

Diffraction

Abstract

The physics of diffraction.

Index Terms

Fraunhofer, Fresnel

A point source in the sky observed by a telescope will appear as a blob due to the diffraction of light. The diffraction limits the resolution of a telescope. In this post we will take a closer look at the physics and mathematics of the diffraction. Figure 1 illustrates self interfering light creating a pattern on the screen[1].

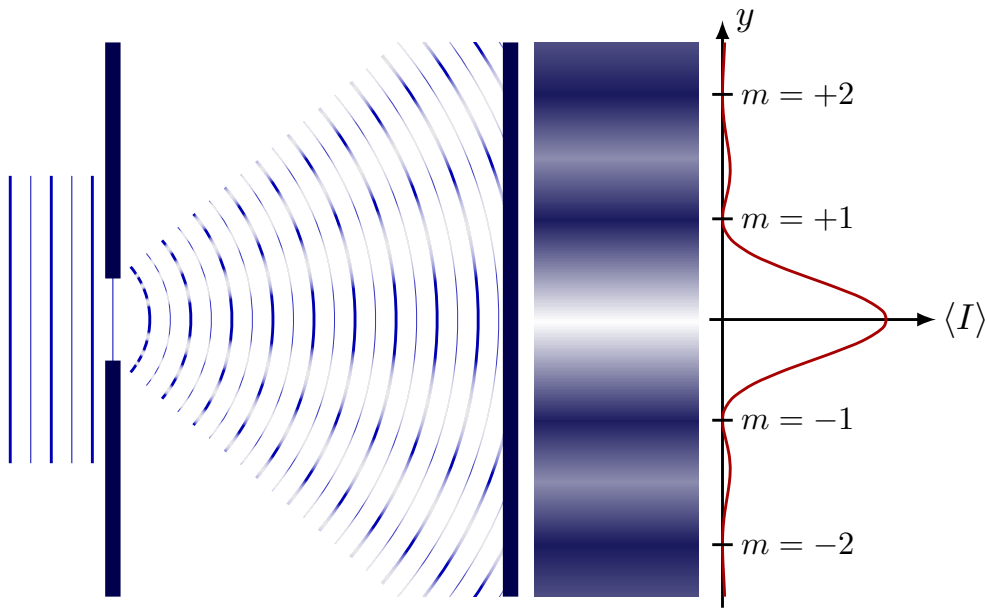


Fig. 1: Cross-section diagram of the aperture stop and the screen.

The proper mathematics to study the diffraction is pretty intense, and we are going to build gently towards the full treatment. Let us start with a light source in free space and understand the propagation of waves.

I. WAVE EQUATION

We will need the Maxwell's equations, which are shown below:

$$\text{Gauss' Law for electric fields: } \nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0}, \quad (1)$$

$$\text{Gauss' Law for magnetic fields: } \nabla \cdot \mathbf{B} = 0, \quad (2)$$

$$\text{Faraday's Law: } \nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}, \quad (3)$$

$$\text{Ampere's Law: } \nabla \times \mathbf{B} = \mu_0 \mathbf{J} + \mu_0 \epsilon_0 \frac{\partial \mathbf{E}}{\partial t}. \quad (4)$$

Consider the Maxwell's equations away from the sources ($\mathbf{J} = 0$ and $\rho = 0$). Lets take $\nabla \times$ of Eq. (3):

$$\nabla \times (\nabla \times \mathbf{E}) = -\frac{\partial}{\partial t} (\nabla \times B). \quad (5)$$

We need the following vector identity:

$$(\nabla \times [\nabla \times \mathbf{E}])^i = \epsilon^{ijk} \epsilon^{klm} \partial_j \partial_l E_m = (\delta^{il} \delta^{jm} - \delta^{im} \delta^{jl}) \partial_j \partial_l E_m = \nabla^i (\nabla \cdot \mathbf{E}) - \nabla^2 E^i. \quad (6)$$

Using the Ampere's law from Eq. (4) with $\mathbf{J} = 0$, we get

$$\left(\mu_0 \epsilon_0 \frac{\partial^2}{\partial t^2} - \nabla^2 \right) \mathbf{E} = 0. \quad (7)$$

This is a wave equation with speed $c = \frac{1}{\sqrt{\mu_0 \epsilon_0}}$, which is the speed of light. This shows that any change in the electric fields will travel at the speed of light. If we really wanted, we could do the identical derivation for magnetic fields.

