) \triangleq \begin{cases} \sum_j \gamma_{t,j} \varphi_{t,j} & \text{given } \exists t \text{ s.t. } \vec{q}^\perp \in T_t \\ 0 & \text{otherwise} \end{cases} \quad (\text{S1.13})$$

$$\text{with} \quad \varphi_{t,j} \triangleq \varphi(\vec{u}_{t,j}) e^{-ikz(\vec{u}_{t,j})},$$

t is the index of the (unique) triangle that contains \vec{q}^\perp (assuming the point is on the aperture), $\gamma_{t,j}$ are the barycentric coordinates of \vec{q}^\perp on the triangle, $\varphi_{t,j}$ are shorthands denoting the values φ takes on the vertices of the triangle, and where $z(\vec{u}_{t,j})$ is the z coordinate of the diffracting mesh at point $\vec{u}_{t,j}$ before projection. Note that we account for the depth light travels across the diffracting aperture.

The spatial beam profile φ may then take an arbitrary form, however our results remain accurate as long as the piecewise-linear approximation is a good approximation, i.e. $\varphi_{\text{pl}} \approx \varphi$. Of course, for a smooth function φ , the PLA approximation can be made arbitrary good by refining the tessellation of the triangular mesh.

S2 EDGE-DIFFRACTED WAVES

We will now derive a closed-form approximation for the Fraunhofer diffraction integral:

$$\psi(\vec{\xi}) = \frac{k}{2\pi R} \int_{\mathcal{A}^\perp} d^2\vec{q}^\perp \varphi(\vec{q}^\perp) g(\vec{q}^\perp, \vec{\xi}), \quad (\text{S2.1})$$

$$\text{with } g(\vec{q}^\perp, \vec{\xi}) \triangleq e^{-ik\vec{\xi} \cdot \vec{q}^\perp} \quad (\text{S2.2})$$

(as in Eq. (S1.3), but with the aperture being the projected polyhedron). As discussed in Subsection 4.1, we use the complement of the polyhedron to diffract for practical reasons. For succinctness of notation, above we denote the Fourier kernel as g .

Recall that we may understand the aperture as a triangular mesh, viz. $\mathcal{A}^\perp = \cup_t T_t$, as discussed in Subsection S1.2. The integral in Eq. (S2.1) then becomes a sum of integrals over the triangles T_t :

$$\psi(\vec{\xi}) = \frac{k}{2\pi R} \sum_t \int_{T_t} d^2\vec{q}^\perp \varphi g. \quad (\text{S2.3})$$

We now consider the integral over some triangle T . We replace φ with its piecewise-linear approximant φ_{PL} , as defined in Eq. (S1.13). Note that $\nabla \cdot (\vec{\xi}g) = -ik\xi^2 g$ and $\nabla^2 g = -k^2 \xi^2 g$, where $\nabla \cdot$ and ∇^2 denote the divergence and Laplacian operators, respectively, hence

$$g\varphi_{\text{PL}} = \frac{i}{k\xi^2} \varphi_{\text{PL}} \nabla \cdot (\vec{\xi}g) \quad \text{and} \quad g\vec{\xi} \cdot \nabla \varphi_{\text{PL}} = -\frac{\nabla^2 g}{k^2 \xi^2} \vec{\xi} \cdot \nabla \varphi_{\text{PL}}, \quad (\text{S2.4})$$

where $\xi = |\vec{\xi}|$. Then, applying all of the above and using the divergence theorem (we explicitly indicate the vector field \vec{F} under each term to which we apply the divergence theorem), we may write the diffraction integral over a triangle as

$$\begin{aligned} \int_T d^2\vec{q}^\perp g\varphi_{\text{PL}} &= \frac{i}{k\xi^2} \int_T d^2\vec{q}^\perp \left(\underbrace{\varphi_{\text{PL}} \nabla \cdot (\vec{\xi}g)}_{\vec{F} = \vec{\xi}g\varphi_{\text{PL}}} + \underbrace{g\vec{\xi} \cdot \nabla \varphi_{\text{PL}} + \nabla^2 g \frac{\vec{\xi} \cdot \nabla \varphi_{\text{PL}}}{k^2 \xi^2}}_{\vec{F} = \nabla g} \right) \\ &= \frac{i}{k\xi^2} \oint_{\delta T} ds \hat{\mathbf{m}} \cdot (\vec{\xi}g\varphi_{\text{PL}}) + \frac{i\vec{\xi} \cdot \nabla \varphi_{\text{PL}}}{k^3 \xi^4} \oint_{\delta T} ds \hat{\mathbf{m}} \cdot \nabla g, \end{aligned} \quad (\text{S2.5})$$

where δT is the boundary of the triangle (parametrized counterclockwise), and $\hat{\mathbf{m}}$ is the outward-pointing normal. We are able to take the gradient $\nabla \varphi_{\text{PL}}$ out of the integral, as it is constant over the triangle (φ_{PL} is linear, by construction):

$$\nabla \varphi_{\text{PL}} = -\frac{1}{2S} [(\varphi_1 - \varphi_3)l_1 \hat{\mathbf{m}}_1 + (\varphi_2 - \varphi_3)l_2 \hat{\mathbf{m}}_2], \quad (\text{S2.6})$$

where we used Eq. (S1.12).

It should be noted that a typical formulation of the divergence theorem requires the vector field to be continuously differentiable in a neighbourhood of the triangle T , however clearly the derivative of φ_{PL} is undefined on the edges of the triangle. φ_{PL} can always be extended to be defined over a triangle enlarged by ϵ , and understand the result in Eq. (S2.5) in the limit $\epsilon \rightarrow 0$.

From Eq. (S2.5) we may deduce:

$$\int_T d^2\vec{q}^\perp g\varphi_{\text{PL}} = \frac{i}{k\xi^2} \oint_{\delta T} ds \left(\varphi_{\text{PL}} - i \frac{\vec{\xi} \cdot \nabla \varphi_{\text{PL}}}{k\xi^2} \right) g\hat{\mathbf{m}} \cdot \vec{\xi}. \quad (\text{S2.7})$$

Consider the two (highlighted) terms inside the parentheses on the right-hand side in the equation above. We assume that $|\varphi_{\text{PL}}| \gg 0$, i.e.

the incident light's intensity over the triangle, is non-negligible, as otherwise the diffraction over this triangle contributes little to the total diffracted field. Furthermore, the product $k\xi$ is very small only within the direct term (Subsection 4.1) of the diffraction pattern. For the diffracted terms $k\xi \gg 0$. Finally, we assume that $|\nabla \varphi_{\text{PL}}| \ll 1$, i.e. light's intensity varies slowly over the triangle—a reasonable assumption for many light profiles. Therefore,

$$|\varphi_{\text{PL}}| \gg \frac{1}{k\xi} |\nabla \varphi_{\text{PL}}|. \quad (\text{S2.8})$$

Usually, in most settings of interest, in the equation above the left-hand side term dominates the right-hand side term by multiple orders-of-magnitude.

Then, based on the above, we drop the second term (in blue) in Eq. (S2.7), yielding:

$$\int_T d^2\vec{q}^\perp g\varphi_{\text{PL}} \approx \frac{i}{k\xi^2} \oint_{\delta T} ds (\hat{\mathbf{m}} \cdot \vec{\xi}) \varphi_{\text{PL}} g. \quad (\text{S2.9})$$

The integral over the boundary can be trivially written as a sum of line integrals over each of the triangle's edges. Substituting into Eq. (S2.3) then results in a sum of line integrals over all edges of all triangles. However, recall that every edge in our triangular mesh is either an edge of the aperture \mathcal{A}^\perp or is completely shared by exactly two adjacent triangles (as defined in Subsection S1.2). Hence, it is easy to see that the line integrals over a shared edge of adjacent triangles cancel. To formally see that, consider Eq. (S2.16) and substitute $\hat{\mathbf{m}}_j \rightarrow -\hat{\mathbf{m}}_j$, $\vec{e}_j \rightarrow -\vec{e}_j$, $a_j \rightarrow b_j$, $b_j \rightarrow a_j$, and recall that sinc and cos are even functions. We must conclude that the only line integrals that do not annihilate are these that are not shared, i.e. exactly those that form the boundary of \mathcal{A}^\perp :

$$\psi(\vec{\xi}) = \frac{k}{2\pi R} \sum_t \int_{T_t} d^2\vec{q}^\perp \varphi g = \frac{i}{2\pi R \xi^2} \oint_{\delta \mathcal{A}^\perp} ds (\hat{\mathbf{m}} \cdot \vec{\xi}) \varphi_{\text{PL}} g. \quad (\text{S2.10})$$

We assume counterclockwise parametrization of the boundary $\delta \mathcal{A}^\perp$.

