

8.5: The Huygens Principle

As the reader could see, the Born approximation is very convenient for tracing the basic features of (and the difference between) the phenomena of interference and diffraction. Unfortunately, this approximation, based on the relative weakness of the scattered wave, cannot be used for more typical experimental implementations of these phenomena, for example, Young's two-slit experiment, or of diffraction on a single slit or orifice – see, e.g. Fig. 10. Indeed, at such experiments, the orifice size a is typically much larger than the light's wavelength λ , and as a result, no clear decomposition of the fields to the “incident” and “scattered” waves is possible inside it.

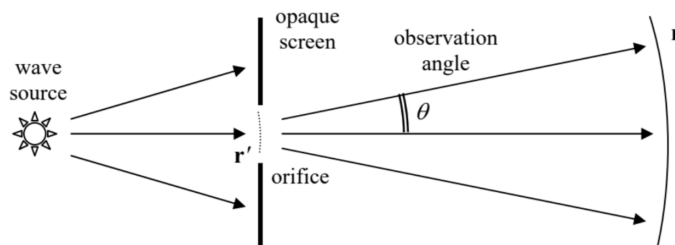


Fig. 8.10. Deriving the Huygens principle.

However, another approximation, called the Huygens (or “Huygens-Fresnel”) principle,³² is very instrumental for the description of such experiments. In this approach, the wave beyond the screen is represented as a linear superposition of spherical waves of the type (17), as if they were emitted by every point of the incident wave's front that has arrived at the orifice. This approximation is valid if the following strong conditions are satisfied:

$$\lambda \ll a \ll r, \quad (8.83)$$

where r is the distance of the observation point from the orifice. In addition, as we have seen in the last section, at small λ/a the diffraction phenomena are confined to angles $\theta \sim 1/ka \sim \lambda/a \ll 1$. For observation at such small angles, the mathematical expression of the Huygens principle, for the complex amplitude $f_\omega(\mathbf{r})$ of a monochromatic wave $f(\mathbf{r}, t) = \text{Re}[f_\omega e^{-i\omega t}]$, is given by the following simple formula

$$f_\omega(\mathbf{r}) = C \int_{\text{orifice}} f_\omega(\mathbf{r}') \frac{e^{ikR}}{R} d^2r'. \quad (8.84)$$

Here f is any transverse component of any of the wave's fields (either \mathbf{E} or \mathbf{H}),³³ R is the distance between point \mathbf{r}' at the orifice and the observation point \mathbf{r} (i.e. the magnitude of vector $\mathbf{R} \equiv \mathbf{r} - \mathbf{r}'$), and C is a complex constant.

Before describing the proof of Eq. (84), let me carry out its sanity check – which also will give us the constant C . Let us see what does the Huygens principle give for the case when the field under the integral is a plane wave with the complex amplitude $f_\omega(z)$, propagating along axis z , with an unlimited $x-y$ front, (i.e. when there is no opaque screen at all), so we should take the whole $[x, y]$ plane, say with $z' = 0$, as the integration area in Eq. (84) – see Fig. 11.

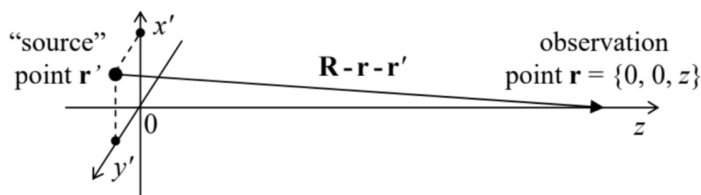


Fig. 8.11. Applying the Huygens principle to a plane incident wave.

Then, for the observation point with coordinates $x = 0$, $y = 0$, and $z > 0$, Eq. (84) yields

$$f_\omega(z) = C f_\omega(0) \int dx' \int dy' \frac{\exp\{ik(x'^2 + y'^2 + z^2)^{1/2}\}}{(x'^2 + y'^2 + z^2)^{1/2}}. \quad (8.85)$$

Before specifying the integration limits, let us consider the range $|x'|, |y'| \ll z$. In this range, the square root participating in Eq. (85) twice, may be approximated as

$$(x'^2 + y'^2 + z^2)^{1/2} \equiv z \left(1 + \frac{x'^2 + y'^2}{z^2} \right)^{1/2} \approx z \left(1 + \frac{x'^2 + y'^2}{2z^2} \right) \equiv z + \frac{x'^2 + y'^2}{2z}. \quad (8.86)$$