We would like solve the wave equation:

$$[\nabla^2 - \partial_\tau^2] \mathbf{E}(\mathbf{r}, t) = 0, \quad (8)$$

where $\tau = ct$.

The spherical coordinates are illustrated in Fig. 2.

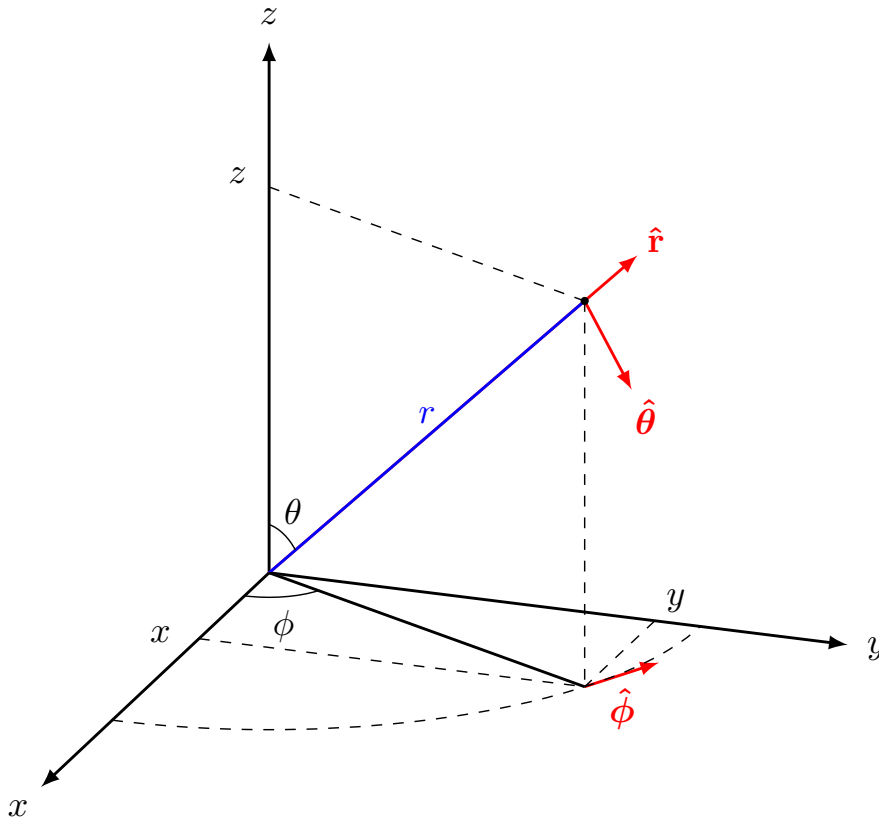


Fig. 2: The spherical coordinates.

In the spherical coordinates, the Laplace operator reads:

$$\nabla^2 = \left[\frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2}{\partial \phi^2} \right]. \quad (9)$$

However, the angle dependence drops out due to spherical symmetry.

$$[\nabla^2 - \partial_\tau^2] \mathbf{E}(\mathbf{r}, t) = \left[\frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) - \partial_\tau^2 \right] \mathbf{E}(\mathbf{r}, t) = \left[\frac{2}{r} \frac{\partial}{\partial r} + \frac{\partial^2}{\partial r^2} - \partial_\tau^2 \right] \mathbf{E}(\mathbf{r}, t) = \frac{\partial^2}{\partial r^2} [r\mathbf{E}(\mathbf{r}, t)] - \partial_\tau^2 [r\mathbf{E}(\mathbf{r}, t)] \quad (\mathbf{10})$$

where we used the Leibniz's formula:

$$\frac{d^m}{dx^m} [f(x)g(x)] = \sum_{k=0}^m \binom{m}{k} \frac{d^k f}{dx^k} \frac{d^{m-k} g}{dx^{m-k}}. \quad (11)$$

This is just like the one-dimensional wave equation, and we can find the solutions as incoming and outgoing waves:

$$(\partial_r + \partial_\tau) (\partial_r - \partial_\tau) [r\mathbf{E}(\mathbf{r}, t)]. \quad (12)$$

The outgoing wave solution is

$$\mathbf{E}(\mathbf{r}, t) = \frac{\mathcal{E}_0}{\lambda r} e^{-i(kr - \omega t + \delta)}, \quad (13)$$

where k is the wave-number, λ is the wavelength, ω is the angular frequency, \mathcal{E}_0 is the polarization, and δ is the initial phase. We defined the overall coefficient such that the expression actually represents the electric field surface density. The power radiated by the source is proportional to \mathbf{E}^2 and it behaves as $1/r^2$, which is the typical radius dependence.

