7

Diffraction

7.1. The Huygens-Fresnel Postulate

Diffraction has played an important role in the development of Optics and, more generally, of Physics. It should be remembered that the wave nature of light had been recognized only after Arago and Fresnel had made their experiments on diffraction by circular holes and disks.

The propagation of light had been experimentally analyzed well before any theoretical formulation (Maxwell's equations) or before the introduction of the necessary mathematical tools (vector analysis, Fourier development, . . .). The physicists working during the eighteenth century proved to be very imaginative and pragmatic when they described the phenomena that they were observing.

7.1.1. *Intuitive Approach of the Huygens-Fresnel Principle*

The Huygens-Fresnel principle is a typical example; the description of the propagation of light waves using wavelets is completely intuitive and is in full agreement with experimental results. Its validity is extended to many other domains and is not restricted to electromagnetic waves.

Diffraction is characteristic of any physical phenomenon leading to the propagation of waves, such as electromagnetic waves, acoustic waves, mechanical surface waves, De Broglie waves. . . . Diffraction occurs any time that some *obstacle* is interposed on the wave trajectory; its effects are especially noticeable when the cross section of a beam is limited by a diaphragm having dimensions that cannot be considered as infinite, that's to say large, as compared to the wavelength.

Chapter 7 has been reviewed by Dr. Ludovic Brassé from Teemphotonics.

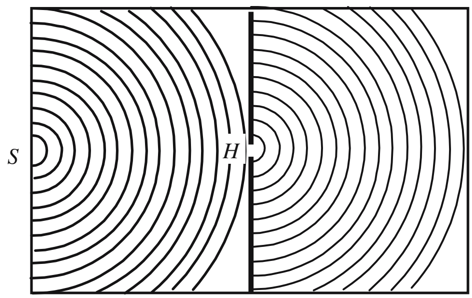


Figure 7.1. Diffraction of mechanical surface waves by a small aperture. The hole can be assimilated to a point source which, apart from a $\pi/2$ phase difference, has the same amplitude as the vibration produced, at point H, by the wave emitted by the source S.

To introduce the Huygens-Fresnel principle, we will start from the observation of the diffraction of mechanical surface waves propagating along the free interface of a liquid; the experimental arrangement is shown in Figure 7.1: a tank is divided into two parts by a wall inside of which a small hole H has been drilled. An electrically activated vibrator is disposed at point S, generating circular waves. All the walls are covered with a special foam to eliminate reflected waves. The two parts of the tank only communicate through the small hole H; the molecules of the liquid are set in motion by the wave coming from S. As far as the second half of the tank is concerned, everything occurs as if a vibrator is disposed at point H and generates circular waves.

According to the Huygens-Fresnel postulate, the vibrations of the molecules located at the hole are the same as the vibrations that would be created by the source S in the absence of the wall. In the arrangement of Figure 7.2, the hole is quite large and the effects of diffraction are only noticeable in the vicinity of the edges.

Diffraction and Interference

In most practical situations, the light that arrives at a given point has several origins: light coming directly from the source and light that has been diffracted by different obstacles. What is observed is the result of the interfer-

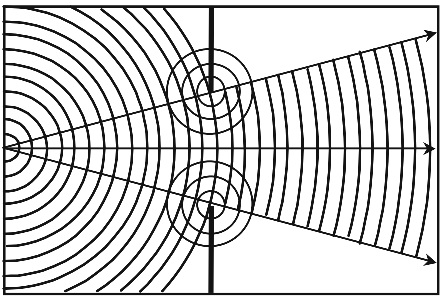


Figure 7.2. Diffraction of a circular wave by a large size obstacle. The effect of diffraction is only noticeable in the close vicinity of the edges of the diaphragm.

ence between all these radiations; if the initial source is coherent enough, fringes will be seen. The arrangement of Young's holes is an interesting example involving, simultaneously, diffraction and interference.

Qualitative Considerations on Diffraction Patterns

The basic properties of a planar wave are the following:

- Only one wave vector is involved.
- The wave surfaces have an infinite extension, since they are planar, with the mathematical meaning of this word.
- The complex amplitude of the vibrations remains constant and doesn't go to zero at infinity.

When the transverse extension of a planar wave is limited by some hole drilled in an opaque screen, the wave is said to be *stopped down.* The transmitted wave is no longer planar: the surface waves are not planar and the amplitude goes to zero on the border of the diffracted pattern. Such a wave (see Section 2.3.4) should be considered as the superposition of an infinity of elementary planar waves, having different wave vectors and different amplitudes. The amplitude of these planar waves is all the larger as the orientation of their wave vectors is closer to the wave vector of the incident planar wave. In Figure 7.3 has been drawn a curve, called an indicatrix, which is obtained by plotting, from the center of the diaphragm and in the direction of the wave vector, a segment having a length proportional to the intensity of the corresponding planar wave. All the diffracted wave vectors are practically inside the cones shown in the figure, the shape and size of the diaphragm determine their angle α. Practically $\alpha = K\lambda l/a$, where a is of the order of the transverse

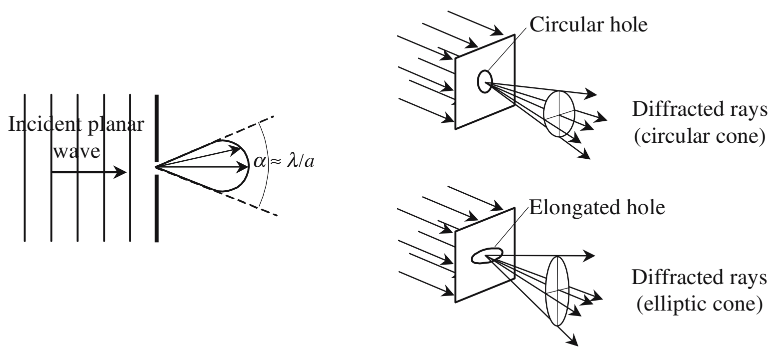


Figure 7.3. Diffraction of a planar wave by a hole. The smaller the hole, the wider the cone inside of which the diffracted light is spread. For an elongated hole, the diffraction pattern is more spread in a direction orthogonal to the larger size of the hole.

size of the diaphragm, and K is a proportionality coefficient of the order of 1. For a circular diaphragm having a diameter a, the coefficient K is 1.22. A small hole transmits only a little light, but the light is all the more spread the smaller it is. The indicatrix of a very small hole is a half-sphere.

Very often the diaphragm has the shape of a thin rectangular slot, the length of which can be considered as infinite compared to its width a. There is no diffraction in a direction parallel to the slot; it will be shown that in the other direction the angle of diffraction is $\alpha = \lambda/a$.

Huygens-Fresnel Wavelets

To solve a problem of electromagnetic propagation, we have to determine the electromagnetic vector at a given time and at all the points of the geometric space, knowing its repartition, over a given surface and at some previous time. The given surface can be the surface of the emitting source. The fields are supposed to be harmonic and can be expressed by the following formula:

$$A_{(x,y,z)} e^{-j\varphi(x,y,z)} e^{j\omega t}.$$

The source S of Figure 7.4 continuously produces harmonic waves with a period equal to T. In the same figure are shown, at time $t = 5T$, the six surface waves that have been emitted at the respective instants $t = 0$, T, $2T$, $3T$, $4T$, $5T$.

A first way of formulating the Huygens-Fresnel principle consists in the replacement of the source S by a set of point sources (called *auxiliary Huygens-Fresnel sources*), disposed over a wave surface (Σ) and emitting *in-phase* spherical waves (called *Huygens-Fresnel wavelets*) having the same amplitude that was produced at the same place as the source S. It is considered that the surface waves are the envelope of the wavelets emitted by all the auxiliary sources.

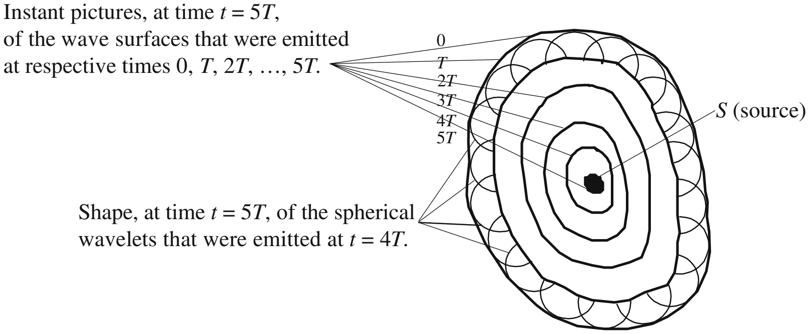


Figure 7.4. Illustration of the Huygens-Fresnel principle in the case where the auxiliary sources are distributed over a surface.

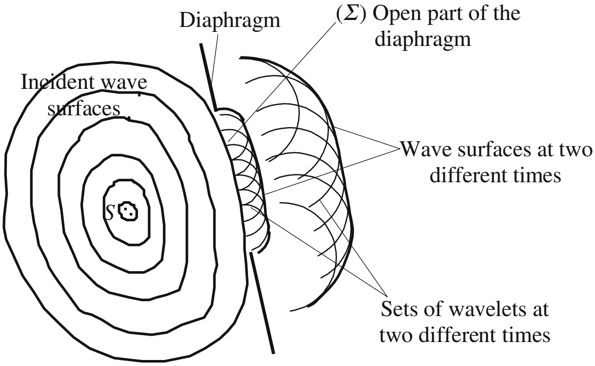


Figure 7.5. Interpretation of diffraction by diaphragm. To reconstitute the same field as produced by the source, it is necessary to let all the auxiliary sources interfere. If the diaphragm obliterates some of them, the reconstituted field is different as it was in the absence of a diaphragm.

As an example, we have drawn the wavelets that have been emitted at time $t = 4T$, their envelope, at time $t = 5T$, coincides with the wave surface that had been emitted at time $t = 0$.