We will now consider a line integral over a single edge of the boundary of the projected polyhedron $\delta \mathcal{A}^\perp$. Let the j -th edge be described by its centre point \vec{v}_j ; its edge vector \vec{e}_j that connects the edge's vertices, and hence $e_j = |\vec{e}_j|$ is the edge length; and its outward-facing edge normal $\hat{\mathbf{m}}_j$ (it should be clear context whether $\hat{\mathbf{m}}$ refers to the outward-facing normals of a triangle or that of the aperture). Then, we define the *edge-diffracted wave*:

$$\psi_j(\vec{\xi}) \triangleq \frac{i}{2\pi R} \frac{e_j}{\xi^2} (\hat{\mathbf{m}}_j \cdot \vec{\xi}) \int_{-1/2}^{1/2} ds \varphi_{\text{PL}}(\vec{q}^\perp) e^{-ik\vec{\xi} \cdot \vec{q}^\perp}, \quad (\text{S2.11})$$

where we now have $\vec{q}^\perp = \vec{v}_j + s\vec{e}_j$. Note that the dot product $\hat{\mathbf{m}}_j \cdot \vec{\xi}$ is constant over the line integral. The total diffracted field (Eq. (S2.10)) becomes

$$\psi = \sum_j \psi_j. \quad (\text{S2.12})$$

The piecewise-linear approximant φ_{PL} in Eq. (S2.11) reduces to a linear interpolation over the edge between the values φ takes on

the edge's vertices, denoted a_j and b_j , respectively, viz.

$$\varphi_{\text{PL}}(\vec{q}^\perp) = \left(s + \frac{1}{2}\right)b_j + \left(\frac{1}{2} - s\right)a_j, \quad (\text{S2.13})$$

$$\text{with } a_j \triangleq \varphi\left(\vec{v}_j - \frac{1}{2}\vec{e}_j\right)e^{-ikz\left(\vec{v}_j - \frac{1}{2}\vec{e}_j\right)} \quad \text{and} \quad (\text{S2.14})$$

$$b_j \triangleq \varphi\left(\vec{v}_j + \frac{1}{2}\vec{e}_j\right)e^{-ikz\left(\vec{v}_j + \frac{1}{2}\vec{e}_j\right)}, \quad (\text{S2.15})$$

and with $s \in [-\frac{1}{2}, \frac{1}{2}]$. We substitute the above into the expression for the edge-diffracted wave (Eq. (S2.11)), and integrate:

$$\begin{aligned} \psi_j(\vec{\xi}) &\triangleq \frac{1}{2\pi R} \frac{e_j}{\xi^2} (\hat{\mathbf{m}}_j \cdot \vec{\xi}) e^{-ik\vec{\xi} \cdot \vec{v}_j} \\ &\times \left[\frac{a_j - b_j}{k\vec{\xi} \cdot \vec{e}_j} \left(\cos \frac{k\vec{\xi} \cdot \vec{e}_j}{2} - \text{sinc} \frac{k\vec{\xi} \cdot \vec{e}_j}{2} \right) + i \frac{a_j + b_j}{2} \text{sinc} \frac{k\vec{\xi} \cdot \vec{e}_j}{2} \right]. \end{aligned} \quad (\text{S2.16})$$

The *primary* and *canonical* spaces. Because \vec{e}_j and $\hat{\mathbf{m}}_j$ are orthogonal, we may write

$$e_j \frac{\hat{\mathbf{m}}_j \cdot \vec{\xi}}{\xi^2} = (ke_j)^2 \frac{1}{k} \frac{ke_j \hat{\mathbf{m}}_j \cdot \vec{\xi}}{\left(ke_j \hat{\mathbf{m}}_j \cdot \vec{\xi}\right)^2 + \left(k\vec{e}_j \cdot \vec{\xi}\right)^2}, \quad (\text{S2.17})$$

which motivates us to define the transform

$$\Xi_j^{-1} \triangleq k \begin{pmatrix} (e_j)_x & (e_j)_y \\ (e_j)_y & -(e_j)_x \end{pmatrix} = k \begin{pmatrix} | & | \\ \vec{e}_j & e_j \hat{\mathbf{m}}_j \\ | & | \end{pmatrix}, \quad (\text{S2.18})$$

where $(e_j)_{x,y}$ are the Cartesian components of \vec{e}_j . The transform Ξ_j scales space by $\frac{1}{ke_j} = \frac{\lambda}{2\pi e_j}$ and rotates space such that the edge lies in direction of the y -axis and its boundary normal lies in direction of the x -axis. That is, $\Xi_j(k\vec{e}_j) = (0, 1)^\top$ and $\Xi_j(ke_j \hat{\mathbf{m}}_j) = (1, 0)^\top$. Recall that the boundary $\delta\mathcal{A}^\perp$ is parametrized counterclockwise, therefore $\hat{\mathbf{m}}_j$ must always point to the right of \vec{e}_j , i.e. out of \mathcal{A}^\perp .

Substitute Eqs. (S2.17) and (S2.18) into Eq. (S2.16), yielding the final expression for the edge-diffracted wave:

$$\psi_j(\vec{\xi}) = \frac{ke_j^2}{R} e^{-ik\vec{\xi} \cdot \vec{v}_j} \left[(a_j - b_j) \alpha_1(\Xi_j^{-\top} \vec{\xi}) + i \frac{a_j + b_j}{2} \alpha_2(\Xi_j^{-\top} \vec{\xi}) \right], \quad (\text{S2.19})$$

with the following important auxiliary functions:

$$\alpha_1(\vec{\zeta}) \triangleq \frac{\zeta_y}{2\pi\zeta^2\zeta_x} \left(\cos \frac{\zeta_x}{2} - \text{sinc} \frac{\zeta_x}{2} \right) \quad (\text{S2.20})$$

$$\text{and } \alpha_2(\vec{\zeta}) \triangleq \frac{\zeta_y}{2\pi\zeta^2} \text{sinc} \frac{\zeta_x}{2}, \quad (\text{S2.21})$$

where $\zeta_{x,y}$ are the Cartesian components of the argument $\vec{\zeta}$, and $\zeta^2 = \vec{\zeta} \cdot \vec{\zeta}$, as before. We highlight in colour the terms in Eqs. (S2.16) and (S2.17) that correspond to the terms highlighted in the same colour in Eqs. (S2.20) and (S2.21).

We designate as *primary space* the space where our edge $\{\vec{v}_j, \vec{e}_j\}$, as well as the wavelength-dependant edge-diffracted wave ψ_j , induced by that edge, live. It is in that space that we wish to render the diffraction pattern $|\psi|^2$ produced by the superposition of all edges. We will keep using $\vec{\xi}$ to denote points on the observation

plane in primary space. For a given edge, we denote as *canonical space* the space that arises from the linear transformation

$$\vec{\zeta} = \Xi_j^{-\top} \vec{\xi}, \quad (\text{S2.22})$$

and where the wavenumber is a fixed $k = 1$. Eq. (S2.19) shows that an edge-diffracted wave can then always be written as a superposition of linearly-transformed α_1 and α_2 , which are generic functions that depend neither on geometry nor light's properties. α_1, α_2 then serve as the *generic building blocks that compose any diffraction pattern* that is produced by an arbitrary aperture \mathcal{A}^\perp illuminated by light of any wavenumber k . We will use $\vec{\zeta}$ to refer to coordinates on the observation plane in canonical space.

S2.1 Extracting the 0th-Order Diffraction Lobe

As described in Subsection 4.1, our rendering algorithm requires extracting the central, 0th-order lobe (i.e., energy that propagates into the same directions as the direct term, Eq. (11)) out of the diffracted field ψ . We do so in this Subsection.

It is well known that Gaussians whose space-spectrum products fulfil the equality in the uncertainty relation Eq. (A.1) have the most compact support in space-spectrum space (i.e., the optical phase space) [Torre 2005]. In Appendix A, we prove Lemma A.1 that shows that the space-spectrum bandwidth product is invariant under transformation between the canonical and primary spaces. Thus, our aim is to extract a central Gaussian lobe that fulfils the equality in the uncertainty relation in canonical space out of $\alpha_{1,2}$, and a consequence of Lemma A.1 is that the extracted lobe retains the same (minimal) space-spectrum bandwidth product in primary space as well.