II. HUYGENS-FRESNEL PRINCIPLE

The Huygens-Fresnel principle[2], states that every individual point on a wavefront can be treated as a point source. This principle is illustrated[3] in Fig. 3.

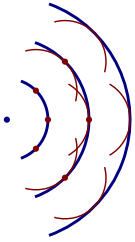


Fig. 3: An illustration of the Huygens' principle.

Let's assume that there is an aperture located at $z = 0$ and a screen at $z = z_0$. Per the Huygens-Fresnel principle, every differential area on the aperture can be considered as light source, and we just need to integrate over the aperture surface. We will assume that the distance to the screen is much larger than the aperture size. This enables us to be able to expand the radial distance from a point on the aperture \mathbf{r}' to a point on the screen \mathbf{r} :

$$|\mathbf{r} - \mathbf{r}'| = \sqrt{z_0^2 + (x - x')^2 + (y - y')^2} = z_0 \sqrt{1 + \frac{(x - x')^2}{z_0^2} + \frac{(y - y')^2}{z_0^2}} \simeq z_0 + \frac{(x - x')^2}{2z_0} + \frac{(y - y')^2}{2z_0} \quad (14)$$

Inserting this back in Eq. (13) yields:

$$\mathbf{E}(\mathbf{r}) \simeq \frac{e^{-ikz_0}}{\lambda z_0} \iint dx' dy' \mathcal{E}_0(x', y') e^{-ik \frac{(x-x')^2}{2z_0} - ik \frac{(y-y')^2}{2z_0}}. \quad (15)$$

If the aperture is a rectangle, and the source is far away, i.e., $\mathcal{E}_0(x', y') \simeq \text{constant}$ the integrals in Eq. (15) can be separated:

$$\mathcal{I} \equiv \int_a^b dx' e^{-ik \frac{(x'-x)^2}{2z_0}} = \sqrt{\frac{z_0}{2k}} \int_{\sqrt{\frac{k}{2z_0}(a-x)}}^{\sqrt{\frac{k}{2z_0}(b-x)}} du e^{-iu^2} = \sqrt{\frac{z_0}{2k}} \left(\int_0^{\sqrt{\frac{k}{2z_0}(b-x)}} e^{-iu^2} du - \int_0^{\sqrt{\frac{k}{2z_0}(a-x)}} e^{-iu^2} du \right) \quad (16)$$

Now we need to study these interesting integrals.

A. Fresnel Integrals

The Fresnel integrals are defined as follows:

$$\begin{aligned} S(t) &= \int_0^t dr \sin r^2, \\ C(t) &= \int_0^t dr \cos r^2. \end{aligned} \quad (17)$$

For a general value of t , the integrals need to be evaluated numerically. However, the asymptotic values $C(t)$ and $S(t)$ can be calculated via the closed contour integral below:

$$I = \oint_C dz e^{-z^2}, \quad (18)$$

where the contour C is illustrated in Fig. 4.