Once the Huygens-Fresnel principle is admitted, the interpretation of diffraction is immediate: to reconstitute exactly a field identical to the one that is radiated by the source, we need to consider all the auxiliary sources; if a diaphragm obliterates some of them, the field is modified. The utilization of the wavelets for determining the diffraction pattern in the situation of Figure 7.5 is surely not rigorous, although commonly used.

7.1.2. *Huygens-Fresnel Mathematical Formulation*

Starting from the wave equation (2.3), Helmholtz and Kirchhoff gave a demonstration of the Huygens-Fresnel principle, which is then no longer a principle.

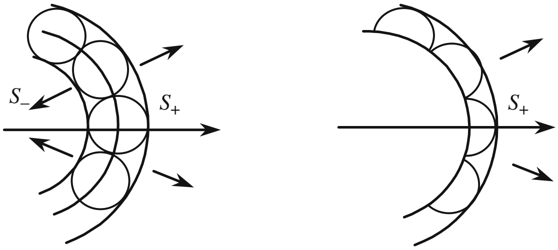


Figure 7.6. According to the Huygens-Fresnel principle, two waves S_+ and S_-, should be considered. Mathematicians could find good arguments to eliminate the wave, which has no experimental existence.

Their demonstration is valid in the case when the auxiliary sources are spread over a closed surface. Quite interesting from a theoretical point of view, this theory doesn't bring any additional physical features with regard to the intuitive approach; in addition, there are very few situations where rigorous calculations can be carried over. However, thanks to computers, numerical integration techniques have largely contributed to new developments of this aspect of electromagnetism.

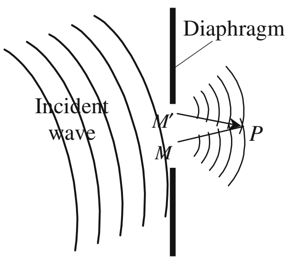
Amplitude of Vibration of the Auxiliary Sources

A second formulation of the Huygens-Fresnel principle is to place over the open part (Σ) of the diffracting diaphragm, auxiliary sources that have complex amplitudes proportional to the amplitudes $F(M)$ of the vibration at the same place, but in the absence of the diaphragm. These elementary amplitudes become meaningful only after an integration over the surface (Σ), they are proportional to the area $d\sigma$ of an elementary surface drawn around point M and will be written as $KF(M)\, d\sigma$, K is a proportionality coefficient that must be evaluated. The elementary field dE created at a point P of a spherical wavelet coming from point M is given by

$$dE = KF(M)\frac{e^{jkPM}}{MP}d\sigma.$$

A full development of the theory gives, for the proportionality coefficient, a value, $K = j/\lambda$, which can be interpreted in the following way: dE and $F(M)$ have the same dimension, thus K is homogeneous to the inverse of a length; the only length already met in the problem is the wavelength, which can be considered as a good reason for making $K = 1/\lambda$. The existence of the imaginary factor $j = e^{j\pi/2}$ is easily understood if we think of a spherical wave that diverges from the center of the wavelet and if we remember that there is a π phase shift for a spherical wave crossing its focus.

To obtain a mathematical expression for the Huygens-Fresnel principle, we will use the two formulas (7.1) and (7.2), given in Figure 7.7.



$F(M)$ = vibration created by the incident wave in M.
Elementary vibration sent from M to P:

$$dE = KF(M)\frac{e^{-jkPM}}{MP}d\sigma. \tag{7.1}$$

Global vibration in P:

$$E = \frac{j}{\lambda}\iint\limits_{\text{diaphragm}} F(M)\frac{e^{-jkPM}}{MP}d\sigma. \tag{7.2}$$

Figure 7.7. Interference of the wavelets emitted by the points located in the open part of the diaphragm.

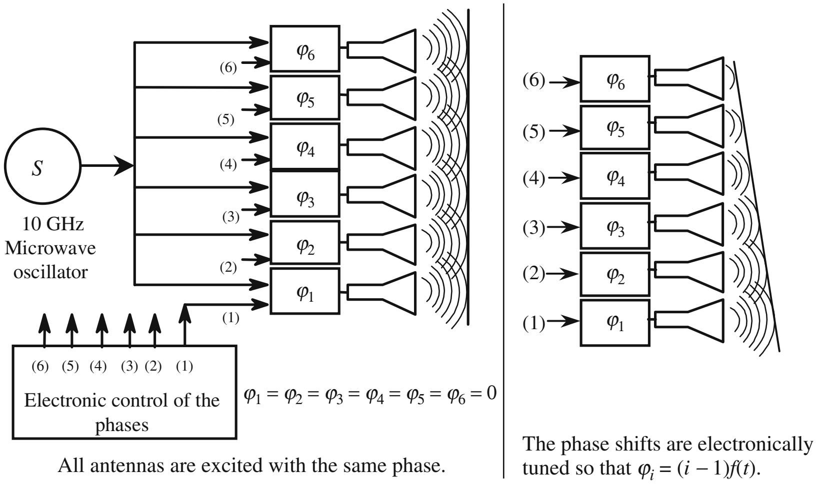


Figure 7.8. Concrete illustration of the Huygens-Fresnel principle: electronic angular scanning of the wave emitted by a microwave antenna network.

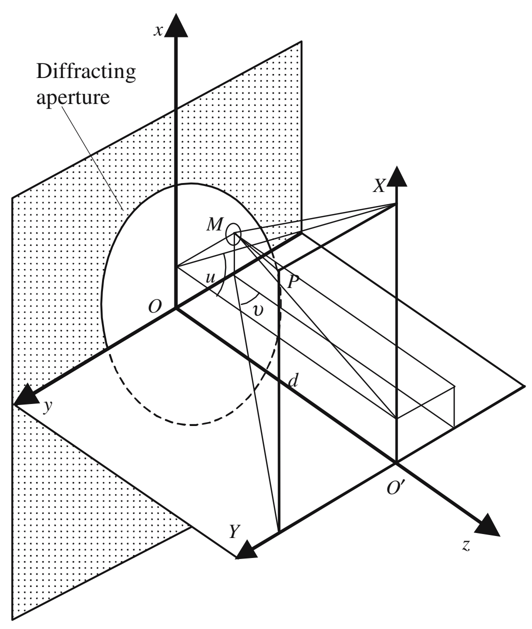
Figure 7.8 shows an arrangement often used in microwave to obtain an angular scanning of a radar electromagnetic wave. Antennas are regularly disposed along a plane; they are connected to a unique source through phase shifters, each of them shifting the phase of a specific amount that can be electrically tuned by an external signal.

If all the antennas are excited in phase, the superposition of all the waves that they emit makes a planar wave emitted parallel to the plane of the antennas. By a suitable choice of the phase shifts, wave surfaces of any shape can be synthesized. If the phase shifts vary according to an arithmetic progression, a planar wave is generated obliquely; if the step of the progression is varied versus time, an angular scanning is achieved.

7.1.3. *Resolution of a Problem of Diffraction*

7.1.3.1. *Notations*

Figure 7.9 summarizes any diffraction experimental arrangement: an incident wave illuminates an aperture opened in a diaphragm, which may have any shape but most often is a plane (xOy). The diffraction pattern is observed on a plane $(XO'Y)$ that is parallel to (xOy). Usual experimental conditions are such that, on one hand, the distance d between the two planes is far larger than the wavelength and that, on the other hand, the line MP joining any point of the diffracting diaphragm to any point of the plane observation makes a small angle with Oz; to be more precise, the two angles, u and v, of Oz with the projection of OM, respectively, on the plane xOz (or XOz) or on the plane yOz (or $YO'z$) are small enough to be assimilated to their sine or tangent.



$$E_{(X,Y)} = \frac{j}{\lambda} \iint_{\text{diaphragm}} f_{(x,y)} \frac{e^{-jkPM}}{MP} dx\,dy,$$

$$E_{(X,Y)} \approx \frac{j}{\lambda d} \iint_{\text{diaphragm}} f_{(x,y)} e^{-jkPM} dx\,dy,$$

$$E_{(X,Y)} \approx \frac{j}{\lambda d} \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f_{(x,y)} e^{-jkPM} dx\,dy. \qquad (7.3)$$

Figure 7.9. Notations used to study the diffraction by an aperture centered on the point O. $M(x, y)$ is a point inside the diffracting aperture. $P(X, Y)$ is the point where the diffracted field $F(X, Y)$ is observed. MP is assimilated to the distance $d = OO'$.

7.1.3.2. *Near Field Diffraction (Fresnel), Far Field Diffraction (Fraunhofer)*

Near field diffraction, preferably called *Fresnel diffraction*, is the most general and the most difficult problem of diffraction. The distance between the diffracting aperture and the point of observation (distance d between the two planes xOy and $XO'Y$ of Figure 7.9), is of course far larger than the wavelength, but can be comparable to the size of the aperture.

Far field diffraction is preferably called *Fraunhofer diffraction*. Let us suppose that the plane $XO'Y$ is covered by a screen on which the diffraction pattern is observed and let us see what happens when the distance d is increased: the diffraction pattern becomes larger and larger and less contrasted, but it tends to keep, whatever the distance, the same aspect that is called the *far field diffraction pattern*.

In fact, it's not necessary to go as far as infinity to observe a far field pattern, since infinity can be brought to the focal plane of a lens. Joseph von Fraunhofer was the first to propose this method of studying diffraction, it's probably the reason why the following epitaph has been written on his tombstone: "Approximavit sidera," "He made the stars closer."