To that end, let the intensity distribution of the central, 0th-order lobe of the diffraction pattern in canonical space be the Gaussian

$$\chi_{\sigma_\zeta}(\vec{\zeta}) \triangleq e^{-\frac{1}{2\sigma_\zeta^2} \zeta^2}, \quad (\text{S2.23})$$

under the condition that

$$\sigma_\zeta = \frac{1}{2\sigma'_q}. \quad (\text{S2.24})$$

The equation above ensures that χ is the most compact lobe physically permissible, i.e., χ is a Gaussian that fulfils the equality in the uncertainty relation.

The variance on the aperture in canonical space, denoted σ'_q , is the characteristic length of that aperture, i.e. an edge length of 1. Then,

$$\left(\sigma'_q\right)^2 = \frac{1}{12}, \quad (\text{S2.25})$$

i.e. the variance of a continuous distribution of length 1. Substituting Eqs. (S2.24) and (S2.25) into Eq. (S2.23) yields the expression for the 0th-order lobe field in canonical space. Then, let an edge-diffracted wave ψ_j with its 0th-order lobe removed be denoted as

$$\hat{\psi}_j(\vec{\zeta}) \triangleq \sqrt{1 - \chi_{\sigma_\zeta}(\Xi_j^{-\top} \vec{\zeta})} \psi_j(\vec{\zeta}), \quad (\text{S2.26})$$

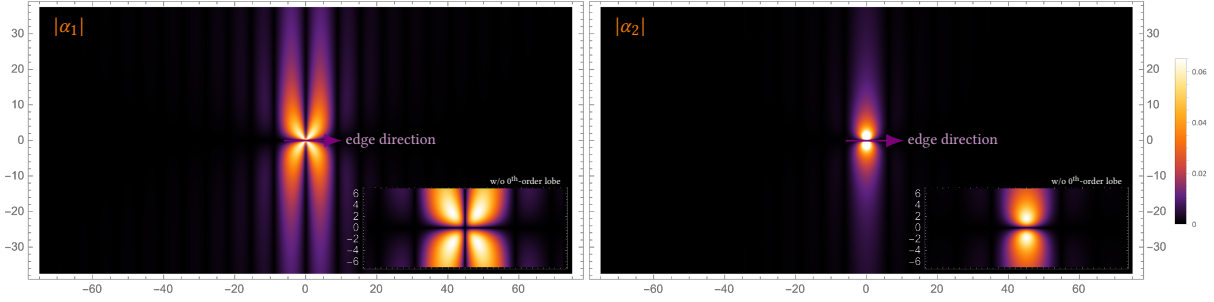


Fig. 1. **The auxiliary functions α_1 and α_2 in canonical space.** We plot the patterns encoded by $|\alpha_{1,2}|$, i.e. square roots of the intensities (square root is taken for visualization purposes). In canonical space, the edge is centred at the origin, has a length of 1, and points in direction of the x -axis (direction illustrated using the violet arrow). α_2 corresponds to diffraction by an edge where light's peak amplitude is constant across the edge's length, therefore diffracted light is strongly concentrated in the direction tangent to the edge, as expected. α_1 corresponds to diffraction with varying amplitude across the edge's length, hence spreads more of the diffracted energy to the sides and admits more pronounced fringes. (insets) α_1 and α_2 with their 0th-order lobe removed, i.e. we plot $(1 - \chi(\vec{\zeta}))^{1/2} |\alpha_{1,2}(\vec{\zeta})|$. Notice the missing energy at the very centre of the pattern.

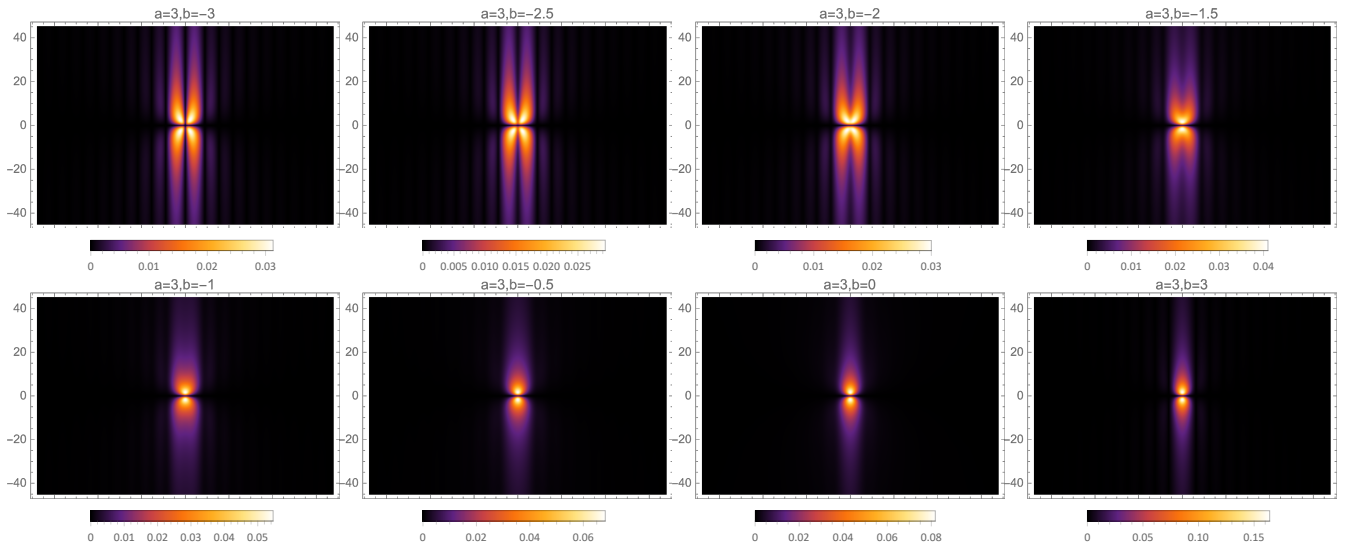


Fig. 2. **Edge-diffraction patterns in canonical space for various incident light profiles.** We plot the square root of intensity of the patterns $|(a - b)\alpha_1(\vec{\zeta}) + i\frac{a+b}{2}\alpha_2(\vec{\zeta})|$ (as in Eq. (S2.19)), where a is the peak amplitude on the left side of the edge, and b is the peak amplitude on its right side. We set $a = 3$ and vary b . We note that the pattern slowly morphs from the wider-scattering behaviour of α_1 to a pattern dictated by α_2 , where energy is almost-entirely confined to the tangent direction.

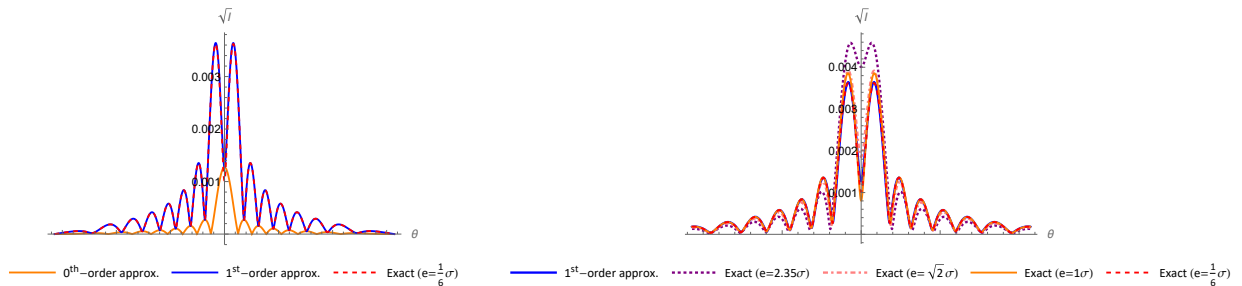


Fig. 3. **Analysis of amplitude approximation.** Plots of square root of intensity as functions of diffracting angle θ ($\theta = 0$ aligns with the y -axis in canonical space). We set the peak amplitudes $a = 3$ and $b = -2$, with a, b as in Fig. 2 and in text. (left) Comparison of 0th-order approximation (constant peak amplitude), our 1st-order approximation (piecewise linear) and exact (peak amplitude varies as a Gaussian with standard deviation σ). 0th-order approximation is equivalent to discarding α_1 and only using the α_2 term in Eq. (S2.19). Edge length is taken to be $e = \frac{1}{6}\sigma$. Note that the 0th-order approximation (constant amplitude) is insufficient. (right) Analysis of our 1st-order approximation (PLA) via comparison with the exact solution for various ratios $\frac{e}{\sigma}$. We note that the PLA begins to deteriorate roughly once the edge length e surpasses the standard deviation σ and breaks down entirely at $e = \text{FWHM} (\approx 2.35\sigma)$.