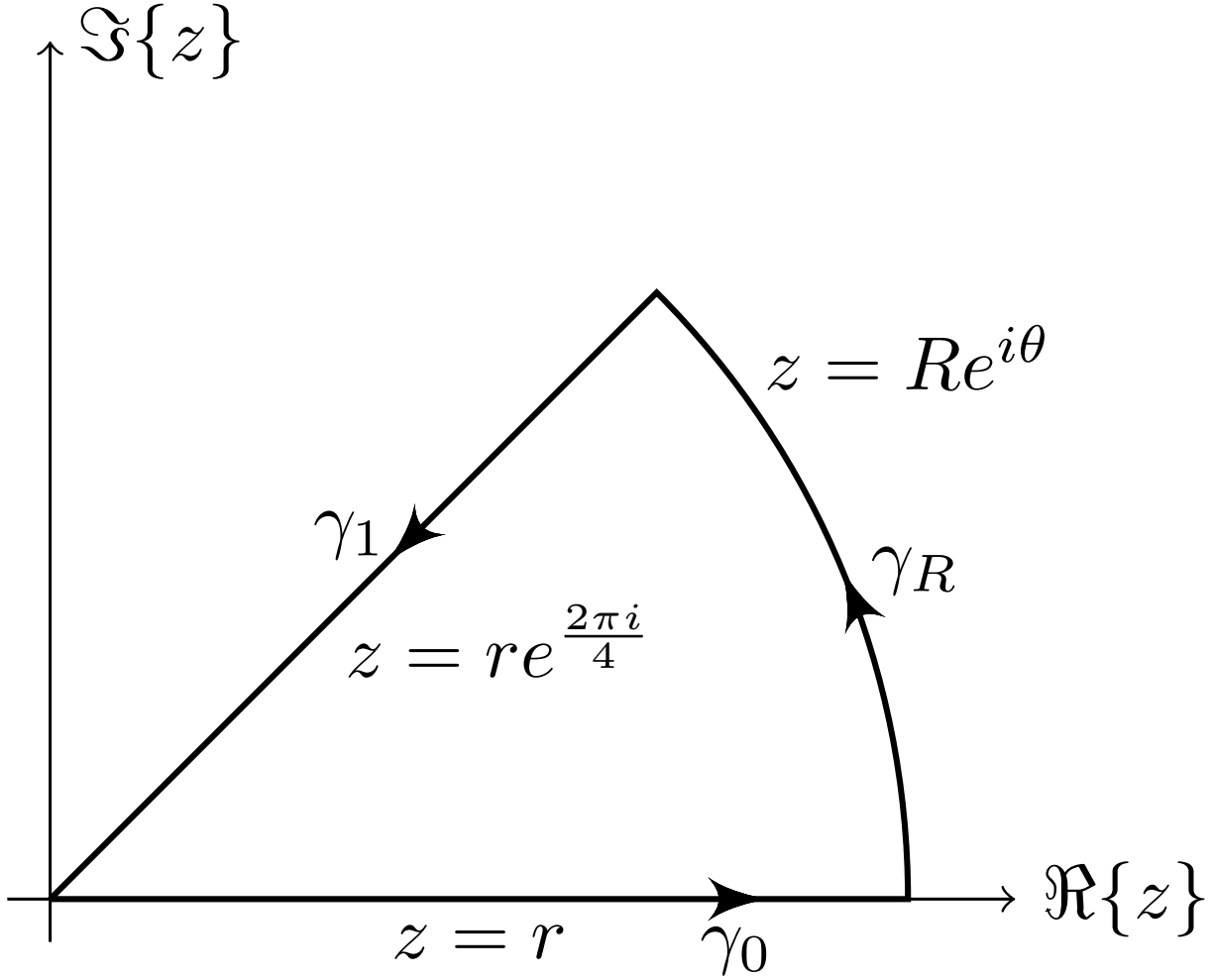


Fig. 4: The contour to evaluate the integral in Eq. [eqref:cont](#). The return path, γ_1 , is chosen such that the integrand reduces to the regular Gaussian.

Let's first evaluate the integral on γ_0 in the limit $R \rightarrow \infty$:

$$I_{\gamma_0} = \lim_{R \rightarrow \infty} \int_0^R dr e^{-r^2} = \frac{\sqrt{\pi}}{2}, \quad (19)$$

where the details of the derivation can be found [here](#). Now consider the (absolute value of the) integral on γ_R in the limit $R \rightarrow \infty$:

$$\begin{aligned} |I_{\gamma_R}| &= \left| \lim_{R \rightarrow \infty} R \int_0^{\frac{\pi}{4}} d\theta e^{i\theta} e^{-R^2(\cos^2 \theta - \sin^2 \theta + 2i \cos \theta \sin \theta)} \right| = \left| \lim_{R \rightarrow \infty} R \int_0^{\frac{\pi}{4}} d\theta e^{i\theta + i \sin(2\theta)} e^{-R^2 \cos(2\theta)} \right| \\ &\leq \left| \lim_{R \rightarrow \infty} R \int_0^{\frac{\pi}{4}} d\theta e^{-R^2 \cos(2\theta)} \right|. \end{aligned} \quad (20)$$

Let's try to put a bound on $\cos(2\theta)$ in the range $0 \leq \theta \leq \pi/4$. At $\cos(2\theta)|_{\theta=0} = 1$ and $\cos(2\theta)|_{\theta=\pi/4} = 0$. We can draw a line that connects these two points: $1 - \frac{4\theta}{\pi}$. Since $\frac{d^2}{d\theta^2} \cos(2\theta) = -4 \cos(2\theta) < 0$ for $0 < \theta < \pi/4$, we know that $\cos(2\theta) < 1 - \frac{4\theta}{\pi}$ in this range. This observation is illustrated in Fig. 5.