7.1.3.3. *Approximate Expressions for the Diffraction Integrals*

Since the angles u and v are small, formula (7.3) can be simplified thanks to a series of developments in the arguments of the complex exponential, the method is very similar to what we did in Section 2.6,

$$MP = \sqrt{d^2 + (X-x)^2 + (Y-y)^2} \approx d\left[1 + \frac{(X-x)^2}{2d^2} + \frac{(Y-y)^2}{2d^2}\right].$$

The presence of the square root of polynomials inside the argument of an exponential is never pleasant, the development considerably simplifies the calculation

$$F(X,Y) = \frac{j}{\lambda d}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y) e^{-j\frac{k}{2d}\left[(X-x)^2+(Y-y)^2\right]}\, dx\, dy. \tag{7.4}$$

Let e_x, e_y, and e_z, respectively, be the three projections of a unit $\boldsymbol{e}$ vector parallel to MP, on the three axes Ox, Oy, and Oz we can write

$$e_x = \sin u \approx u << 1, \quad e_y = \sin v \approx v << 1,$$
$$e_z^2 = 1 - e_x^2 - e_y^2 \cong 1 - u^2 - v^2 \cong 1 \quad \forall (u,v).$$

Because of the assumption of small angles, the orientation of MP involves only the two parameters u and v. We now look at what happens when $d \to \infty$, the following asymptotic formulas can be written:

$$\frac{X}{d} \approx u, \quad \frac{Y}{d} \approx v, \quad \frac{x^2}{d} \to 0, \quad \frac{y^2}{d} \to 0,$$
$$\frac{(X-x)^2 + (Y-y)^2}{d} \approx \frac{X^2+Y^2}{d} - 2(ux+vy). \tag{7.5}$$

Formula (7.4) then becomes

$$F(X,Y) = \frac{j}{\lambda d} e^{-jkd} e^{-jk\frac{X^2+Y^2}{d}} \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y) e^{-jk(ux+vy)}\, dx\, dy, \tag{7.6}$$

$$F(u,v) = K \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y) e^{-jk(ux+vy)}\, dx\, dy, \tag{7.7}$$

$$F(k_x,k_y) = K \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y) e^{-j(k_x x+k_y y)}\, dx\, dy, \tag{7.8}$$

$$K = \frac{j}{\lambda d} e^{-jkd} e^{-jk\left(X^2+Y^2\right)/2d},$$

where k_x and k_y are the two projections on the axes OX and OY of a wave vector parallel to the unit vector $\boldsymbol{e}$.

The modulus of the complex number K is easily obtained and is equal to $1/\lambda d$; on the other hand, the determination of this is more difficult and would need more information about the asymptotic behavior of $(X^2 + Y^2)/2d$. In fact we don't really care about the phase, since expressions like (7.4) and (7.6) are mostly designated to evaluate the intensity of the diffracted light: in the end, $F(X, Y)$ will be multiplied by its complex conjugate and the phase will disappear.

Relationship Between Diffraction and Fourier Transform

Formula (7.8) is very interesting for two reasons:

- It is the decomposition of the diffracted field in planar harmonic waves.
- $F(k_x, k_y)$ is the two-dimensional Fourier transform of $f(x, y)$.

A Fourier transform makes a correspondence between a function $f(x, y)$ defined in the Ox, Oy geometric space and a function $F(k_x, k_y)$ defined in the space of the wave vectors. A planar wave of vector (k_x, k_y) is focused by a lens at a focal point ϕ_k; let ξ and η be the coordinates of ϕ_k (see Figure 7.10), the repartition $F(\xi, \eta)$ of the complex amplitudes of the light vibration over the plane $(F\xi, F\eta)$ is the Fourier transform of $f(x, y)$.

7.2. Fraunhofer Diffraction

7.2.1. *Definition and Conditions of Observation*

We refer to Figure 7.10; because of the Malus theorem, the stigmatism of the lens implies that it takes the same time to cover the optical paths $MI\phi_k$, $H'I'\phi_k$, $H''I''\phi_{kk}$, and $H_0I_0\phi_k$: when arriving at the point ϕ_k the rays emitted by the

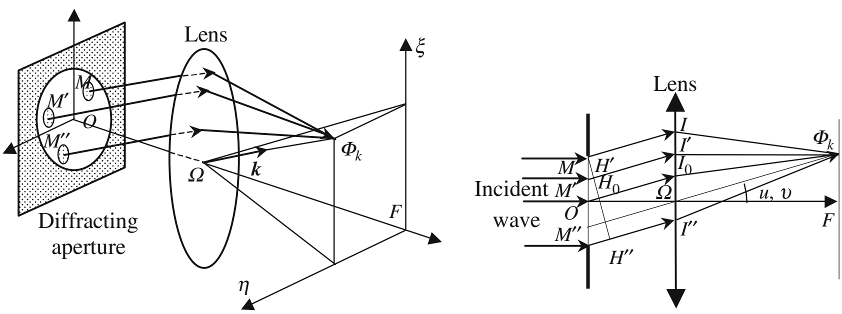


Figure 7.10. All the rays diffracted by the aperture in the direction (u, v) are focused at the same focal point ϕ_{uv} (coordinates $\xi = fu$, $\eta = fv$).

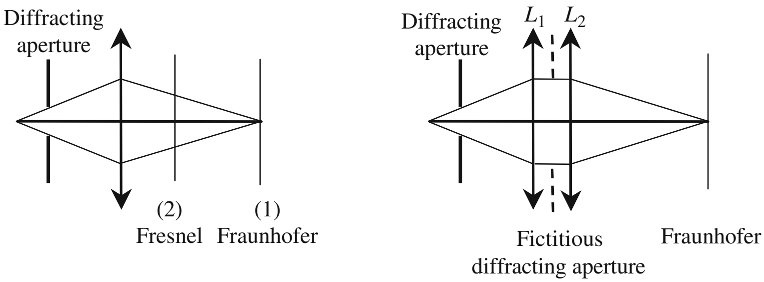


Figure 7.11. If the observation screen is in (2), we have Fresnel diffraction, while (1) corresponds to Fraunhofer diffraction. The lens of the left-hand figure (focal length f) is replaced by two separate lenses (focal lengths f_1 and f_2, such that $1/f_1 + 1/f_2 = 1/f$). The right-hand figure can then be considered as a far field diffraction arrangement through the fictitious diffracting aperture.

auxiliary sources M, M', O, and M'' have exactly the same phase differences that they had at points M, H', H_0, and M''.

We suppose that the incident wave is a planar wave propagating parallel to Oz; the phase difference φ_M between the two planar waves, respectively, is diffracted by the points O and M in the direction defined by the angles u and v,

$$\varphi_M = kOM = k(ux + vy).$$

Superposing, at the focal point ϕ_{uv}, all the planar waves diffracted in the (u, v) direction by the different points of the diffracting aperture gives the same result as expressed by formula (7.7).

Diffraction in the Vicinity of a Geometric Image

The case of Fraunhofer diffraction is more general than far field diffraction. We will say that we are in the situation of Fraunhofer diffraction every time that the diffraction pattern is observed in the vicinity of the image of an object through some optical imaging device. It is shown in Figure 7.10 how a Fraunhofer diffraction problem can be studied as a far field diffraction problem.

7.2.2. *Diffraction by One or Several Slits*

7.2.2.1. *One Slit*

Direct Determination of the Diffraction Pattern

The problem is to calculate the diffraction pattern of a rectangular aperture having a length L that is considerably larger than its width a. The notations and experimental conditions are described in Figure 7.12. Since the incident wave propagates orthogonal to the plane of the slit, all the auxiliary sources vibrate with the same phase, let b_0 be the amplitude of their vibrations.

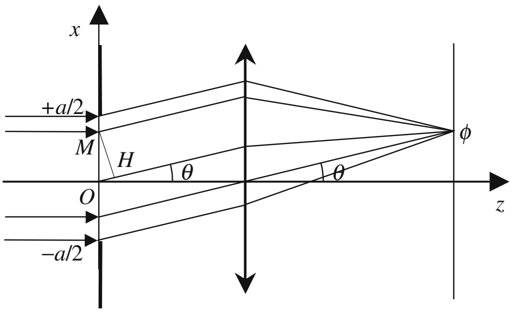


Figure 7.12. Fraunhofer diffraction by a thin slit. The diffraction pattern is observed in the focal plane of the lens.

We first consider all the rays that arrive at the focal point ϕ associated to the direction of propagation corresponding to the angle θ, the diffracted ray coming from point M has covered a shorter path than the ray coming from point O, the difference is equal to OH and the phase difference is equal to $\varphi = kOH = kOM\sin\theta$. The global vibration diffracted by the slit in the direction θ and at point ϕ is given by the integral

$$B(\theta) = \frac{jb_0 L}{\lambda} \int_{-a/2}^{+a/2} e^{-jk\sin\theta}\, dx,$$

$$B(\theta) = \frac{jb_0 L}{\lambda} \frac{1}{jk\sin\theta} \int_{(-jka\sin\theta)/2}^{(+jka\sin\theta)/2} e^{-u}\, du = -j\frac{b_0 L}{\pi\sin\theta} \sin\left(\frac{ka\sin\theta}{2}\right), \tag{7.9}$$

$$B(\theta) = -jB_0 \frac{\sin\beta}{\beta} = \operatorname{sinc}\beta, \quad \text{with} \quad \beta = \frac{ka}{2}\sin\theta \quad \text{and} \quad B_0 = \frac{b_0 aL}{\lambda},$$

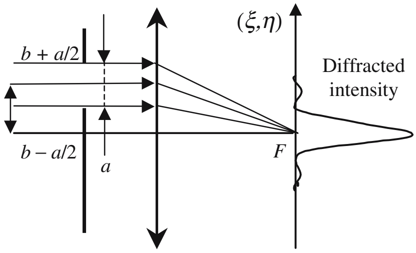
$$I(\theta) = B_0^2 \operatorname{sinc}^2\beta.$$

Determination of the Diffraction Pattern Using the Fourier Transform

$$\begin{aligned} \operatorname{Rect}\left[\frac{x}{a}\right] &= 1 \quad \text{if } |x| \le \frac{a}{2}, \\ \operatorname{Rect}\left[\frac{x}{a}\right] &= 0 \quad \text{if } |x| > \frac{a}{2}. \end{aligned} \tag{7.10}$$

It is established in mathematics that the Fourier transform $F(u)$ of a rectangle function is a *sinc* and that, omitting the proportionality coefficient, we have

$$F(u) = \text{FT}\left\{\operatorname{Rect}\left[\frac{x}{a}\right]\right\} = \frac{\sin\left(\dfrac{kau}{2}\right)}{\dfrac{kau}{2}} = \operatorname{sinc}\left(\frac{kau}{2}\right).$$



$$f(x) = \mathrm{Rect}\left[\frac{(x-b)}{a}\right],$$

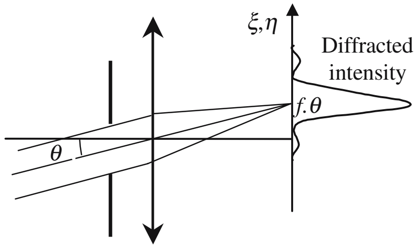
$$\mathrm{FT}\left\{\mathrm{Rect}\left[\frac{(x-b)}{a}\right]\right\} = e^{jkbu}\,\mathrm{sinc}\left(\frac{kau}{2}\right),$$

$$F(\xi) = e^{jkb\xi/f}\,\mathrm{sinc}\left(\frac{ka\xi}{2f}\right),$$

$$I = I_0\,\mathrm{sinc}^2\left(\frac{ka\xi}{2f}\right).$$

Figure 7.13. As compared to Figure 7.12, the slit has been translated parallel to itself. The diffraction pattern remains exactly the same, except that phases at point F are not the same.