where we use Eq. (S2.22) to transform $\vec{\xi}$ to the canonical space. The total diffracted field, with the 0th-order lobe removed, is then

$$\widehat{\psi} = \sum_j \widehat{\psi}_j. \quad (\text{S2.27})$$

Discarding the 0th-order lobe out of the edge-diffracted wave and its associated pattern is crucial for our rendering algorithm. By construction, Eqs. (S2.23) and (S2.24) ensure that the removed 0th-order lobe in canonical space is a Gaussian with a space-spectrum bandwidth product that is the least permissible by the uncertainty principle. Lemma A.1, ensures that it remains a minimum-uncertainty Gaussian after transformation to primary space, i.e. to the pattern produced by an arbitrary edge, under illumination by light of any wavenumber. Such minimum-uncertainty Gaussians occupy the least physically-permissible space-spectrum volume [Torre 2005], making the removed 0th-order lobe as compact as possible. Resolving details under the uncertainty limit is not possible, we are guaranteed that the removed lobe must align with the direct term d .

The converse of the statement above does not have to hold: that is, some of the energy contained in $\widehat{\psi}$ might still propagate in directions of the direct term. That is fine: as we render the diffracted term, some of its energy will overlap, correctly, with the energy of the direct term. We only remove as much energy out of the edge-diffracted wave ψ_j as can formally be guaranteed to always coincide with the direct term, for any diffracting polyhedron \mathcal{A}^\perp , and under illumination by light of any wavenumber.

Width of the central, 0th-order lobe. The full width at half maximum (FWHM) of the Gaussian χ_{σ_ζ} , with $\sigma_\zeta = \sqrt{3}$, in canonical space is about $2.355\sqrt{3}$. After transforming to the primary space of an edge-diffraction pattern (via Eq. (S2.22)), the FWHM becomes

$$\begin{aligned} \text{FWHM} &= |\Xi_j|^{1/2} \times \text{FWHM}_{\text{canonical}} \\ &= \frac{1}{ke_j} 2.355\sqrt{3} \approx 0.65 \frac{\lambda}{e_j}. \end{aligned} \quad (\text{S2.28})$$

That is, the width of the 0th-order lobe scales linearly with wavelength and the reciprocal of the edge length. As to be expected due to the uncertainty relation: Longer wavelengths and smaller apertures induce greater uncertainty on the observation plane.

S2.2 Intensity and Power Contained in an Edge-Diffracted Wave

The intensity of an edge-diffracted wave (Eq. (S2.19)) is denoted as

$$\begin{aligned} I_j(\vec{\xi}) &\triangleq \left| \widehat{\psi}_j(\vec{\xi}) \right|^2 \\ &= \frac{|\Xi_j|^{-2}}{k^2 R^2} \left| (a_j - b_j) \alpha_1(\vec{\xi}) + i \frac{a_j + b_j}{2} \alpha_2(\vec{\xi}) \right|^2, \end{aligned} \quad (\text{S2.29})$$

where we use the fact $|\Xi|^{-1} = k^2 e_j^2$ and $\zeta = \Xi_j^{-\top} \vec{\xi}$, as before. The intensity of an edge-diffracted wave with its 0th-order lobe removed (Eq. (S2.26)) is

$$\widehat{I}_j(\vec{\xi}) \triangleq \left| \widehat{\psi}_j(\vec{\xi}) \right|^2 = \left(1 - \chi_{\sigma_\zeta}(\vec{\xi}) \right) \left| \psi_j(\vec{\xi}) \right|^2. \quad (\text{S2.30})$$

Let the total power contained in the total diffracted field $\psi = \sum_j \psi_j$ be denoted as $\mathcal{P}_{\mathcal{A}^\perp}^{(\text{PL})}$, where the superscript makes it explicit

that we use the piecewise-linear approximant (PLA) φ_{PL} for the incident field of light. Energy conservation mandates that this power must be equal to the power incident upon the aperture, under the PLA approximation (in similar manner to Eq. (S1.4)), viz.

$$\begin{aligned} \mathcal{P}_{\mathcal{A}^\perp}^{(\text{PL})} &\triangleq R^2 \int_{\mathbb{R}^2} d^2 \vec{\xi} \left| \psi(\vec{\xi}) \right|^2 = \int_{\mathcal{A}^\perp} d^2 \vec{q}^\perp \left| \varphi_{\text{PL}}(\vec{q}^\perp) \right|^2 \\ &= \sum_t \int_{T_t} d^2 \vec{q}^\perp \left| \varphi_{\text{PL}}(\vec{q}^\perp) \right|^2, \end{aligned} \quad (\text{S2.31})$$

$\triangleq \mathcal{P}_{T_t}^{(\text{PL})}$

where we have rewritten the integral over the aperture as a sum of integrals over the triangles T_t comprising \mathcal{A}^\perp . Recall the definitions of φ_{PL} and $\varphi_{t,j}$ (Eq. (S1.13)), and integrate in barycentric coordinates (using Eq. (S1.10)):

$$\begin{aligned} \mathcal{P}_{T_t}^{(\text{PL})} &= \int_{T_t} d^2 \vec{q}^\perp \left| \varphi_{\text{PL}}(\vec{q}^\perp) \right|^2 = 2|S_t| \int_0^1 d\gamma_1 \int_0^{1-\gamma_1} d\gamma_2 \left| \sum_{j=1}^3 \gamma_j \varphi_{t,j} \right|^2 \\ &= \frac{|S_t|}{6} \text{Re} \sum_{1 \leq j \leq l \leq 3} \varphi_{t,j} \varphi_{t,l}^*, \end{aligned} \quad (\text{S2.32})$$

where S_t is the signed area (Eq. (S1.9)) of the triangle T_t , Re is the real-part operator, and $\gamma_3 = 1 - \gamma_1 - \gamma_2$. Eq. (S2.32) gives a closed-form expression for $\mathcal{P}_{T_t}^{(\text{PL})}$, i.e. the power incident upon a triangle, under the PLA approximation, and summing over all triangles (Eq. (S2.31)) yields the total power $\mathcal{P}_{\mathcal{A}^\perp}^{(\text{PL})}$.

For importance sampling purposes, we are also interested in the total power contained in an edge-diffracted wave, with its 0th-order lobe removed, which we denote as

$$\begin{aligned} \widehat{\mathcal{P}}_j^{(\text{PL})} &\triangleq R^2 \int_{\mathbb{R}^2} d^2 \vec{\xi} \left| \widehat{\psi}_j(\vec{\xi}) \right|^2 = R^2 |\Xi_j^\top| \int_{\mathbb{R}^2} d^2 \vec{\zeta} \left| \widehat{\psi}_j(\Xi_j^\top \vec{\zeta}) \right|^2 \\ &= e_j^2 \left[|a_j - b_j|^2 \mathfrak{S}_1 + \left| \frac{a_j + b_j}{2} \right|^2 \mathfrak{S}_2 + 2(b_j \text{Im } a_j - a_j \text{Im } b_j) \mathfrak{S}_{12} \right], \end{aligned} \quad (\text{S2.33})$$

$$\text{where } \mathfrak{S}_{1,2} \triangleq \int_{\mathbb{R}^2} d^2 \vec{\zeta} \left(1 - \chi_{\sigma_\zeta}(\vec{\zeta}) \right) \alpha_{1,2}(\vec{\zeta})^2, \quad (\text{S2.34})$$

$$\text{and } \mathfrak{S}_{12} \triangleq \int_{\mathbb{R}^2} d^2 \vec{\zeta} \left(1 - \chi_{\sigma_\zeta}(\vec{\zeta}) \right) \alpha_1(\vec{\zeta}) \alpha_2(\vec{\zeta}), \quad (\text{S2.35})$$

and we performed the variable exchange $\vec{\xi} \rightarrow \Xi_j^\top \vec{\zeta}$. The integrals $\mathfrak{S}_{1,2,12}$ above are difficult to compute analytically, instead we compute them numerically to eight digits of precision:

$$\mathfrak{S}_1 \approx 0.0045255085, \quad \mathfrak{S}_2 \approx 0.11487543, \quad (\text{S2.36})$$

$$\text{and } \mathfrak{S}_{12} \approx 9.06 \times 10^{-11}. \quad (\text{S2.37})$$

Because the term \mathfrak{S}_{12} is miniscule, in the paper we neglect that term in the expression for $\widehat{\mathcal{P}}_j^{(\text{PL})}$, for simplicity.