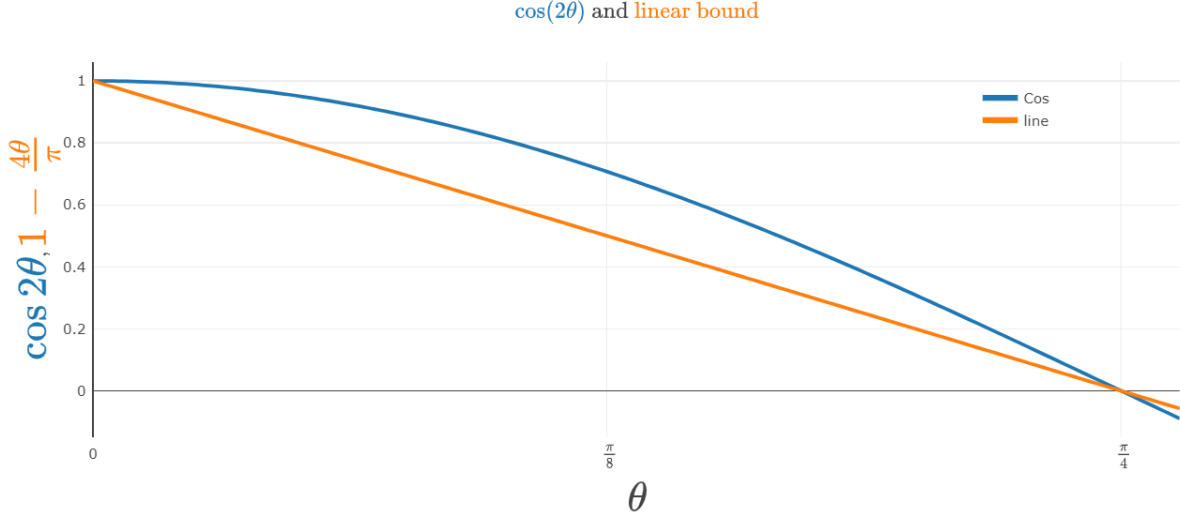


Fig. 5: $\cos(2\theta)$ and a bound on it with the line $1 - \frac{4\theta}{\pi}$.

We can now go back to Eq. (22) make use of the bound:

$$\begin{aligned}
 |I_{\gamma_R}| &\leq \left| \lim_{R \rightarrow \infty} R \int_0^{\frac{\pi}{4}} d\theta e^{-R^2 \cos(2\theta)} \right| \leq \left| \lim_{R \rightarrow \infty} R \int_0^{\frac{\pi}{4}} d\theta e^{-R^2(1 - \frac{4\theta}{\pi})} \right| = \left| \lim_{R \rightarrow \infty} R e^{-R^2} \int_0^{\frac{\pi}{4}} d\theta e^{R^2 \frac{4\theta}{\pi}} \right| \\
 &\leq \left| \lim_{R \rightarrow \infty} R \frac{\pi}{4R^2} (1 - e^{-R^2}) \right| = \left| \lim_{R \rightarrow \infty} \frac{\pi}{4R} (1 - e^{-R^2}) \right| = 0.
 \end{aligned} \tag{21}$$

Finally, let's look at the integral on γ_1 in the limit $R \rightarrow \infty$:

$$\begin{aligned}
 I_{\gamma_1} &= \lim_{R \rightarrow \infty} R \int_R^0 dr e^{\frac{i\pi}{4}} e^{-r^2 \frac{i\pi}{2}} = -\frac{1+i}{\sqrt{2}} \lim_{R \rightarrow \infty} \int_0^R dr (\cos r^2 - i \sin r^2) = -\frac{1+i}{\sqrt{2}} \int_0^\infty dr (\cos r^2 - i \sin r^2) \\
 &= -\frac{1}{\sqrt{2}} \left(\int_0^\infty dr \cos r^2 + \int_0^\infty dr \sin r^2 + i \left[\int_0^\infty dr \cos r^2 - \int_0^\infty dr \sin r^2 \right] \right).
 \end{aligned} \tag{22}$$