The method for obtaining the Fourier transform of the function Rect(x/a) is indeed very similar to the calculations of formulas (7.9) and (7.10). The interest in using the Fourier method is that we can take full advantage of all the results that have been accumulated by mathematicians concerning the Fourier transform. Figures 7.13, 7.14, and 7.15 are good examples of the advantages of the method, which also justify an affirmation that we made earlier, see Figure 7.3: a diffraction pattern is all the more spread as the diffracting aperture is smaller.



$$f(x) = \mathrm{Rect}\left[\frac{x-b}{a}\right],$$

$$FT\left\{\mathrm{Rect}\left[\frac{x-b}{a}\right]\right\} = e^{jkbu}\,\mathrm{sinc}\left[\frac{kau}{2}\right],$$

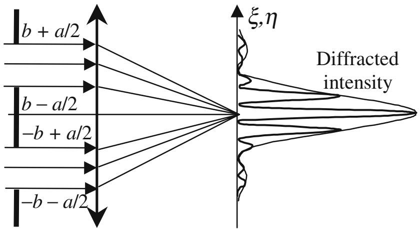
$$F(\xi) = e^{jkb\xi/f}\,\mathrm{sinc}\left[\frac{ka\xi}{2f}\right],$$

$$I = I_0\,\mathrm{sinc}^2\left[\frac{ka\xi}{2f}\right].$$

Figure 7.14. The respective positions of the slit and of the lens are the same as in Figure 7.12, but the incidence is now oblique. The diffraction pattern has the same aspect, but is now centered on the focal point associated to the direction θ.

7.2.2.2. *Several Equidistant Slits*

Two Slits



$$f(x) = \mathrm{Rect}\left[\frac{x-b}{a}\right] + \mathrm{Rect}\left[\frac{x-b}{a}\right],$$

$$FT\{f(x)\} = \left(e^{jkbu} + e^{-jkbu}\right)\mathrm{sinc}\left[\frac{kau}{2}\right]$$

$$= 2\cos(kbu)\,\mathrm{sinc}\left[\frac{kau}{2}\right],$$

$$F(\xi) = 2\cos\left(kb\frac{\xi}{f}\right)\mathrm{sinc}\left[\frac{ka\xi}{2f}\right],$$

$$I = 4I_0\,\mathrm{sinc}^2\left[\frac{\pi a\xi}{\lambda f}\right]\cos^2\left[2\pi\frac{b\xi}{\lambda f}\right].$$

Figure 7.15. The two slits give two diffraction patterns that have exactly the same amplitude repartitions, see Figure 7.13, but different phase repartitions. The phase difference depends on the position in the focal plane and produces an interference pattern described by a sine, superimposed on the diffraction pattern described by a sinc.

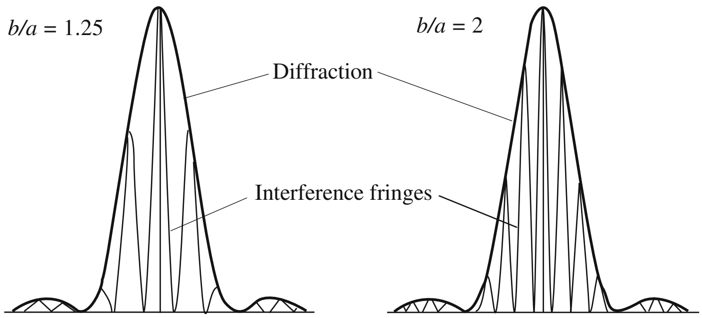


Figure 7.16. Variation of the light intensity in the focal plane of Figure 7.15. The interference fringes, analogous to Young's slit fringes, are "modulated" by diffraction fringes (see Figure 6.1).

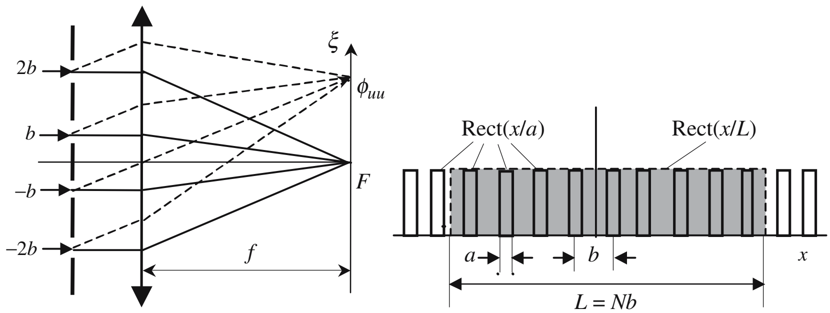


Figure 7.17. Diffraction by N identical and regularly spaced slits.

Many Slits

Before studying the diffraction pattern given by a great (possibly infinite) number of slits, we will introduce some useful mathematical functions:

- Dirac's pulse $\delta(\zeta)$, where ζ is a real dimensionless variable,

$$\delta(\zeta = 0) \to \infty,$$
$$\delta(\zeta \neq 0) = 0,$$
$$\int_{-\infty}^{\infty} \delta(\zeta)\,d\zeta = 1. \tag{7.11.a}$$

- Dirac's pulse occurring for a given value, n, of the variable $\delta_n(\zeta) = \delta(\zeta - n)$,

$$\delta_n(\zeta = n) = \delta(0) \to \infty,$$
$$\delta_n(\zeta \neq n) = 0,$$
$$\int_{-\infty}^{\infty} \delta_n(\zeta)\,d\zeta = 1. \tag{7.11.b}$$

- Product of convolution, $f(\zeta) \otimes \delta_n(\zeta)$, of a function $f(\zeta)$ by $\delta_n(\zeta)$,

$$f(\zeta) \otimes \delta_n(\zeta) = \int_{-\infty}^{\infty} f(\zeta)\delta_n(\zeta)\,d\psi = f(\zeta = n).$$

- Fourier transform of a product of convolution,

$$\mathrm{FT}(f(\xi) = F(\eta)), \quad \mathrm{FT}(g(\xi) = G(\eta)),$$
$$\mathrm{FT}[f(\xi)g(\xi)] = F(\eta) \otimes G(\eta),$$
$$\mathrm{FT}[f(\xi) \otimes g(\xi)] = F(\eta)G(\eta).$$

- Dirac's comb: succession of equally spaced Dirac pulses,

$$\mathrm{Comb}(\zeta) = \sum_{n=-\infty}^{\infty} \delta_n(\zeta), \quad n \text{ is an integer.} \tag{7.11.c}$$

- Infinite comb of slits (see Figure 7.17):

$$\mathrm{SlitComb}(x) = \mathrm{Comb}\left(\frac{x}{b}\right) \otimes \mathrm{Rect}\left(\frac{x}{a}\right) = \left\{\sum_{n=-\infty}^{\infty} \delta_n\left(\frac{x}{b}\right)\right\} \otimes \mathrm{Rect}\left(\frac{x}{a}\right). \qquad (7.11.d)$$

The rectangle function Rect[x/a] has already been introduced in Section 7.2.2.1, formula (7.10).

- An infinite comb of slits is easily reduced to a finite comb of only N slits (see Figure 7.17), by multiplying by the function Rect[x/Nb]:

$$N/\mathrm{SlitComb}(x) = \left\{\mathrm{Comb}\left(\frac{x}{b}\right) \otimes \mathrm{Rect}\left(\frac{x}{a}\right)\right\} \mathrm{Rect}\left(\frac{x}{Nb}\right). \qquad (7.11.e)$$

Diffraction Pattern of a Comb of Slits

The Fourier transform is a convenient and powerful tool for studying diffraction patterns, and provides a nice illustration of the respective roles of diffraction and interference and also of multiple interference. Making the Fourier transform of the various formulas (7.11) and the products of convolution needs rather tricky mathematical manipulations, we will develop the full calculations only in the cases where the number of slits is infinite (formulas (7.11.c) and (7.11.d)).

The Fourier transform of Dirac's pulse is another Dirac pulse; in the same way, the Fourier transform of a Dirac comb is another Dirac comb. Observed in the focal plane of a lens, the Fraunhofer diffraction pattern of an infinite number of parallel equidistant infinitely narrow slits, illuminated by a monochromatic planar wave, is an infinite set of parallel and infinitely thin bright lines.