Total power in the diffracted term. We turn our attention to the diffracted term w , as defined in Eq. (12). In Appendix B we derive an approximative expression for the total power contained in the 0th-order lobe, \widehat{P}_0 , see Eq. (B.16). Then, the total power contained in the diffracted term is

$$\widehat{\mathcal{P}}_{\mathcal{A}^\perp}^{(\text{PL})} \triangleq \mathcal{P}_{\mathcal{A}^\perp}^{(\text{PL})} - \widehat{P}_0^{(\text{PL})}. \quad (\text{S2.38})$$

S3 ADDITIONAL INFORMATION

Ratio of diffracted energy. Clearly, as $\mathcal{P}_{\mathcal{A}^\perp}^{(PL)} \geq \widehat{\mathcal{P}}_{\mathcal{A}^\perp}^{(PL)}$ (Eqs. (S2.31) and (S2.38)), the BSDF is conserving energy, and we may define the *ratio of diffracted energy* as

$$\mathcal{D} \triangleq \int_{\mathbb{R}^2} d^2 \xi f \cos \theta = \widehat{\mathcal{P}}_{\mathcal{A}^\perp}^{(PL)} / \mathcal{P}_{\mathcal{A}^\perp}^{(PL)} \leq 1. \quad (\text{S3.1})$$

The fraction of energy “absorbed” by the BSDF, i.e. $1 - \mathcal{D}$, accounts for the energy that we have taken out of the diffracted field in Subsection S2.1: the 0th-order lobe that aligns with the direct term. Recall that to render the diffracted term, our rendering algorithm needs to “lend” the energy for that term from the energy that falls upon the diffracting aperture \mathcal{A}^\perp and is back-scattered or absorbed. The total power that we need to lend is exactly $\mathcal{D} \cdot \mathcal{P}_{\mathcal{A}^\perp}^{(PL)}$.

Quality of the importance sampling strategy. Consider the intensity of the total field, Eq. (39). Intuitively, if most of the total power is contained in the edge-diffracted lobes, i.e. $\widehat{\mathcal{P}}_{\mathcal{A}^\perp}^{(PL)} \approx \sum_j \mathcal{P}_j^{(PL)}$, then, on average, the shape of the diffraction pattern will be dictated by the edge-diffracted lobes and not the interference factors. On the other hand, when the ratio $\widehat{\mathcal{P}}_{\mathcal{A}^\perp}^{(PL)} / \sum_j \mathcal{P}_j^{(PL)}$ is much larger or smaller than 1, then the interference factors take over and our importance sampling strategy becomes less effective.

The above motivates us to define a heuristic quality metric:

$$\mathcal{Q} \triangleq \mathcal{D} \frac{\mathcal{P}_{\mathcal{A}^\perp}^{(PL)}}{\sum_j \mathcal{P}_j^{(PL)}} = \frac{\widehat{\mathcal{P}}_{\mathcal{A}^\perp}^{(PL)}}{\sum_j \mathcal{P}_j^{(PL)}}. \quad (\text{S3.2})$$

Our importance sampling strategy is considered good when $\mathcal{Q} \approx 1$. We are able to exercise a measure of control over \mathcal{Q} : Because the power contained in an edge-diffracted wave $\mathcal{P}_j^{(PL)}$ is quadratic in the edge’s length e_j (see Eq. (S2.33)), splitting an edge into 2 edges of equal length results in about $\frac{1}{2} \mathcal{P}_j^{(PL)}$ reduction (up to the values of the amplitudes a_j, b_j) of the sum in the denominator in Eq. (S3.2). Indeed, by recursively splitting the longest edge we may make that sum arbitrarily small. And vice versa, merging edges increases said sum.

Therefore, when $\mathcal{Q} \neq 1$, our importance sampling strategy can be improved by using a preprocess pass:

- (1) While $\mathcal{Q} \ll 1$, identify a long edge and split it. This also serves to make our PLA approximation more accurate.
- (2) On the other hand, while $\mathcal{Q} \gg 1$, attempt to merge the smallest edge into one of its adjacent edges. Note that because the total power of an edge-diffracted wave is quadratic in its edge length, small edges contribute little to the total diffracted field.

Accuracy of the direct term. The direct term is computed via ray optics, and therefore predicts an energy distribution that is just as accurate as classical ray optics. It is easy to see that it is not possible to do better with just linear ray tracing (unless an oracle is able to predict when a ray is passing through an aperture or close to geometry). For many applications, the error in the direct term is negligible: e.g., overestimating the direct term makes the bright areas in the city scene (Fig. 7 in the paper) on average a fraction

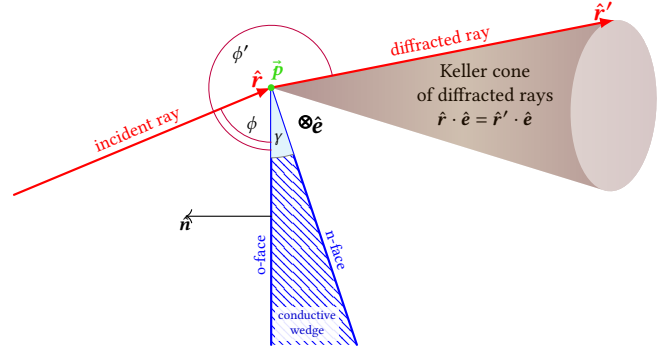


Fig. 4. UTD wedge diffraction.

brighter ($\ll 1$ db), but it is the diffraction lobes that make a significant difference (> 10 db) in the very dark areas.

For applications that require better accuracy, the inaccuracies induced by our decomposition into direct and diffracted terms could be avoided via fully-fledged beam tracing (sometimes done for acoustics propagation): our method is used to construct the free-space diffraction BSDF whenever a beam impinges upon geometry, but instead of using the clamped term Eqs. (S2.26) and (S2.27) (where we removed the 0th-order lobe), we construct the BSDF using the terms in Eqs. (S2.12) and (S2.19). However, beam tracing is expensive, and we try to keep the method practical, as this practicality is what enables simulations in unprecedentedly complex scenes.

S4 THE UNIFORM THEORY OF DIFFRACTION

We consider the uniform theory of diffraction (UTD), limited to diffracting conductive semi-infinite wedges. UTD is an asymptotic (high-frequency) method that employs ray tracing, and formalizes the diffraction coefficients of rays when they impinge upon a wedge edge.

Let the (infinite) edge of a wedge be oriented in direction \hat{e} , and let the wedge opening be of angle γ . It is common to designate the faces of the wedge as “o-face” and “n-face”. Let the normal of the o-face be \hat{n} . See Fig. 4.