At $z \gg \lambda$, the denominator of Eq. (85) is a much slower function of x' and y' than the exponent in the nominator, and in the former case, it is sufficient (as we will check a posteriori) to keep just the main, first term of the expansion (86). With that, Eq. (85) becomes

$$f_\omega(z) = C f_\omega(0) \frac{e^{ikz}}{z} \int dx' \int dy' \exp \frac{ik(x'^2 + y'^2)}{2z} = C f_\omega(0) \frac{e^{ikz}}{z} I_x I_y, \quad (8.87)$$

where I_x and I_y are two similar integrals; for example,

$$I_x = \int \exp \frac{ikx'^2}{2z} dx' = \left(\frac{2z}{k} \right)^{1/2} \int \exp \{ i\xi^2 \} d\xi = \left(\frac{2z}{k} \right)^{1/2} \left[\int \cos(\xi^2) d\xi + i \int \sin(\xi^2) d\xi \right], \quad (8.88)$$

where $\xi \equiv (k/2z)^{1/2} x'$. These are the so-called Fresnel integrals. I will discuss them in more detail in the next section (see, in particular, Fig. 13), and for now, only one property³⁴ of these integrals is important for us: if taken in symmetric limits $[-\xi_0, +\xi_0]$, both of them rapidly converge to the same value, $(\pi/2)^{1/2}$, as soon as ξ_0 becomes much larger than 1. This means that even if we do not impose any exact limits on the integration area in Eq. (85), this integral converges to the value

$$f_\omega(z) = C f_\omega(0) \frac{e^{ikz}}{z} \left\{ \left(\frac{2z}{k} \right)^{1/2} \left[\left(\frac{\pi}{2} \right)^{1/2} + i \left(\frac{\pi}{2} \right)^{1/2} \right] \right\}^2 \equiv \left(C \frac{2\pi i}{k} \right) f_\omega(0) e^{ikz}, \quad (8.89)$$

due to contributions from the central area with a linear size corresponding to $\Delta\xi \sim 1$, i.e.

$$\Delta x \sim \Delta y \sim \left(\frac{z}{k} \right)^{1/2} \sim (\lambda z)^{1/2}, \quad (8.90)$$

so that the net contribution from the front points \mathbf{r}' well beyond the range (90) is negligible.³⁵ (Within our assumptions (83), which in particular require λ to be much less than z , the diffraction angle $\Delta x/z \sim \Delta y/z \sim (\lambda/z)^{1/2}$, corresponding to the important area of the front, is small.) According to Eq. (89), to sustain the unperturbed plane wave propagation, $f_\omega(z) = f_\omega(0) e^{ikz}$, the constant C has to be taken equal to $k/2\pi i$. Thus, the Huygens principle's prediction (84), in its final form, reads

$$f_\omega(\mathbf{r}) = \frac{k}{2\pi i} \int_{\text{orifice}} f_\omega(\mathbf{r}') \frac{e^{ikR}}{R} d^2r', \quad \text{Huygens principle} \quad (8.91)$$

and describes, in particular, the straight propagation of the plane wave (in a uniform medium).

Let me pause to emphasize how nontrivial this result is. It would be a natural corollary of Eqs. (25) (and the linear superposition principle) if all points of the orifice were filled with point scatterers that re-emit all the incident waves into spherical waves. However, as it follows from the above example, the Huygens principle also works if there is nothing in the orifice but the free space!

This is why let us discuss a proof of the principle,³⁶ based on Green's theorem (2.207). Let us apply it to the function $f = f_\omega$ where f_ω is the complex amplitude of a scalar component of one of the wave's fields, which satisfies the Helmholtz equation (7.204),

$$(\nabla^2 + k^2) f_\omega(\mathbf{r}) = 0, \quad (8.92)$$

and the function $g = G_\omega$ that is the temporal Fourier image of the corresponding Green's function. The latter function may be defined, as usual, as the solution to the same equation with the added delta-functional right-hand side with an arbitrary coefficient, for example,

$$(\nabla^2 + k^2) G_\omega(\mathbf{r}, \mathbf{r}') = -4\pi\delta(\mathbf{r} - \mathbf{r}'). \quad (8.93)$$