The infinitely narrow slits are now replaced by an infinite set of narrow slits (width equal to a). The diffraction pattern is obtained from the Fourier transform of formula (7.11.d),

$$\mathrm{FT}\left[\mathrm{Comb}\left(\frac{x}{b}\right) \otimes \mathrm{Rect}\left(\frac{x}{a}\right)\right] = \left\{\sum_{n=-\infty}^{\infty} \delta_n(kbu)\right\} \mathrm{sinc}\left(\frac{kau}{2}\right), \qquad (7.12.a)$$

$$I = I_0\left\{\sum_{n=-\infty}^{\infty} \delta_n\left(2\pi \frac{b}{\lambda}\frac{\xi}{f}\right)\right\} \mathrm{sinc}^2\left(\pi \frac{a}{\lambda}\frac{\xi}{f}\right). \qquad (7.12.b)$$

The first term of (7.12.a) comes from the multiple interference of all the waves diffracted by the different apertures: constructive interference only occurs in the directions for which the waves diffracted by two neighboring slits show a phase difference exactly equal to 2π. The second term comes from the diffraction by each individual slit. Formula (7.12.b) gives the variation of the light intensity on the focal plane of the lens of Figure 7.17: a set of equidistant and infinitely thin fringes have an appreciable intensity only between the first two zeros of the *sinc-function* ($|\xi| < \lambda f/a$); the narrower the slit is, the greater is the number of fringes.

The calculation, in the case of a finite number of slits with a finite width a (formula (7.11.e)), is less trivial, because of the presence of a usual product and of a convolution product. The limited number of slits corresponds to the term Rect($x/(Nb)$) in formula (7.11.e); its Fourier transform is the *sinc-function* sinc[$kNbu/2$]. In the focal plane the diffraction pattern is a comb. The light intensity of each individual tooth is given by $\text{sinc}^2[\pi Nb\xi/\lambda f]$, it has a principal maximum for $\xi = 0$, surrounded by two minima equal to zero that are obtained for

$$\xi = \pm\lambda f/Nb. \tag{7.13}$$

Two Waves and Multiple Wave Interference

In the previous arrangements diffraction and interference are produced by a regular distribution of diffracting objects, the phase shift varies versus the angle and versus the position of the object. The angular repartition of the intensity of the diffracted light shows important *principal maxima* for some directions, between them *subsidiary maxima* are found, their intensity being far smaller. Between two consecutive maxima there is a minimum having an intensity equal to zero.

The directions of the principal maxima are those for which the phase shift between two neighboring objects is precisely equal to 2π. The diffracted intensity decreases all the more rapidly as the total number of diffracting objects is larger.

- In the case of only two diffracting objects (Young's slits, for example), the variation of the intensity is a sine.
- In the case of an infinite number of diffracting objects, the variation law of the intensity is a succession of Dirac pulses.
- In the case of a finite number N of diffracting objects, the variation law is a comb tooth having a width that is all the thinner as N is larger.

Resolving Power of Diffraction Grating Spectroscope

A set of regularly spaced slits is illuminated by a planar wave that has two spectral components of respective wavelengths λ_1 and λ_2. The diffraction pattern, observed on the focal plane of a lens, is made of the superposition of the two sets of fringes of two colors. Such an arrangement is a spectroscope, its ability to separate two wavelengths is all the more interesting as the diffraction peaks have a narrower angular width.

If $\Delta\lambda$ is the smallest difference of two wavelengths that can be discriminated by the spectroscope, by definition, the resolving power of the spectroscope is equal to

$$R = \frac{\lambda}{\Delta\lambda}.$$

It is considered that two wavelengths are discriminated if the distance between the two principal maxima corresponding to λ_1 and λ_2 is larger than the distance from a principal maximum to the first minimum. The positions of the principal maxima come from Comb($b\xi/\lambda f$), and are given by

$$\xi_{\mathrm{Max},\lambda 1} = n\lambda_1 f/b, \quad \xi_{\mathrm{Max},\lambda 2} = n\lambda_2 f/b \quad \rightarrow \quad (\xi_{\mathrm{Max},\lambda 1} - \xi_{\mathrm{Max},\lambda 1}) = n(\lambda_1 - \lambda_2) f/b.$$

The distance to a principal minimum is obtained from (7.13): $\Delta\xi = \lambda f/Nb$, the resolving power is equal to

$$R = \frac{\lambda}{\Delta\lambda} = \frac{\lambda}{(\lambda_1 - \lambda_2)} = nN. \tag{7.14}$$

The total number N of diffracting slits mainly determines the resolving power.

7.2.3. *Far Field Diffraction by a Circular Aperture*

The circular aperture of radius a of Figure 7.18 is described by the following expression:

$$C(x, y) = 1 \quad \text{if } (x^2 + y^2) \leq a^2,$$
$$C(x, y) = 0 \quad \text{if } (x^2 + y^2) > a^2.$$

The Fourier tranform of $C(x, y)$ is a Bessel function of the first order,

$$F(u, v) = \mathrm{FT}[C(x, y)] = 2\frac{J_1(Z)}{Z} \quad \text{with} \quad \rho = \sqrt{u^2 + v^2} \quad \text{and} \quad Z = 2\pi\frac{a}{\lambda}\rho.$$

The function $J_1(Z)/Z$ plays, in the case of a circular aperture, the same role as the *sinc-function* in the case of a rectangular slit. The graph of $(J_1(Z)/Z)^2$ looks like that of a square *sinc*: a central maximum surrounded by subsidiary less intense maxima, with a minimum equal to zero between two consecutive maxima. The first minimum is obtained for $Z = 3.83$, the corresponding radius, which is the radius of the Airy disk, see Figure 7.18, is equal to

$$\rho_{\text{Airy disk}} = f\frac{3.83}{\pi}\frac{\lambda}{2a} = 1.22\frac{\lambda}{2a}f. \tag{7.15}$$

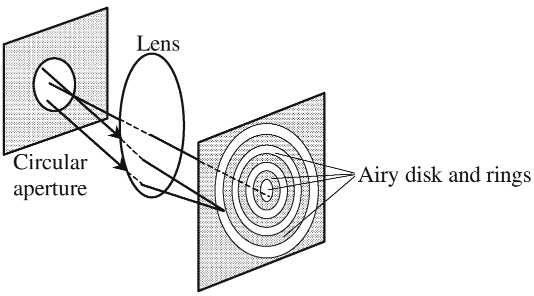


Figure 7.18. The diffraction pattern of a circular aperture, observed in the focal plane of a lens, is made of a disk surrounded by rings. The central disk is often called an "Airy disk."

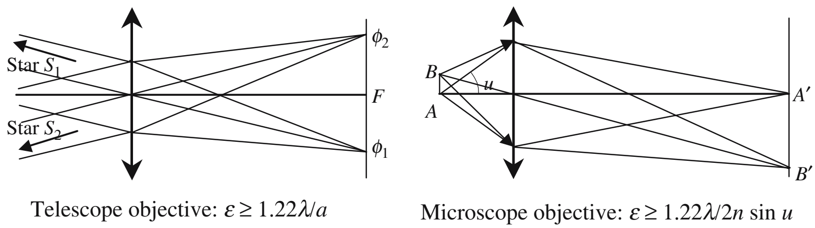


Figure 7.19. Resolving power of an objective. For the two images, (ϕ_1 and ϕ_2) or (A' and B'), to be separated when observed through an eyepiece, it is considered that their distance should be greater than the radius of the first dark Airy ring.

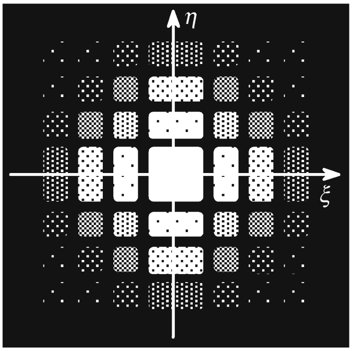
Resolving Power of an Imaging System

Because of diffraction, the image of a point source is not a point, but an Airy disk; the images of two neighboring points are seen separately if their Airy disks don't overlap too much. It is considered that this is case if the center of one disk coincides with the first zero of the other (this approximation is called the *Rayleigh criterion*).

The distance between two points that are seen separately is the *limit of resolution* of the objective. In the case of a telescope the limit of resolution is an angle equal to $1.22\lambda/2a$ according to formula (7.15). The case of a microscope objective raises some difficulty since it involves angles that are not small, it is found that the angular limit of resolution is $1.22\lambda/2n \sin u$, where u is the greatest value of the angle that makes, with the axis, a ray entering into the objective. It is interesting to notice that, in both cases, the resolving power is all the better as the objective transmits a higher light flux to the eyepiece.

7.2.4. *Far Field Diffraction of a Rectangular Aperture*

An (a, b) rectangular aperture is represented by the product of two rectangle functions with respective arguments equal to x/a and y/b. The corresponding diffraction pattern is the product of two *sinc-functions*. In Figure 7.20 is shown the diffraction pattern of a square aperture: it is a regular arrangement of squares and rectangles having their respective centers at the nodes of a network with a step equal to λ/a. The light intensity rapidly decreases with the distance to the origin, and only the most central have been represented. At the center is a square, the side of which is equal to $2\lambda/a$.



$$f(x, y) = \mathrm{Rect}\left[\frac{x}{a}\right]\mathrm{Rect}\left[\frac{y}{a}\right],$$

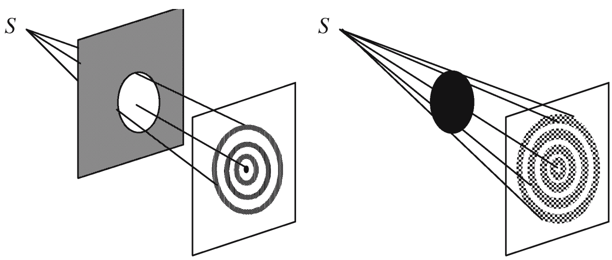
$$F(u, v) = \mathrm{sinc}\left[\frac{\pi a}{\lambda}u\right]\mathrm{sinc}\left[\frac{\pi a}{\lambda}v\right],$$

$$I = I_0\,\mathrm{sinc}^2\left[\frac{\pi a}{\lambda}u\right]\mathrm{sinc}^2\left[\frac{\pi a}{\lambda}v\right].$$

Figure 7.20. Far field diffraction pattern of a square aperture (side a).