Assume that a ray, propagating in direction \hat{r} , is incident upon the wedge edge at point \hat{p} , such that $\cos \beta = \hat{r} \cdot \hat{e}$ is the cosine between the ray and the wedge edge. Let \hat{r}' be a diffracted ray with $\cos \beta' = \hat{r}' \cdot \hat{e}$. Rays diffract into a 1-dimensional set of directions from \hat{p} where $\beta = \beta'$ (known as the “law of edge diffraction” [Keller 1962]), giving rise to the *Keller cone* of diffracted rays. The incident field is assumed to be a plane wave. Then, the wedge diffraction coefficient is given by the UTD formula [Pakyns 2016]:

$$f_{\text{UTD}}(\hat{r}, \hat{r}', z') \triangleq \mathcal{U}^{+, -} + \mathcal{U}^{-, -} + R_{s, h}(\mathcal{U}^{+, +} + \mathcal{U}^{-, +}) \quad (\text{S4.1})$$

with
$$\mathcal{U}^{\mathfrak{s}_1, \mathfrak{s}_2} \triangleq -\frac{e^{-i\pi/4}}{2n\sqrt{2\pi k}} \cot \frac{\pi + \mathfrak{s}_1(\phi + \mathfrak{s}_2\phi')}{2n} \times \mathcal{F} \left[k\alpha^{\mathfrak{s}_1} (\phi + \mathfrak{s}_2\phi') z' \sin^2 \beta \right],$$

where $\mathfrak{s}_{1,2} \in \{-, +\}$ denote signs, ϕ, ϕ' are the angles between the o-face and the incident and diffracted rays, respectively, z' is the distance from the diffraction point \hat{p} to the diffracted field observation

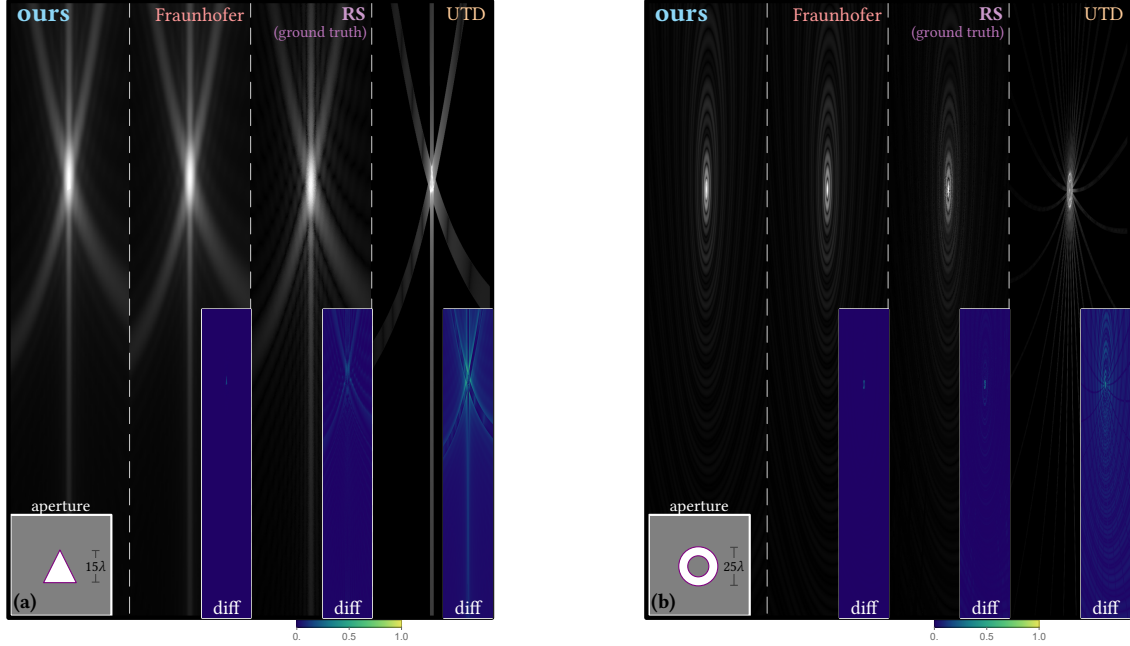


Fig. 5. **Comparison with UTD and numerically-integrated diffraction patterns.** Similar experiment to Fig. 6 in the paper. We compare with UTD, as well as Fraunhofer and the exact Rayleigh-Sommerfeld (RS) diffraction integrals. Note that UTD struggles with apertures that contain edges that are small, especially the ring aperture (b): The circles of the ring aperture are discretized into 32 small edges each. Instead of the expected circular diffraction pattern, edges diffract into 1-dimensional Keller cones, producing the visible line patterns, which are the intersections of each Keller cone with the wall.

point, and we define the shorthand $n \triangleq 2 - \gamma/\pi$. The *UTD reflection coefficients* $R_{s,h}$ depend on the wedge’s material and light’s polarization. For a conductive wedge $R_{s,h} = \pm 1$. k is the wavenumber, as before.

The utility functions \mathcal{F} and α^\pm are defined next. α^\pm relates the angular separation between the shadow boundary of the incident ray and the diffracted ray:

$$\alpha^\pm(\vartheta) \triangleq 2 \cos^2 \left[\pi n \times \text{round} \left(\frac{\vartheta \pm \pi}{2\pi n} \right) - \frac{\vartheta}{2} \right]. \quad (\text{S4.2})$$

The function \mathcal{F} is known as the *transition function*, for $x \geq 0$:

$$\mathcal{F}(x) \triangleq 2i\sqrt{x}e^{ix} \int_{\sqrt{x}}^{\infty} dv e^{-iv^2}. \quad (\text{S4.3})$$

Note that $\mathcal{F}(x)$ is defined for non-negative inputs. For negative x , the transition function becomes $\mathcal{F}(|x|)^\star$, where \star denotes complex conjugation. To evaluate \mathcal{F} on a computer, we note that it is related to the well-known Fresnel integrals, which can be expressed via the complex-valued error function. Therefore, we rewrite Eq. (S4.3) as

$$\mathcal{F}(x) = \sqrt{\frac{\pi}{2}} x (1+i) e^{ix} \left[1 - \text{erf}(\sqrt{ix}) \right]. \quad (\text{S4.4})$$

Evaluating \mathcal{F} then reduces to evaluating the complex error function, to which plenty of C libraries are available. The expression above applies to negative and non-negative x : the principal branches of the square roots are taken.

In all of our UTD results in this paper, we assume every edge is a semi-infinite half-plane, that is a wedge with $\gamma = 0$ (i.e. $n = 2$) where the o-face and n-face coincide. In the half-plane case $n = 2$,

we may slightly simplify the formulae above, specifically Eq. (S4.2) reduces to: $\alpha^\pm(\vartheta) = 2 \cos^2(\vartheta/2)$.

S4.1 Difficulties with UTD

Some practical difficulties arise with UTD:

- (1) Because UTD diffracts into a 1-dimensional set of directions (the Keller cone), the typical issues with BRDFs that scatter into a discrete or 1-dimensional sets of directions arise here as well: for example, next-event-estimation (NEE) is difficult, and rendering a diffraction pattern that arises on a diffuse surface (e.g., a wall) is akin to rendering caustics.
- (2) The diffraction coefficient function f_{UTD} in Eq. (S4.1) requires knowledge of the distance of propagation to the next interaction/observation point, i.e. a BSDF based on UTD can not be formulated as an angular scattering function.
- (3) The formulae above are limited to diffraction by a semi-infinite wedge. See Fig. 5b: when a curved edge (in this case, a circular aperture) is discretized into many small edges, the wedge-based UTD above fail to reproduce the expected circular diffraction pattern. While UTD can be applied to curved edges, this requires generalizing the formulae above to such curved edges (see Paknys [2016]). However, it is very common to model geometry via meshes, where curved edges do not explicitly arise, therefore computing the implied curvature of edges is cumbersome in practice.
- (4) Finally, as discussed in the paper, UTD’s most limiting difficulty is the requirement to employ mutually-interfering rays, prohibiting the formulation of a linear rendering equation.

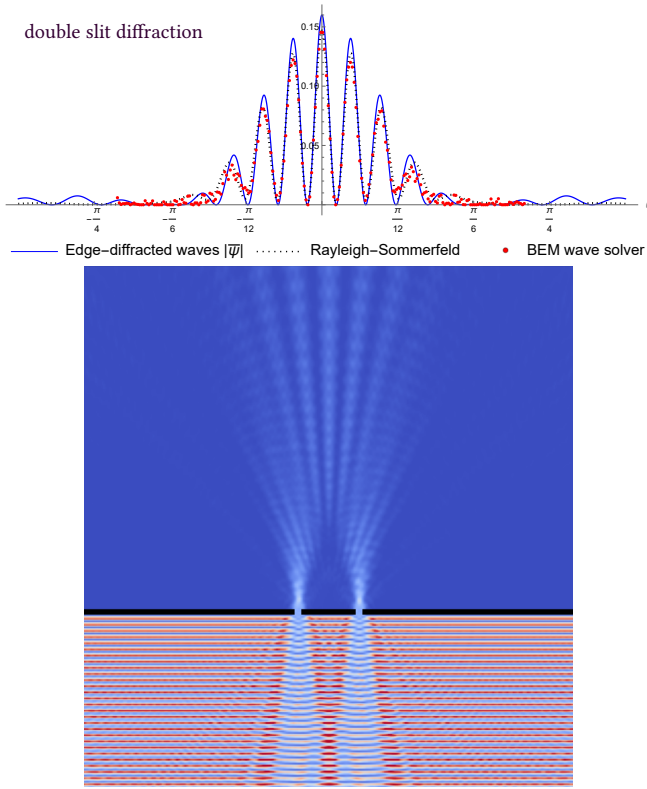


Fig. 6. **Comparison with a wave solver.** Additional comparison of our method (blue) with boundary element method (BEM) wave solver (discretized solution, red points) for a double-slit diffraction aperture. The exact Rayleigh-Sommerfeld diffraction integral (dashed black) is also plotted. (top) Visualization of the spatial field distribution of double-slit diffraction, with light incident from the bottom. We used the `bempp-cl` solver [Betcke and Scroggs 2021].

