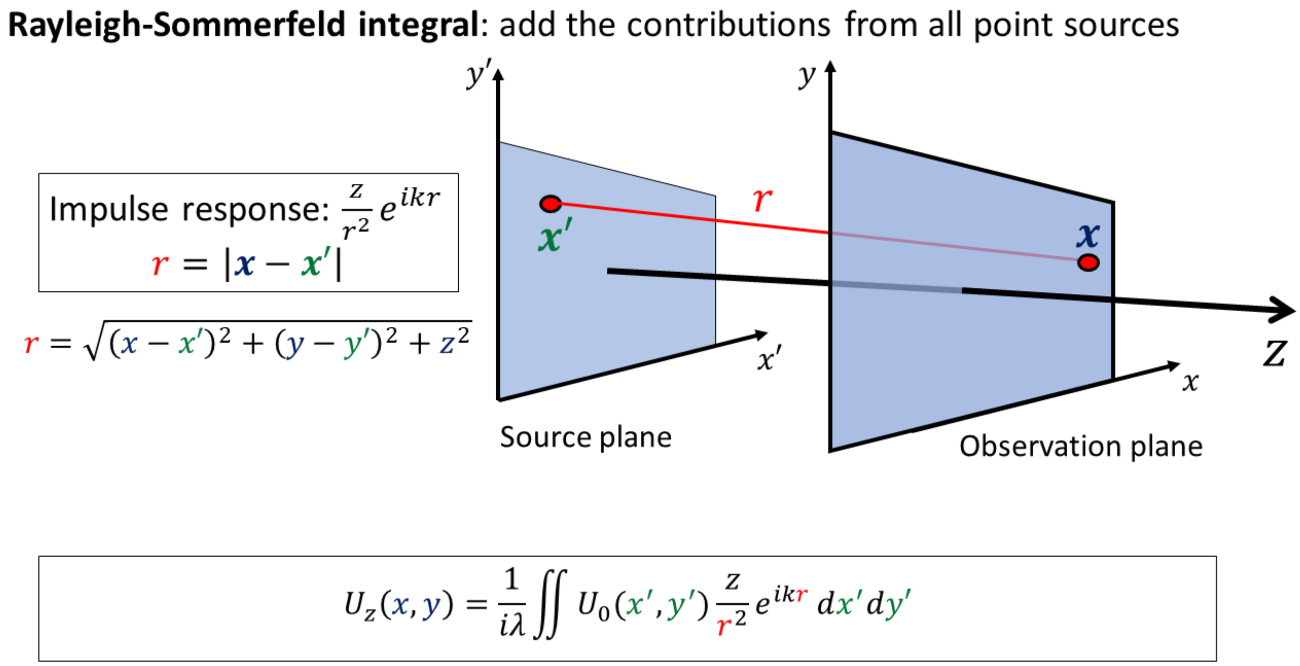

Figure 5.3: The Rayleigh-Sommerfeld states that to propagate a field from an initial plane $z = 0$ to another plane z, all the impulse response that originate from the initial plane must be summed together.

To find the commonly used Fresnel and Fraunhofer diffraction integrals, we find approximate expressions for r under the assumption that z is sufficiently large. We approximate r in the following way:

- The r in the denominator are approximated as $r = z$.
- The r in the exponent is approximated using a Taylor expansion:

$$
\begin{aligned}
r &= \sqrt{(x - x')^2 + (y - y')^2 + z^2} \\
&= z\sqrt{1 + \frac{(x - x')^2 + (y - y')^2}{z^2}} \\
&\approx z\left(1 + \frac{(x - x')^2 + (y - y')^2}{2z^2}\right) \\
&= z + \frac{(x - x')^2 + (y - y')^2}{2z} \quad \text{Fresnel approximation} \\
&\approx z + \frac{x^2 + y^2}{2z} - \frac{xx' + yy'}{z} \quad \text{Fraunhofer approximation}
\end{aligned}
\tag{5.5}
$$

  In the Fresnel approximation we approximate the spherical wave fronts as parabolic wave fronts, and in the Fraunhofer approximation we approximate them as plane waves for a small range of observation points (x, y), such that the paraxial approximation is valid (see e.g. Sections 3.1.8 and 3.2.1).

One might ask: why should we approximate the r in the exponent differently than the r in the denominator? The reason is that a small error in the exponent affects the value of the integrand much more than a small error in the denominator. For example, for $r \gg \lambda/2$, observe that $\frac{1}{r+\lambda/2} \approx \frac{1}{r}$, but $e^{ik(r+\lambda/2)} = -e^{ikr}$ regardless of how large r is compared to $\lambda/2$. Therefore, the r in the exponent must be approximated much more carefully than the r in the denominator.