As we have computed individual pieces of the integral Eq.(18), we can assemble them and state that they need to add to 0 since e^{-z^2} is analytic everywhere. Therefore we have:

$$\begin{aligned}
 \oint_C dz e^{-z^2} = 0 &= I_{\gamma_0} + I_{\gamma_R} + I_{\gamma_1} \\
 &= \frac{\sqrt{\pi}}{2} + 0 - \frac{1}{\sqrt{2}} \left(\int_0^\infty dr \cos r^2 + \int_0^\infty dr \sin r^2 + i \left[\int_0^\infty dr \cos r^2 - \int_0^\infty dr \sin r^2 \right] \right)
 \end{aligned} \tag{23}$$

Matching the real and imaginary parts, we get:

$$\begin{aligned}
 \frac{1}{\sqrt{2}} \left(\int_0^\infty dr \cos r^2 + \int_0^\infty dr \sin r^2 \right) &= \frac{\sqrt{\pi}}{2}, \\
 \frac{1}{\sqrt{2}} \left(\int_0^\infty dr \cos r^2 - \int_0^\infty dr \sin r^2 \right) &= 0,
 \end{aligned} \tag{24}$$

from which we get

$$\int_0^\infty dr \cos r^2 = \int_0^\infty dr \sin r^2 = \frac{1}{2} \sqrt{\frac{\pi}{2}} \simeq 0.626. \tag{25}$$

Now we can conclude with the plots of the Fresnel integrals in Fig. 6.

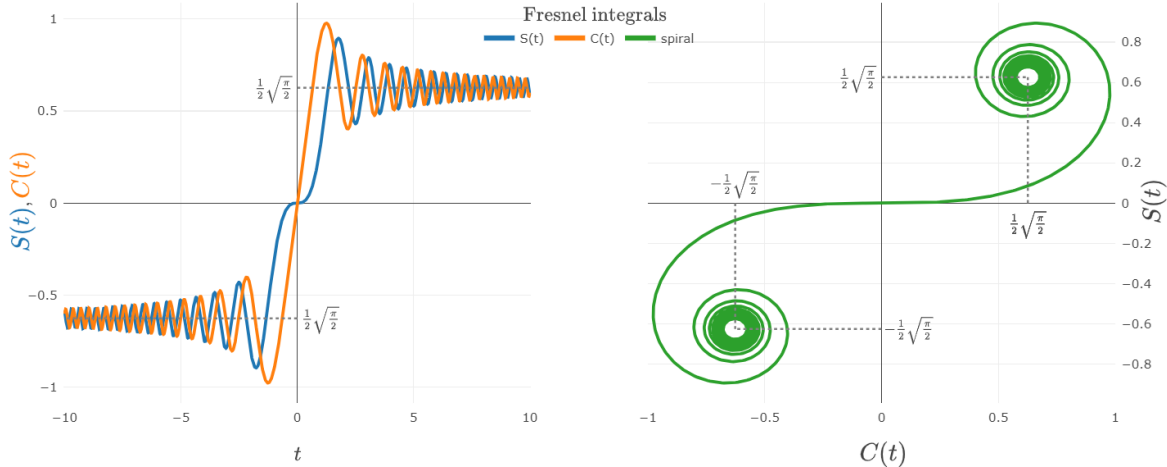


Fig. 6: Left: Fresnel integrals as a function of their argument, Right: parametric plot of the integrals forming the Euler spiral.

Going back to Eq. (16), we see that we can express it in terms of the Fresnel integrals:

$$\begin{aligned} \mathcal{I} &\equiv \int_a^b dx' e^{-ik \frac{(x'-x)^2}{2z_0}} = \sqrt{\frac{z_0}{2k}} \left(\int_0^{\sqrt{\frac{k}{2z_0}(b-x)}} e^{-iu^2} du - \int_0^{\sqrt{\frac{k}{2z_0}(a-x)}} du e^{-iu^2} \right) \\ &= \sqrt{\frac{z_0}{2k}} \{C(u) - C(v) + i[S(v) - S(u)]\}, \end{aligned} \quad (26)$$