This has profound performance implications (see Fig. 5 in the paper). As rays may always annihilate each other, required ray count increases exponentially with path depth (hence, with scene complexity).

Our method does not suffer from any of the difficulties summarised above.

Fig. 5a, and Figs. 4 and 5 (in the paper), show that our method predicts similar diffraction fringes compared to UTD. Fraunhofer diffraction and UTD formalise very different theories, admitting distinct strengths and weaknesses. Because UTD rays carry phase, UTD predicts more accurate fringe positions when the observation points are not very far from the aperture (as in Fig. 4 in the paper). On the other hand, UTD assumes infinite edge lengths, hence Fraunhofer diffraction may predict more accurate results in the presence of short edges: note the horizontal diffraction streaks produced by our method in Fig. 5a, as well as Fig. 5 in the paper, which UTD fails to reproduce well. Extensive comparison between Fraunhofer diffraction and UTD is beyond the scope of this paper (see Yang and Brown [2011]). So is an analysis of path tracing in the

presence of mutually-interfering rays, which to our knowledge has never been done.

REFERENCES

- Timo Betcke and Matthew Scroggs. 2021. `Bempp-cl`: A fast Python based just-in-time compiling boundary element library. *Journal of Open Source Software* 6, 59 (March 2021), 2879. <https://doi.org/10.21105/joss.02879>
- Max Born and Emil Wolf. 1999. *Principles of optics : electromagnetic theory of propagation, interference and diffraction of light*. Cambridge University Press, Cambridge New York.
- Joseph B Keller. 1962. Geometrical theory of diffraction. *Josa* 52, 2 (1962), 116–130.
- U-Rae Kim, Wooyong Han, Dong-Won Jung, Jungil Lee, and Chaehyun Yu. 2021. Electrostatic potential of a uniformly charged triangle in barycentric coordinates. *Eur. J. Phys.* 42, 4 (July 2021), 045205.
- Leonard Mandel and Emil Wolf. 1995. *Optical coherence and quantum optics*. Cambridge University Press, Cambridge.
- Robert Paknys. 2016. Uniform theory of diffraction. In *Applied Frequency-Domain Electromagnetics*. John Wiley & Sons, Ltd, Chichester, UK, Chapter 8, 268–316.
- Amalia Torre. 2005. *Linear ray and wave optics in phase space: bridging ray and wave optics via the Wigner phase-space picture*. Elsevier.
- Ming Yang and Anthony K. Brown. 2011. Comparisons of UTD-Based and FK Models for Propagation Through Windows. *IEEE Antennas and Wireless Propagation Letters* 10 (2011), 1043–1046. <https://doi.org/10.1109/lawp.2011.2169230>

A DIFFRACTION LOBES AND THEIR SPACE-SPECTRUM BANDWIDTH

Due to the nature of our diffraction rendering algorithm, we would like to extract from the edge-diffracted field ψ_j the central (0^{th} -order) lobe, i.e., the non-diffracted energy. We recall that under wave optics light may not be perfectly resolved in a singular point or direction, due to the *uncertainty relation* between position and momentum [Mandel and Wolf 1995]:

$$\underbrace{\sigma_q \sigma_k}_{\text{space-spectrum bandwidth product}} \geq \frac{1}{2} \quad (\text{A.1})$$

(per dimension), where σ_q and σ_k are the standard deviations of light’s spatial position on the aperture and wavevector, respectively. The quantity $\sigma_q \sigma_k$ is often referred to as *space-spectrum bandwidth product*.

Noting that at the far field we have $k\sigma_\xi = \sigma_k$, where σ_ξ the standard deviation in light’s spatial position ξ on the observation plane, we immediately deduce that

$$\sigma_\xi \geq \frac{1}{2k} \frac{1}{\sigma_q}. \quad (\text{A.2})$$

Indeed, as wavelength becomes shorter or as the aperture’s size becomes greater (and hence the variance in light’s position on the aperture σ_q^2 increases), the variance in light’s position on the observation plane σ_ξ can be made smaller, and the diffraction lobes are sharper and less blurry.

We may now note an important Lemma.

LEMMA A.1 (INVARIANCE OF THE SPACE-SPECTRUM BANDWIDTH). *The space-spectrum bandwidth product is invariant under the transform Eq. (S2.22). That is, the space-spectrum bandwidth products of a edge-diffracted wave and its canonical edge-diffracted counterpart (related by Eq. (S2.22)) are equal.*

PROOF. Treating the spatial positions on the observation plane as random vectors, let

$$\sigma_{\xi}^2 \triangleq \left| \text{Var} \left(\vec{\xi} \right) \right| = \left| \mathbf{E} \left[\vec{\xi} \vec{\xi}^{\top} \right] \right| \quad \text{and} \quad \sigma_{\zeta}^2 \triangleq \left| \text{Var} \left(\vec{\zeta} \right) \right| = \left| \mathbf{E} \left[\vec{\zeta} \vec{\zeta}^{\top} \right] \right|,$$

i.e. the determinants of the covariance matrices, where $\vec{\zeta}$ is the position on the canonical edge-diffraction pattern, as before. Then, applying Eq. (S2.22),

$$\sigma_{\xi}^2 = \left| \mathbf{E} \left[\vec{\xi} \vec{\xi}^{\top} \right] \right| = \left| \mathbf{E} \left[\mathbf{\Xi}_j^{\top} \vec{\zeta} \vec{\zeta}^{\top} \mathbf{\Xi}_j \right] \right| = \left| \mathbf{\Xi}_j \right|^2 \sigma_{\zeta}^2 = \left(\frac{1}{ke_j} \right)^2 \sigma_{\zeta}^2. \quad (\text{A.3})$$

We turn our attention to the spatial variances on the aperture. The transform from the canonical edge to an arbitrary edge scales space isotropically by e_j . Therefore, the spatial variances on the aperture of an arbitrary edge and the canonical edge, denoted σ_q^2 and $(\sigma'_q)^2$, respectively, are related via

$$\left(\sigma'_q \right)^2 = e_j^{-2} \sigma_q^2. \quad (\text{A.4})$$

Then, using Eqs. (A.3) and (A.4), the space-spectrum bandwidth product becomes

$$\sigma_q \sigma_k = \sigma_q k \sigma_{\xi} = e_j \sigma'_q k \frac{1}{ke_j} \sigma_{\zeta} = \sigma'_q \sigma_{\zeta},$$

as required (recall that $k = 1$ in the canonical space). \square

B DERIVATION OF TOTAL POWER CONTAINED IN 0^{TH} -ORDER LOBE

Typical of Fraunhofer diffraction, the intensity of the pattern peaks at the centre $\vec{\xi} = 0$, which is the peak of the 0^{th} -order lobe. We assume that around that peak it falls off roughly as a Gaussian, viz.

$$\frac{|\psi(\vec{\xi})|^2}{|\psi(0)|^2} \propto e^{-\frac{1}{2} \vec{\xi}^{\top} \mathbf{\Sigma}_0^{-1} \vec{\xi}}. \quad (\text{B.1})$$

Applying the piecewise-linear approximation φ_{pl} , the amplitude of ψ at the peak is easy to compute in closed-form directly from the diffraction integral Eq. (S2.1):

$$\begin{aligned} \psi(0) &= \frac{k}{2\pi R} \int_{\mathcal{A}_+} d^2 \vec{q}^{\perp} \varphi(\vec{q}^{\perp}) = \frac{k}{2\pi R} \sum_t \int_{T_t} d^2 \vec{q}^{\perp} \varphi(\vec{q}^{\perp}) \\ &= \frac{k}{6\pi R} \sum_t |S_t| \sum_j \varphi_{t,j}, \end{aligned} \quad (\text{B.2})$$

where we rewrite the integral over the aperture as a sum of integrals over the triangles composing the aperture and integrate in barycentric coordinates. $\varphi_{t,j}$ are as defined in Eq. (S1.13).