If we apply the approximations of Eq. (5.5) to the Rayleigh-Sommerfeld integral of Eq. (5.3), we find the following expressions for the Fresnel and Fraunhofer integrals:

$$\begin{aligned} U_z(x,y) &\approx \frac{e^{ikz}}{i\lambda z} e^{ik\frac{x^2+y^2}{2z}} \iint U_0(x',y') e^{ik\frac{x'^2+y'^2}{2z}} e^{-ik\frac{xx'+yy'}{z}} \,\mathrm{d}x'\,\mathrm{d}y' \\ &= \frac{e^{ikz}}{i\lambda z} e^{ik\frac{x^2+y^2}{2z}} \mathcal{F}\left\{U_0(x',y') e^{ik\frac{x'^2+y'^2}{2z}}\right\}\left(\frac{x}{\lambda z},\frac{y}{\lambda z}\right) \quad \text{Fresnel propagation,} \end{aligned} \tag{5.6}$$

and

$$\begin{aligned} U_z(x,y) &\approx \frac{e^{ikz}}{i\lambda z} e^{ik\frac{x^2+y^2}{2z}} \iint U_0(x',y') e^{-ik\frac{xx'+yy'}{z}} \,\mathrm{d}x'\,\mathrm{d}y' \\ &= \frac{e^{ikz}}{i\lambda z} e^{ik\frac{x^2+y^2}{2z}} \mathcal{F}\left\{U_0(x',y')\right\}\left(\frac{x}{\lambda z},\frac{y}{\lambda z}\right) \quad \text{Fraunhofer propagation.} \end{aligned} \tag{5.7}$$

Here we used $k = 2\pi/\lambda$ to rewrite the integral as a Fourier transform. To remember these formulas more easily, let's introduce shorthand notation for the *quadratic phase factor*:

$$Q_z(x,y) = e^{ik\frac{x^2+y^2}{2z}}. \tag{5.8}$$

If we ignore the constant factor $\frac{e^{ikz}}{i\lambda z}$ (which is of little interest if we're interested in a monochromatic field in only a single plane z, because the factor doesn't depend on x and y), then we can remember the formulas for Fresnel propagation and Fraunhofer propagation as follows:

- **Fresnel propagation**: multiply the field with a quadratic phase factor, take the Fourier transform, multiply with another quadratic phase factor

$$\begin{aligned} U_z(x,y) &= e^{ik\frac{x^2+y^2}{2z}} \iint U_0(x',y') e^{ik\frac{x'^2+y'^2}{2z}} e^{-ik\frac{x'x+y'y}{z}} \,\mathrm{d}x'\,\mathrm{d}y' \\ &= Q_z(x,y)\mathcal{F}\left\{U_0(x',y')Q_z(x',y')\right\}\left(\frac{x}{\lambda z},\frac{y}{\lambda z}\right). \end{aligned} \tag{5.9}$$

- **Fraunhofer propagation**: take the Fourier transform of the field, multiply with a quadratic phase factor

$$\begin{aligned} U_z(x,y) &= e^{ik\frac{x^2+y^2}{2z}} \iint U_0(x',y') e^{-ik\frac{x'x+y'y}{z}} \,\mathrm{d}x'\,\mathrm{d}y' \\ &= Q_z(x,y)\mathcal{F}\left\{U_0(x',y')\right\}\left(\frac{x}{\lambda z},\frac{y}{\lambda z}\right). \end{aligned} \tag{5.10}$$

  If we only care about the intensity of the propagated field, then the quadratic phase factor $Q_z(x,y)$ becomes irrelevant, and we may say that the far field is given by the Fourier transform of the initial field (as was claimed in Section 3.5.3). Note that as the field propagates (z increases), the field simply expands without changing its shape, because the Fourier transform is evaluated in $(x/\lambda z, y/\lambda z)$.

So we now know how to derive the Fresnel and Fraunhofer approximations for field propagation. But when are these approximations valid? Recall that in Eq. (5.5), we approximated the r in the complex exponential e^{ikr}. Note that this complex exponential has a period of 2π. So if the error that we make when approximating r is ε_r, then we require that

$$k\varepsilon_r \ll 2\pi. \tag{5.11}$$

So how large is the error ε_r that we make when approximating r? In the case of the Fresnel approximation, that error can be found by taking the next term in the Taylor expansion of r. The Taylor expansion of $\sqrt{1+\delta}$ is

$$\sqrt{1+\delta} = 1 + \frac{\delta}{2} - \frac{\delta^2}{8} + \mathcal{O}(\delta^3). \tag{5.12}$$