where $u = \sqrt{\frac{k}{2z_0}(b-x)}$ and $v = \sqrt{\frac{k}{2z_0}(a-x)}$. Remember that the aperture is between a and b with a span $L = b - a$. We can shift the coordinate system so that the aperture is centered at 0 and it extends from $a = -L/2$ to $b = L/2$. Plugging this back in to Eq. (26) we get

$$\mathcal{I} = \sqrt{\frac{z_0}{2k}} \{C(x^+) - C(x^-) + i[S(x^-) - S(x^+)]\} \equiv \sqrt{\frac{z_0}{2k}} \mathcal{F}(x^-, x^+) \quad (27)$$

where $x_{\pm} = x \pm L/2$ and $\mathcal{F}(x^-, x^+) \equiv C(x^+) - C(x^-) + i[S(x^-) - S(x^+)]$. This is the x integral. We have an identical copy for y axis. Putting them together along with the overall coefficient in Eq. (15) we get:

$$\mathbf{E}(\mathbf{r}) \simeq \frac{\mathcal{E}_0 e^{-ikz_0}}{2} \mathcal{F}(x^-, x^+) \mathcal{F}(y^-, y^+), \quad (28)$$

where $y_{\pm} = y \pm L_y/2$.

III. INTERACTIVE PLOTS

We can now take E. (28) and plot the intensity in Fig. 7.

L_x (μm)

L_y (μm)

z (μm)

λ

plot range

Plot Type

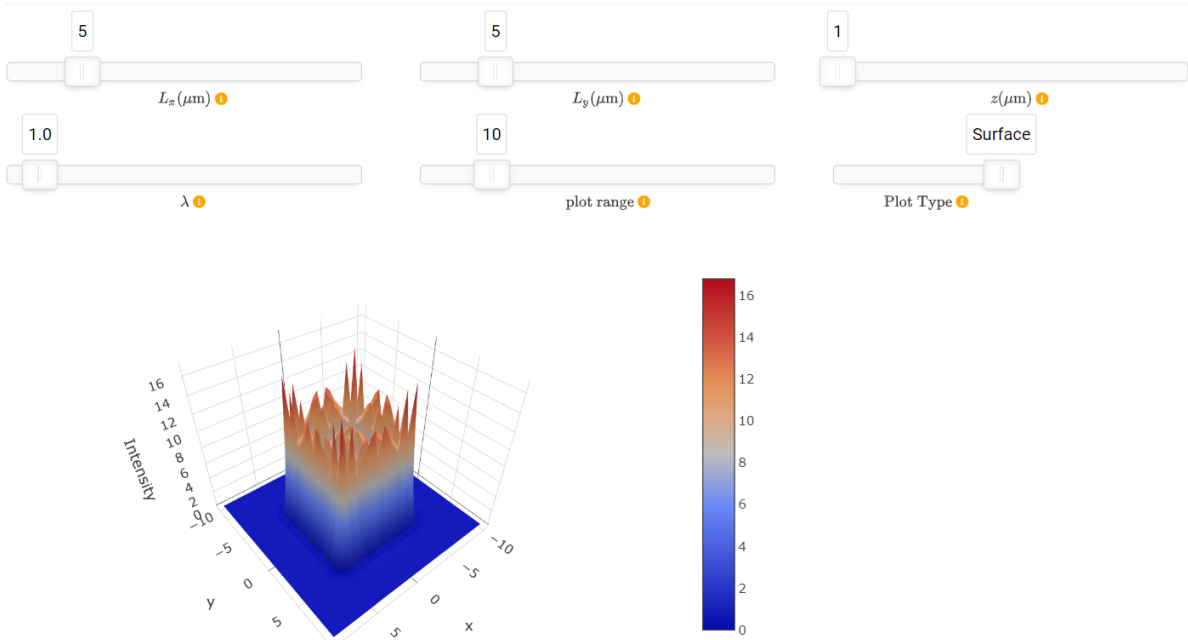


Fig. 7: The plot of intensity for various input parameters.

- [1] Tikz.net, “Diffraction.” https://tikz.net/optics_diffraction/, 2024.
- [2] Wikipedia, “Huygens–Fresnel principle — Wikipedia, the free encyclopedia.” <http://en.wikipedia.org/w/index.php?title=Huygens%E2%80%93Fresnel%20principle&oldid=1181770846>, 2024.
- [3] Tikz.net, “Huygens’ optics.” https://tikz.net/optics_huygens/, 2024.