7.3. Fresnel Diffraction

The integration of the integrals in the case of Fresnel's diffraction is more complicated than in the case of Fraunhofer's diffraction. We are going to describe a method that was used by Fresnel to determine the intensity of the light that is diffracted by an opaque circular disk, at a point located on its axis. Submitted to the French Academy of Sciences in 1818, this calculation had been rejected by Poisson who objected that one consequence was that some light should exist only in the middle of the geometric shadow. Arago who participated in the Academy meeting made the experiment in his laboratory and could demonstrate experimentally the existence of the litigious bright point, bringing a clear argument in favor of the wave theory of light. The experimental arrangement is described in Figure 7.21.



(a) Circular aperture (b) Opaque disk

Figure 7.21. Diffraction by a circular aperture or by an opaque disk. The two diffraction patterns have the same general aspect. In the case of the opaque disk, as soon as the screen is sufficiently remote, there is always a bright dot at the center. In the case of the circular aperture the center is either bright or dark, according to the position of the screen.

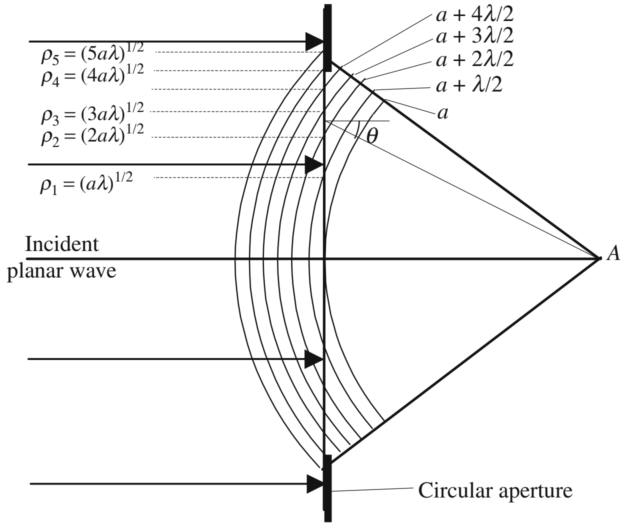


Figure 7.22. Definition of the Fresnel zones associated to a given point A. The intersection of a set of spheres, centered at point A and having radii, respectively, equal to a, $a + \lambda/2$, $a + 2\lambda/2$, $a + 3\lambda/2, \ldots$, with the plane of the aperture, are concentric circles. A Fresnel zone is delimited by two consecutive circles. The first zone is a circle, the following zones are circular rings.

Fresnel Zones

The Fresnel method of integration of the integrals is made possible because of a suitable division of the domain of integration: the elementary areas are defined in Figure 7.22. From geometric considerations it can be shown that all the Fresnel zones have equal areas $\pi a\lambda$ and that the radius of the nth zone is equal to $\rho_n = (na\lambda)^{1/2}$.

To determine the diffracted vibration sent to point A, we start by adding the elementary vibrations sent by all the points of the same zone, and then we add the contributions of the different zones. Since all the Fresnel zones have the same area, the moduli of the contributions of the different zones are almost equal. In fact, the moduli decrease slowly as the radius is increased, and this is for two reasons. First, the distance to point A increases; the second, and most important, reason is due to the fact that the amplitude of the vibration sent to point A decreases with the angle θ between the normal to the plane of the zone and the line that joins a point of the zone to point A.

The total vibration at point A can be calculated using the Huygens-Fresnel principle; it is found, but the result is quite obvious because of the method that has been chosen to divide the surface of integration, that the vibrations arriving from two consecutive rings have opposite phases. The vibration at point A is the sum of alternately positive and negative terms, let s_i be the vibration coming from the Fresnel zone labeled I, and we have

$$s = \sum_i s_i = s_1 + s_2 + s_3 + s_4 + s_5 + \cdots.$$

In the case of the circular aperture of Figure 7.21(a), the total number of Fresnel zones is either even or odd. In the first case the waves interfere destructively: the intensity at point A is almost equal to zero, the center of the diffraction pattern is dark. On the contrary, an odd number of zones gives a maximum of light at the center.

The case of an opaque disk is different since the more central Fresnel zones are not illuminated; let p be the label of the first nonobliterated zone, the total vibration is now given by

$$s = \sum_{i=p}^{\infty} s_i = s_p + s_{p+1} + s_{p+2}, \ldots,$$

$$s = \frac{s_p}{2} + \left(\frac{s_p}{2} + \frac{s_{p+1}}{2}\right) + \left(\frac{s_{p+1}}{2} + \frac{s_{p+2}}{2}\right) + \cdots. \tag{7.16}$$

To be rigorous we should consider the convergence of the sum; we will admit that the two terms in parentheses cancel and that the first term, $s_p/2$, is the only remaining one. Finally, there is always a maximum of light at the center.

Fresnel Zone Plate

Although they have been considered, for a long time, as esthetical curiosities, the Fresnel zone plates have now found interesting applications for making lenses in cases where it is difficult to substantially increase the index of refraction with regard to the surroundings, for example, in integrated optics or in X-ray imaging systems.

The principle of a zone plate is rather simple; opaque rings obliterate one Fresnel zone over two, the purpose being to keep only, in the development of formula (7.16), the terms that have the same sign.

If $a = 1\,\mathrm{m}$ and $\lambda = 0.5\,\mu\mathrm{m}$, the radius of the first zone is 0.707 mm. A zone plate can readily be obtained by drawing, using a large scale, circles having radii varying as the square root of the successive integers. The rings between two consecutive circles are then darkened and photographed using a given magnification; the resulting slide is a zone plate. The radius of the nth ring is $\rho_n \approx \sqrt{na\lambda}$, its thickness is equal to the difference $\rho_{n+1} - \rho_n \approx \sqrt{a\lambda}/2\sqrt{n} = \rho_0/2\sqrt{n}$. It can be shown that the optical quality of the image increases with the number of zones, the transverse size (aberration) of the focal point is of the order of the thickness of the last ring. In the case of a photographic process, this thickness is fixed by the photoplate.

If a planar monochromatic wave is sent orthogonal to a Fresnel zone plate, an accumulation of light is found in the vicinity of the different points F_n, see

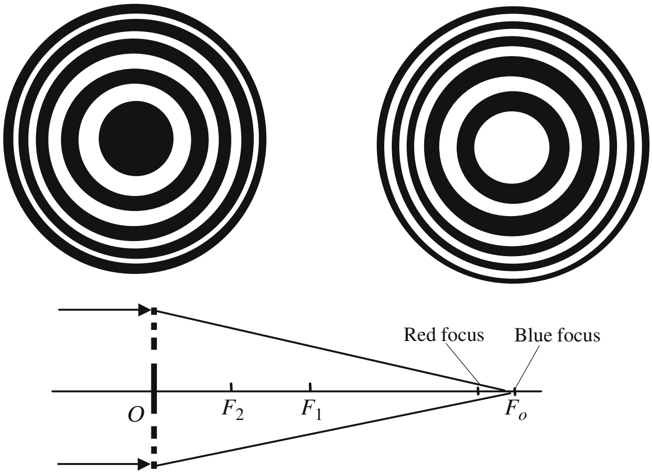


Figure 7.23. Examples of Fresnel zone plates. On the left-hand plate, the even zones are suppressed; on the right-hand plate, the odd zones have been suppressed.

Figure 7.23, that satisfy the following relation, in which n is an integer, R is the radius of the first zone, and λ is the wavelength,

$$OF_n = \frac{1}{2n+1}\frac{R^2}{\lambda}.$$

A Fresnel zone plate appears to be a lens that has several focal points. Of course the chromatic aberrations are important since the focal length is proportional to the inverse of the wavelength. If, instead of using a parallel incident beam, a spherical wave centered at some point P is sent to the plate, the transmitted light is made of several spherical waves centered at the points Q_n defined by

$$\frac{1}{OP} + \frac{1}{OQ_n} = (2n+1)\frac{R^2}{\lambda} = \frac{1}{OF_n}.$$

Number of Fresnel Zones Inside an Aperture-Limited Spherical Wave

The solid angle of a full spherical wave is equal to 4π steradians; the beams that are used in systems working in Gauss conditions are far less open. As soon as the aperture of a spherical wave is quite limited, diffraction effects are observed. A good evaluation of the effect of diffraction is provided by the notion of *Fresnel's number*. The definition of *Fresnel's number*, associated to a given spherical wave and a given circular aperture, is given in Figure 7.24 where a spherical wave is obtained by focusing a planar wave with an aberration-free lens. We consider concentric spheres centered at the focal point; the first one is tangent to the plane of the aperture and its radius is

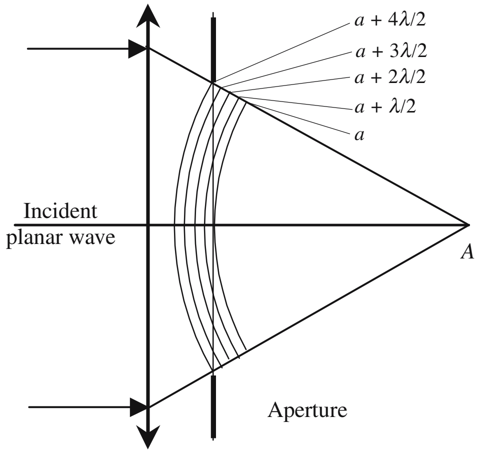


Figure 7.24. The Fresnel number of an aperture-limited spherical wave is equal to the number of Fresnel zones inside the open part of the aperture.

equal to a, the radii of the other spheres are regularly increased $\lambda/2$ by $\lambda/2$. The intersections of the different spheres with the plane of the aperture are circles. A Fresnel zone is a ring between two adjacent circles. The Fresnel number is just the total number of zones inside the aperture. The larger the Fresnel number, the smaller is the diffraction spot in the vicinity of the focus.

7.4. Diffraction Gratings

During the last decades the technology for making diffraction gratings has considerably improved, opening the way to efficient mass production methods. Diffraction gratings are now usual and relatively cheap components.

7.4.1. *Diffraction—Diffusion of Light*

Diffraction is observed any time that an electromagnetic wave hits some small-sized object, by small size we mean of the order of one to fifty times the wavelength. The object under consideration is often made of the collection of smaller (as compared to the wavelength) constituents that will be called *particles*, each of them participates in the diffraction process.