Using Eqs. (92) and (93) to express the Laplace operators of the functions f_ω and G_ω , we may rewrite Eq. (2.207) as

$$\int_V \{ f_\omega [-k^2 G_\omega(\mathbf{r}, \mathbf{r}') - 4\pi\delta(\mathbf{r} - \mathbf{r}')] - G_\omega(\mathbf{r}, \mathbf{r}') [-k^2 f_\omega] \} d^3r = \oint_S \left[f_\omega \frac{\partial G_\omega(\mathbf{r}, \mathbf{r}')}{\partial n} - G_\omega(\mathbf{r}, \mathbf{r}') \frac{\partial f_\omega}{\partial n} \right] d^2r, \quad (8.94)$$

where \mathbf{n} is the outward normal to the surface S limiting the integration volume V . Two terms in the left-hand side of this relation cancel, so that after swapping the arguments \mathbf{r} and \mathbf{r}' , we get

$$-4\pi f_\omega(\mathbf{r}) = \oint_S \left[f_\omega(\mathbf{r}') \frac{\partial G_\omega(\mathbf{r}', \mathbf{r})}{\partial n'} - g_\omega(\mathbf{r}', \mathbf{r}) \frac{\partial G_\omega(\mathbf{r}')}{\partial n'} \right] d^2 r'. \quad (8.95)$$

This relation is only correct if the selected volume V includes the point \mathbf{r} (otherwise we would not get its left-hand side from the integration of the delta function), but does not include the genuine source of the wave – otherwise, Eq. (92) would have a non-zero right-hand side. Now let \mathbf{r} be the field observation point, V be all the source-free half-space (for example, the half-space right of the screen in Fig. 10), so that S is the surface of the screen, including the orifice. Then the right-hand side of Eq. (95) describes the field (at the observation point \mathbf{r}) induced by the wave passing through the orifice points \mathbf{r}' . Since no waves are emitted by the opaque parts of the screen, we can limit the integration by the orifice area.³⁷ Assuming also that the opaque parts of the screen do not re-emit the waves “radiated” by the orifice, we can take the solution of Eq. (93) to be the retarded potential for the free space:³⁸

$$G_\omega(\mathbf{r}, \mathbf{r}') = \frac{e^{ikR}}{R}. \quad (8.96)$$

Plugging this expression into Eq. (82), we get

$$\text{Kirchhoff integral} \quad -4\pi f_\omega(\mathbf{r}) = \oint_{\text{orifice}} \left[f_\omega(\mathbf{r}') \frac{\partial}{\partial n'} \left(\frac{e^{ikR}}{R} \right) - \left(\frac{e^{ikR}}{R} \right) \frac{\partial f_\omega(\mathbf{r}')}{\partial n'} \right] d^2 r'. \quad (8.97)$$

This is the so-called Kirchhoff (or “Fresnel-Kirchhoff”) integral. (Again, with the integration extended over all boundaries of the volume V , this would be an exact mathematical result.) Now, let us make two additional approximations. The first of them stems from Eq. (83): at $ka \gg 1$, the wave’s spatial dependence in the orifice area may be represented as

$$f_\omega(\mathbf{r}') = (\text{a slow function of } \mathbf{r}') \times \exp\{i\mathbf{k}_0 \cdot \mathbf{r}'\}, \quad (8.98)$$

where “slow” means a function that changes on the scale of a rather than λ . If, also, $kR \gg 1$, then the differentiation in Eq. (97) may be, in both instances, limited to the rapidly changing exponents, giving

$$-4\pi f_\omega(\mathbf{r}) = \oint_{\text{orifice}} i(\mathbf{k} + \mathbf{k}_0) \cdot \mathbf{n}' \frac{e^{ikR}}{R} f(\mathbf{r}') d^2 r'. \quad (8.99)$$

Second, if all observation angles are small, we may take $\mathbf{k} \cdot \mathbf{n}' \approx \mathbf{k}_0 \cdot \mathbf{n}' \approx -k$. With that, Eq. (99) is reduced to the Huygens principle in its form (91).