In Eq. (5.5) we have $\delta = \frac{(x-x')^2+(y-y')^2}{z^2}$ and we expand to second order. Therefore, if we define $\rho^2 = (x-x')^2 + (y-y')^2$ the error term is

$$|\varepsilon_r| = z\delta^2/8 = \frac{\rho^4}{8z^3}. \tag{5.13}$$

When we plug this result in the requirement that we defined in Eq. (5.11), and using $k = 2\pi/\lambda$, we find the following requirement for the propagation distance z in order for the Fresnel approximation to be valid

$$\frac{z}{\lambda} \gg \frac{1}{2}\left(\frac{\rho}{\lambda}\right)^{4/3}. \tag{5.14}$$

Now let's find the condition for the Fraunhofer approximation to be valid. In Eq. (5.5), we see that to make the Fraunhofer approximation, the term $\frac{x'^2+y'^2}{2z}$ must be negligible. Let's assume that the field in the initial plane is restricted by an aperture with radius R, so that the maximum value of $x'^2 + y'^2$ is R^2. Then we require according to Eq. (5.11)

$$k\frac{R^2}{2z} \ll 2\pi \quad \Leftrightarrow \quad z \gg \frac{R^2}{2\lambda} \quad \Leftrightarrow \quad \frac{R^2}{\lambda z} \ll 2. \tag{5.15}$$

The dimensionless number $F = \frac{R^2}{\lambda z}$ is called the *Fresnel number*, and it is commonly used to indicate whether the Fraunhofer approximation is valid. It is usually understood that the condition for Fraunhofer propagation is $F \ll 1$. This condition is valid when the propagation distance and/or wavelength is very large compared to the diffracting aperture. Moreover, keep in mind that the Fraunhofer diffraction formula of Eq. (5.10) assumes that the observation points x and y are sufficiently close to the optical axis (z-axis) so that the paraxial approximation is valid (see Eq. (3.40)).

According to the Rayleigh-Sommerfeld integral, each field point emits a spherical wave e^{ikr}/r that is multiplied by an inclination factor z/r, which makes the wave propagate more in the forward direction, and less in the sideways direction. To find the Fresnel and Fraunhofer approximations, we approximate the r in the denominator as z, and we Taylor expand the r in the exponent. To propagate a field using the Fresnel approximation, we multiply it with a quadratic phase factor, and take the Fourier transform (assuming you don't care about the phase of the propagated field). To propagate a field using the Fraunhofer approximation, you simply take the Fourier transform of the field (again ignoring the phase of the propagated field). The Fraunhofer approximation is valid when the propagation distance is so large compared to the diffracting aperture, that the Fresnel number is much smaller than 1.

5.1.3 Partially coherent fields

In the previous sections we saw how we can computationally propagate a field using the Angular Spectrum / Rayleigh-Sommerfeld / Fresnel / Fraunhofer propagation formulas. However, all these methods assume we have perfectly coherent, monochromatic field, which can be described with a complex-valued function $U_0(x, y)$. But we saw in Section 3.6 that light is not always coherent. It can consist of multiple wavelengths (making the light temporally less coherent), or it can be quasi-monochromatic but originate from an extended incoherent source (making the light spatially less coherent). How do we propagate a field if it's partially coherent? In this section, we will see that a partially coherent field can be decomposed into coherent modes (that are mutually incoherent), so that we can straightforwardly apply the propagation formulas we found for coherent fields in the previous sections.

First, we must be able to describe a partially coherent field mathematically. We saw in Section 3.6 that the degree of coherence depends on the correlation between random field fluctuations at different points in space and time. In the context of partial polarization, we saw in Equation (4.28) that this correlation can be computed with the time average

$$\langle U(x_1, t+\tau)U(x_2, t)^*\rangle\,. \tag{5.16}$$

This time average denotes the correlation between the field fluctuations at two points x_1, x_2, and with a time delay τ. We assume the fluctuations are *wide-sense stationary*, which means that correlation only depends on the time *difference* τ, and not on the time t itself. Then, to fully describe a partially coherent field we would need a *5-dimensional function*, which defines the correlation between each pair of points (x_1, y_1), (x_2, y_2), for each time interval τ. This function is called the *mutual coherence function* Γ

$$\Gamma(x_1, y_1, x_2, y_2, \tau) = \langle U(x_1, y_1, t+\tau)U(x_2, y_2, t)^*\rangle\,. \tag{5.17}$$

In case the field is quasi-monochromatic (i.e. it has high temporal coherence, but not necessarily high spatial coherence), then the degree of correlation does not depend on the time difference τ, so we define the *4-dimensional mutual intensity function*

$$J(x_1, y_1, x_2, y_2) = \langle U(x_1, y_1, t)U(x_2, y_2, t)^*\rangle\,. \tag{5.18}$$