The light that is diffracted by the global object is the result of the interference of all the wavelets diffracted by the different particles. Although the following formal distinction is not universally used, we will distinguish two main cases:

- *Diffusion*, where the spatial distribution of the particles is at random. Typical examples are the repartition of water drops in fog, or of nitrogen and oxygen molecules in the atmosphere.

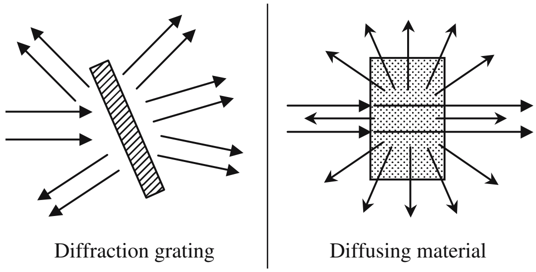


Figure 7.25. Comparison between diffusion and diffraction. For a periodic arrangement, the diffracted beams are observed only along specific directions called the orders of diffraction. Diffused light is emitted in all directions.

- *Diffraction*, where the distribution of the particles is perfectly organized. The two most famous examples are the equidistant grooves of a grating and the arrangement of the atoms inside a perfect crystal.

From a practical point of view, the main difference between diffraction and diffusion is related to the directivity of the diffracted beams. No privileged direction exists in the case of diffusion, the diffused light being emitted in all directions. On the contrary, in the case of a regular arrangement of the particles there are only a few directions in which the wavelets interfere positively, these directions are called the *orders of diffraction*.

A diffraction grating is made of identical, parallel, and equidistant grooves that are considered to have an infinite length. The most important characteristic of a grating is its periodicity, which is the separation of two adjacent grooves. The directions of the diffraction orders are fixed by the periodicity.

Another characteristic of a grating is the shape of its grooves. This shape doesn't play any role in the determination of the direction of the orders, however it determinates how the incident energy is shared by the different orders. This is the reason why a significant effort is made to achieve given profiles, so that only one order will carry away most of the energy.

Working in the far field diffraction regime, a grating must be placed between two lenses, a collimating lens and then a focusing lens. As shown in Figure 7.27, two main different arrangements are used, the light being either reflected or transmitted by the grating. In fact, the transmission arrangement is handicapped by the fact that the light travels through the substrate on which the grating is deposited; consequently, this substrate should be very homogeneous with a perfectly planar and well-polished second interface, in order to avoid unwanted phase modifications.

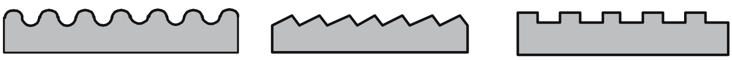


Figure 7.26. Some usual groove profiles. The periodicity ranges from $10\,\mu$m to $0.5\,\mu$m, i.e., 100 to 2000 grooves per millimeter.

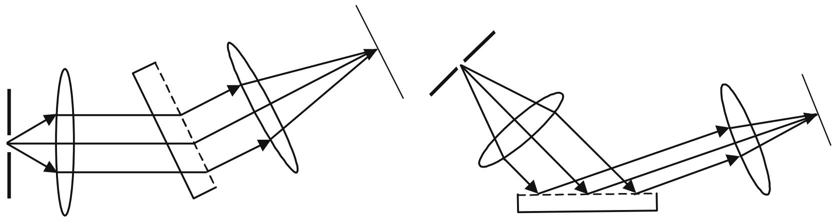


Figure 7.27. A grating can work either by transmission or by reflection. The second arrangement is preferred since it is not concerned with a possible lack of homogeneity of the substrate that supports the grating.

7.4.2. *The Bragg Formula*

Diffraction by a grating is a multiple interference process. The total vibration diffracted in a given direction is obtained by summing up all the waves diffracted by the different particles of the diffracting object. In the case of a grating, the wavelets diffracted by all the particles belonging to the same groove are first added, and then the contributions of the different grooves. As the total number of grooves is enormous, it's a highly multiple interference process.

The directions of constructive interference are those for which the phase difference between the contributions of two neighboring grooves is very close to a multiple of 2π; in the case of an infinite number of grooves the phase difference should be strictly equal to a multiple of 2π. This condition is often referred to as the *phase matching condition.*

The Bragg formula is a relationship between the angle of incidence and the angle of diffraction. It indicates the directions in which the different diffracted beams are emitted, but it doesn't give any indication as to how the energy is shared between them, the shape of the grooves governs the partition rule.

7.4.2.1. *The Bragg Formula in the Optical Case*

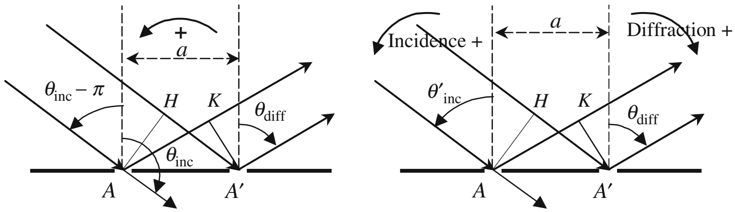
The grating will be considered here as an infinite set of parallel and equidistant infinitely thin grooves; the only parameter to be defined is the periodicity a, which is the separation of two neighboring grooves,

$$(\sin\theta_{\text{diff}} - \sin\theta_{\text{inc}}) = p\frac{\lambda}{a} \quad \text{(same orientation for the two beams),} \tag{7.17.a}$$

$$(\sin\theta_{\text{diff}} + \sin\theta_{\text{inc}}) = p\frac{\lambda}{a} \quad \text{(opposite orientation for the two beams).} \tag{7.17.b}$$

The orientation (7.17.b) is often more convenient.

Formulas (7.17) give, for a given value of the angle of incidence, a set of values for the angle of diffraction, labeled by the integer p. Only the values of p corresponding to absolute values of $\sin\theta_{\text{diff}}$ smaller than one should, a priori,



$\delta = AK - A'H = a[\sin\theta_{diff} - \sin\theta_{inc}]$
$\theta_{diff} < 0$ and $\theta_{inc} > 0$
(a)

$\delta = AK - A'H = a[\sin\theta_{diff} + \sin\theta_{inc}]$
$\theta_{diff} > 0$ and $\theta_{inc} > 0$
(b)

Figure 7.28. Evaluation of the difference δ of the optical lengths of the waves diffracted by two neighboring grooves. According to the authors, two different orientations are commonly used for the angles between the rays and the normal to the plane of the grating.

be considered. An interpretation of the case when the sine is greater than one will be given in Section 7.4.2.3.

The directions of the different orders of diffraction explicitly depend on the wavelength, except for the zero order, which is said to be nondispersive.

7.4.2.2. *The Bragg Formula for a Three-Dimensional Crystal Lattice (Diffracting X-Rays)*

The distance between atoms (angstroms) being far smaller than optical wavelengths, a crystal lattice doesn't diffract light waves; the situation is, of course, different for X-rays. By the way, gratings can also diffract X-rays; the experiment was successfully made at a time when some doubt existed about a common identity between X-rays and light waves. Let us come back to the diffraction of X-rays by a crystal; the practical interest is enormous, since it provides information about the crystal lattice. From the angular repartition of the diffracted beams and the Bragg formula, we obtain the shape and size of the crystal cell; from the comparison of the respective intensities of the different orders of diffraction, we obtain information about the chemical nature of the atoms, or groups of atoms in the cell.

When an X-ray electromagnetic wave reaches the crystal, it penetrates deep inside and each atom reemits X-rays in all directions. Because of the three-dimensional arrangement of the diffracting particles, the expression of the phase-matching condition is rather different from the case of a diffraction grating. We will proceed in two steps: first we consider the diffracting particles belonging to the same reticular plane, and then consider the phase matching between the vibrations produced by two neighboring parallel reticular planes.

We refer to Figure 7.30, the different atoms of a given reticular plane are labeled $A_1, A_2, \ldots, A_i, \ldots$, the phase-matching condition should be written for any couple of points A_iA_k:

$$(\sin\theta_{diff} - \sin\theta_{inc}) = p\frac{\lambda}{A_iA_k}.$$

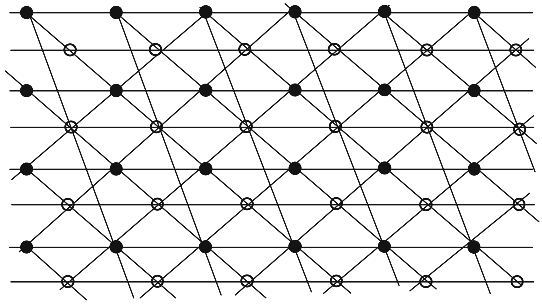


Figure 7.29. Schematic representation of some reticular planes of a two-dimensional lattice with two different atoms in each individual cell.