It is clear that the principle immediately gives a very simple description of the interference of waves passing through two small holes in the screen. Indeed, if the holes’ sizes are negligible in comparison with the distance a between them (though still are much larger than the wavelength!), Eq. (91) yields

$$f_\omega(\mathbf{r}) = c_1 e^{ikR_1} + c_2 e^{ikR_2}, \quad \text{with } c_{1,2} \equiv k f_{1,2} A_{1,2} / 2\pi i R_{1,2}, \quad (8.100)$$

where $R_{1,2}$ are the distances between the holes and the observation point, and $A_{1,2}$ are the hole areas. For the wave intensity, Eq. (100) yields

$$\bar{S} \propto f_\omega f_\omega^* = |c_1|^2 + |c_2|^2 + 2|c_1||c_2| \cos[k(R_1 - R_2) + \varphi], \quad \text{where } \varphi \equiv \arg c_1 - \arg c_2. \quad (8.101)$$

The first two terms in the last expression clearly represent the intensities of the partial waves passed through each hole, while the last one is the result of their interference. The interference pattern’s contrast ratio

$$\mathcal{R} \equiv \frac{\bar{S}_{\max}}{\bar{S}_{\min}} = \left(\frac{|c_1| + |c_2|}{|c_1| - |c_2|} \right)^2, \quad (8.102)$$

is the largest (infinite) when both waves have equal amplitudes.

The analysis of the interference pattern is simple if the line connecting the holes is perpendicular to wave vector $\mathbf{k} \approx \mathbf{k}_0$ – see Fig. 6a. Selecting the coordinate axes as shown in that figure, and using for the distances $R_{1,2}$ the same expansion as in Eq. (86), for the interference term in Eq. (101) we get

$$\cos[k(R_1 - R_2) + \varphi] \approx \cos\left(\frac{kxa}{z} + \varphi\right). \quad (8.103)$$

This means that the term does not depend on y , i.e. the interference pattern in the plane of constant z is a set of straight, parallel strips, perpendicular to the vector \mathbf{a} , with the period given by Eq. (60), i.e. by the Bragg law.³⁹ This result is strictly valid only at

$y^2 \ll z^2$; it is straightforward to use the next term in the Taylor expansion (73) to show that farther on from the interference plane $y = 0$, the strips start to diverge.

Reference

³² Named after Christian Huygens (1629-1695) who had conjectured the wave nature of light (which remained controversial for more than a century, until T. Young's experiments), and Augustin-Jean Fresnel (1788-1827) who has developed a quantitative theory of diffraction, and in particular gave a mathematical formulation of the Huygens principle. (Note that Eq. (91), sufficient for the purposes of this course, is not its most general form.)

³³ The fact that the Huygens principle is valid for any field component should not too surprising. Due to our assumption $a \gg \lambda$, the real boundary conditions at the orifice edges are not important; what is only important that the screen that limits the orifice, is opaque. Because of this, the Huygens principle (84) is a part of the so-called scalar theory of diffraction. (In this course, I will not have time to discuss the more accurate, vector theory, of these effects – see, e.g., Chapter 11 of the monograph by M. Born and E. Wolf, cited at the end of Sec. 7.1.)

³⁴ See, e.g., MA Eq. (6.10).

³⁵ This result very is natural, because the function $\exp\{ikR\}$ oscillates fast with the change of \mathbf{r}' , so that the contributions from various front points are averaged out. Indeed, the only reason why the central part of plane $[x', y']$ gives a non-zero contribution (89) to $f_\omega(z)$ is that the phase exponents stop oscillating as $(x'^2 + y'^2)$ is reduced below $\sim z/k$ – see Eq. (86).

³⁶ This proof was given in 1882 by the same G. Kirchhoff whose circuit rules were discussed in Sec. 4.1 and 6.6.

³⁷ Actually, this is a nontrivial point of the proof. Indeed, it may be shown that the exact solution of Eq. (94) identically is equal to zero if $f(\mathbf{r}')$ and $\partial f(\mathbf{r}')/\partial n'$ vanish together at any part of the boundary, of non-zero area. A more careful analysis of this issue (which is the task of the formal vector theory of diffraction, which I will not have time to pursue) confirms the validity of the described intuition-based approach at $a \gg \lambda$.

³⁸ It follows, e.g., from Eq. (16) with a monochromatic source $q(t) = q_\omega \exp\{-i\omega t\}$, with the amplitude $q_\omega = 4\pi\varepsilon$ that fits the right-hand side of Eq. (93).

³⁹ The phase shift φ vanishes at the normal incidence of a plane wave on the holes. Note, however, that the spatial shift of the interference pattern following from Eq. (103), $\Delta x = -(z/ka)\varphi$, is extremely convenient for the experimental measurement of the phase shift between two waves, especially if it is induced by some factor (such as insertion of a transparent object into one of the interferometer's arms) that may be turned on/off at will.

This page titled [8.5: The Huygens Principle](#) is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by [Konstantin K. Likharev](#) via [source content](#) that was edited to the style and standards of the LibreTexts platform.