The second-order moments of the 0^{th} -order peak is the second derivative w.r.t. $\vec{\xi}$ (the Hessian matrix), evaluated at 0:

$$\begin{aligned} \mathbf{\Sigma}_0^{-1} &= - \frac{1}{|\psi(0)|^2} \frac{d^2}{d\vec{\xi}^2} \left| \psi(\vec{\xi}) \right|^2 \Bigg|_{\vec{\xi}=0} \\ &= - 2 \frac{1}{|\psi(0)|^2} \left[\nabla \psi(\vec{\xi}) \left(\nabla \psi(\vec{\xi}) \right)^{\dagger} + \text{Re} \psi(\vec{\xi})^{\star} \frac{d^2}{d\vec{\xi}^2} \psi(\vec{\xi}) \right] \Bigg|_{\vec{\xi}=0} \\ &\approx - 2 \text{Re} \frac{1}{\psi(0)} \frac{d^2}{d\vec{\xi}^2} \psi(\vec{\xi}) \Bigg|_{\vec{\xi}=0}, \end{aligned} \quad (\text{B.3})$$

where we assume that $\nabla \psi \approx 0$ at $\vec{\xi} = 0$. That is, $\vec{\xi} = 0$ is the peak of the 0^{th} -order lobe with an even fall off around $\vec{\xi} = 0$, which is typically the case for Fraunhofer diffraction.

To compute the second-order derivative that arises in Eq. (B.3), we apply the Leibniz rule and, as usual, integrate over the triangular mesh in barycentric coordinates:

$$\begin{aligned} \frac{d^2}{d\vec{\xi}^2} \psi(\vec{\xi}) \Bigg|_{\vec{\xi}=0} &= - \frac{k^3}{2\pi R} \sum_t \int_{T_t} d^2 \vec{q}^{\perp} \left[\vec{q}^{\perp} (\vec{q}^{\perp})^{\top} \right] \varphi_{\text{pl}}(\vec{q}^{\perp}) \\ &= - \frac{k^3}{2\pi R} \sum_t 2|S_t| \int_0^1 d\gamma_1 \int_0^{1-\gamma_1} d\gamma_2 \left[\vec{q}^{\perp} (\vec{q}^{\perp})^{\top} \right] \sum_j \gamma_j \varphi_{t,j}, \end{aligned} \quad (\text{B.4})$$

where $\vec{q}^{\perp} (\vec{q}^{\perp})^{\top}$ is the outer product. The assignment $\vec{\xi} = 0$ is understood at the limit, and we formally interchange integration with the limit (the integrand is uniformly convergent). Consider the integration over a particular triangle T_t . To convert the outer product above to barycentric coordinates, recall the matrix \mathbf{B}_{T_t} (Eq. (S1.8)), and note that directly from its definition:

$$\begin{pmatrix} \vec{q}^{\perp} \\ 1 \end{pmatrix} = \mathbf{B}_{T_t} \begin{pmatrix} \gamma_1 \\ \gamma_2 \\ 1 - \gamma_1 - \gamma_2 \end{pmatrix} = \begin{pmatrix} \vec{l}_{t,2}\gamma_1 - \vec{l}_{t,1}\gamma_2 + \vec{u}_{t,3} \\ 1 \end{pmatrix}, \quad (\text{B.5})$$

where $\vec{u}_{t,j}$ are the vertices of the triangle T_t and $\vec{l}_{t,j}$ are the edge vectors, as defined in Eq. (S1.7). Therefore,

$$\vec{q}^{\perp} (\vec{q}^{\perp})^{\top} = \mathbf{H}_1 + \mathbf{H}_2\gamma_1 + \mathbf{H}_3\gamma_1^2 + \mathbf{H}_4\gamma_2 + \mathbf{H}_5\gamma_2^2 + \mathbf{H}_6\gamma_1\gamma_2, \quad (\text{B.6})$$

$$\text{where} \quad \mathbf{H}_1 \triangleq \vec{u}_{t,3} \vec{u}_{t,3}^{\top} \quad (\text{B.7})$$

$$\mathbf{H}_2 \triangleq - \left(\vec{u}_{t,3} \vec{l}_{t,2}^{\top} + \vec{l}_{t,2} \vec{u}_{t,3}^{\top} \right) \quad (\text{B.8})$$

$$\mathbf{H}_3 \triangleq \vec{l}_{t,2} \vec{l}_{t,2}^{\top} \quad (\text{B.9})$$

$$\mathbf{H}_4 \triangleq \vec{u}_{t,3} \vec{l}_{t,1}^{\top} + \vec{l}_{t,1} \vec{u}_{t,3}^{\top} \quad (\text{B.10})$$

$$\mathbf{H}_5 \triangleq \vec{l}_{t,1} \vec{l}_{t,1}^{\top} \quad (\text{B.11})$$

$$\mathbf{H}_6 \triangleq - \left(\vec{l}_{t,2} \vec{l}_{t,1}^{\top} + \vec{l}_{t,1} \vec{l}_{t,2}^{\top} \right). \quad (\text{B.12})$$

Note that all the expressions for the matrices \mathbf{H}_j are outer products. Finally, substitute Eqs. (B.4) and (B.6) into Eq. (B.3), and integrate:

$$\begin{aligned} \mathbf{\Sigma}_0^{-1} &= \text{Re} \frac{1}{\psi(0)} \frac{k^3}{\pi R} \sum_t \frac{|S_t|}{60} \left[60\mathbf{H}_1 + 20\mathbf{H}_2 + 10\mathbf{H}_3 + 20\mathbf{H}_4 \right. \\ &\quad \left. + 10\mathbf{H}_5 + \mathbf{H}_6 (2\varphi_{t,1} + 2\varphi_{t,2} + \varphi_{t,3}) \right], \end{aligned} \quad (\text{B.13})$$

yielding the final expression for the covariance $\mathbf{\Sigma}_0^{-1}$.

Total power in the 0^{th} -order lobe. We define the following expression,

$$\tilde{I}_0(\vec{\xi}) \triangleq \chi_{\sigma_{\xi}}(\vec{\xi}) e^{-\frac{1}{2} \vec{\xi}^{\top} \mathbf{\Sigma}_0^{-1} \vec{\xi}} |\psi(0)|^2 \quad (\text{B.14})$$

as our derived approximant for the intensity of the 0^{th} -order lobe. As discussed, $|\psi(0)|^2$ is the peak intensity, and the Gaussian term serves as a second-order approximation to the falloff of the 0^{th} -order lobe around the centre. This is a solid approximation for the intensity at a small region around the centre. $\chi_{\sigma_{\xi}}$, as defined in

Eq. (S2.23), quantifies the angular support of the 0th-order lobe, in an identical manner to Eq. (S2.26). σ_ξ^2 then is the angular variance, which we set to:

$$\sigma_\xi = \frac{\sigma_\gamma}{k e_{\text{avg}}}, \quad (\text{B.15})$$

where $\sigma_\gamma = \sqrt{3}$, as before, and e_{avg} is the average edge length (weighted by power contained in each edge) of all edges participating in the diffraction. σ_ξ then defines the average angular extent of the 0th-order lobe, and is most accurate when edge lengths do not vary vastly.

The total power contained in \tilde{I}_0 is then:

$$\begin{aligned} \tilde{P}_0^{(\text{PL})} &\triangleq \int_{\mathbb{R}^2} d^2 \vec{\xi} R^2 \tilde{I}_0(\vec{\xi}) = R^2 |\psi(0)|^2 \int_{\mathbb{R}^2} d^2 \vec{\xi} e^{-\frac{1}{2} \vec{\xi}^T (\boldsymbol{\Sigma}_0^{-1} + \sigma_\xi^{-2} \mathbf{I}) \vec{\xi}} \\ &= 2\pi \left| \boldsymbol{\Sigma}_0^{-1} + \sigma_\xi^{-2} \mathbf{I} \right|^{-1/2} R^2 |\psi(0)|^2. \end{aligned} \quad (\text{B.16})$$

Note that the R^2 term vanishes with its reciprocal in Eq. (B.2). To compute $\tilde{P}_0^{(\text{PL})}$, compute $\psi(0)$ using Eq. (B.2), $\boldsymbol{\Sigma}_0^{-1}$ using Eq. (B.3), σ_ξ using Eq. (B.15), and then evaluate Eq. (B.16).