Due to the high value of the possible number of pairs, the only possible value is $p = 0$, which means that the angles of incidence and diffraction are equal and that each reticular plane behaves as a mirror on which X-rays are reflected according to the Snell-Descartes law for reflection. Finally, we write the phase-matching condition between the waves that are reflected by two consecutive planes. In the X-ray domain, it's more usual to refer to the direction of the beams with regard to the angle α with the reticular planes, d being the distance between two neighboring planes, the difference of the optical paths is given by

$$IA_2 + A_2J = 2d\sin\alpha \quad \rightarrow \quad 2d\sin\alpha = p\lambda \quad \text{(Bragg formula for a crystal).} \tag{7.18}$$

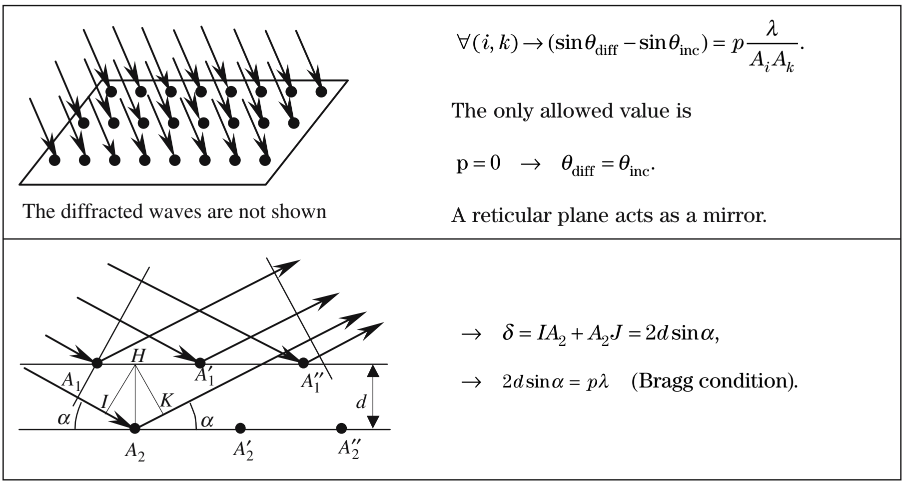
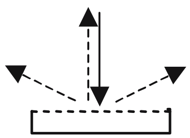
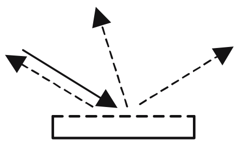


Figure 7.30. X-ray diffraction by a crystal lattice. A given incident beam will generate a diffracted beam if, and only if, its direction is properly oriented so that the Bragg condition is fulfilled.

7.4.2.3. *Numerical Applications of the Bragg Formula in the Optical Case*

Knowing the periodicity, a, of a grating, the wavelength, λ, and the angle of incidence, θ_{inc}, we want to calculate the angles of diffraction θ_{diff}. The number N of grooves per millimeter (or per inch) is often preferred to the periodicity: $N = 1/a$. We use the orientation defined by (7.13.b) (opposite orientations for the incident and diffracted beam).

$N = 1800\ \text{grooves/mm} = 1.8\ \mu\text{m}^{-1}, \quad a = 0.5\ \mu\text{m}, \quad \lambda/a = N\lambda = 0.9.$	
 Normal incidence	$\theta_{inc} = 0,$ $\sin\theta_{diff} = -\sin\theta_{inc} + p\lambda/a = 0.9p,$ $\lvert\sin\theta_{diff}\rvert \le 1 \quad\rightarrow\quad -1 \le 0.9p \le 1 \quad\rightarrow\quad -1.1 \le p \le 1.1,$ $p \in \{-1; 0; 1\}, \quad \sin\theta_{diff} \in \{-0.9; 0; 0.9\},$ $\theta_{diff} \in \{-65.4°; 0; 65.4°\}.$
 Incidence 60°	$\theta_{inc} = 60°,$ $\sin\theta_{diff} = -0.866 + p\lambda/a = -0.866 + 0.9p,$ $\lvert\sin\theta_{diff}\rvert \le 1 \quad\rightarrow\quad -0.148 \le 0.9p \le 2.07,$ $p \in \{0; 1; 2\}, \quad \sin\theta_{diff} \in \{-0.866; 0.034; 0.934\},$ $\theta_{diff} \in \{-60°; 1.95°; 69°\}.$

$N = 100\ \text{mm}^{-1}$; $\lambda = 0.5\,\mu\text{m}$; normal incidence, $\theta_{inc} = 0$. Calculate the total number of diffracted orders:

- $N\lambda = 0.05 \quad\rightarrow\quad \sin\theta_{diff} = 0.05p \quad\rightarrow\quad -20 \le p \le 20.$
- The total number of diffracted orders is 41. This example shows that the number of diffracted orders is all the larger as the number of grooves per unit length is smaller.

$N = 2000\ \text{mm}^{-1}$; $\lambda = 0.5\,\mu\text{m}$; $\theta_{inc} = 10°$. Calculate the total number of diffracted orders:

- $N\lambda = 1.2 \quad\rightarrow\quad \sin\theta_{diff} = -\sin 10° + 1.2p = -017 + 1.2p.$
- $0.69 \le p \le 0.95 \rightarrow$ the only possible value is $p = 0 \rightarrow \theta_{inc} = -\theta_{diff}$. The zeroth order is the only order allowed, the periodicity is too small and the grating is equivalent to a mirror, there is no dispersion.

7.4.2.4. *Littrow's Configuration*

Incidence conditions may be found where the diffracted beam propagates exactly in the opposite direction to the incident beam, $\theta_{\text{inc}} = -\theta_{\text{diff}}$; the grating is then said to work in the Littrow condition, which is defined by

$$\sin\theta_{\text{Littrow}} = \frac{pN\lambda}{2}. \tag{7.19}$$

For a given grating and a given order of diffraction, except the zeroth order, the angle θ_{Littrow} corresponding to the Littrow condition depends on the wavelength.

Numerical Application

- $N\lambda = 1$. There are only two possible values for p in formula (7.19):
  - $p = 0$, $\theta_{\text{Littrow}} = 0$, this case is not interesting, since the zero-order mode is not dispersive.
  - $p = 1 \quad \rightarrow \quad \sin\theta_{\text{Littrow}} = \frac{1}{2} \quad \rightarrow \quad \theta_{\text{Littrow}} = 30°$.
- $N\lambda = 0.6 \quad \rightarrow \quad \sin\theta_{\text{Littrow}} = 0.3p \quad \rightarrow \quad \sin\theta_{\text{Littrow}} \in \{0;\ 0.3;\ 0.6;\ 0.9\}$.
  - There are four different orders for which the Littrow condition can be satisfied.

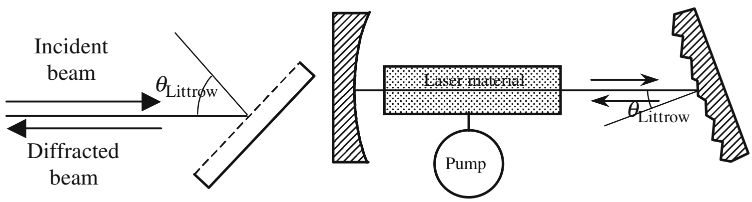


Figure 7.31. In the Littrow configuration the grating sends back the diffracted beam exactly in the opposite direction of the incident beam. As θ_{Littrow} depends on the wavelength such an arrangement is commonly used as the second mirror of a tunable laser, a rotation of the grating changes the laser frequency.

7.4.2.5. *A Diffraction Grating May Support Evanescent Waves*

A formal application of the Bragg formula may give values of $\sin\theta_{\text{diff}}$ that are larger than +1 or smaller than −1. A similar situation has already been encountered when using the Snell-Descartes law of refraction in the conditions of total internal reflection. An interpretation has been given thanks to evanescent waves. The situation here is very analogous: each time that the Bragg formula gives a sine or cosine greater than one, in absolute value, evanescent waves can be observed.

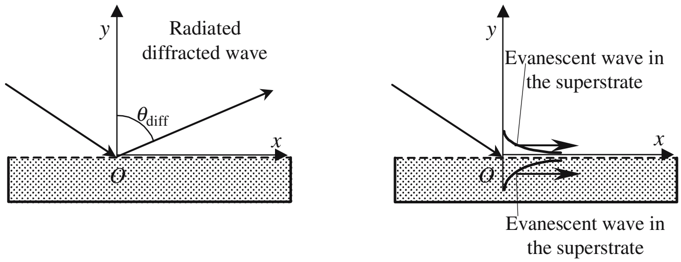


Figure 7.32. Radiated and evanescent modes of diffraction.

We consider a diffracted mode such that the sine of the angle of diffraction is equal to some real number α_p, with an absolute value greater than one,

$$\sin\theta_{\text{diff}} = \alpha_p > 1,$$

the cosine of θ_{diff} is obtained by the usual formula, $\cos^2\theta_{\text{diff}} + \sin^2\theta_{\text{diff}} = 1$,

$$\cos\theta_{\text{diff}} = \pm j\sqrt{\alpha_p^2 - 1} = \pm j\beta_p \quad \text{with} \quad \beta_p^2 = \alpha_p^2 - 1.$$

Let us call $\boldsymbol{V}(x, y)$ the electromagnetic vibration at some point $M(x, y)$, to obtain an expression for $\boldsymbol{V}(x, y)$ we go back to the usual form of a planar wave having a wave vector $\boldsymbol{k}$ making an angle θ_{diff} with the axis Oy, see Figure 7.32.

We set

$$\boldsymbol{OM} = \boldsymbol{r} = x\boldsymbol{x} + y\boldsymbol{y}, \quad \boldsymbol{k} = \frac{2\pi}{\lambda}(\boldsymbol{x}\sin\theta_{\text{diff}} + \boldsymbol{y}\cos\theta_{\text{diff}}),$$

$$\boldsymbol{V}(x, y) = \boldsymbol{A}e^{j\omega t}e^{-j\boldsymbol{k}_p\boldsymbol{r}} = \boldsymbol{A}e^{j\omega t}e^{-j\frac{2\pi}{\lambda}(x\sin\theta_{\text{diff}} + y\cos\theta_{\text{diff}})},$$

$$\boldsymbol{V}(x, y) = \boldsymbol{A}e^{j\omega t}e^{-j\frac{2\pi}{\lambda}\alpha_p x}e^{-\frac{2\pi}{\lambda}\beta_p|y|}, \quad \boldsymbol{A} \text{ is a constant vector.}$$

$\boldsymbol{V}(x, y)$ represents a wave, that is:

A progressive wave in a direction that is parallel to the grating Ox.
An evanescent wave in a direction that is orthogonal to the grating Oy.

7.4.3. *Practical Considerations about Gratings*

7.4.3.1. *Methods for Making Gratings*

Until rather recently the gratings were quite expensive; nowadays, they can be considered as relatively cheap optical components since collective methods of elaboration are now available, a large number of identical gratings being simultaneously duplicated from an initial "master grating." The grating market is an important one, thanks to applications such as spectrometers (biology, chemistry) and, more recently, to optical communications (wavelength multiplexing).

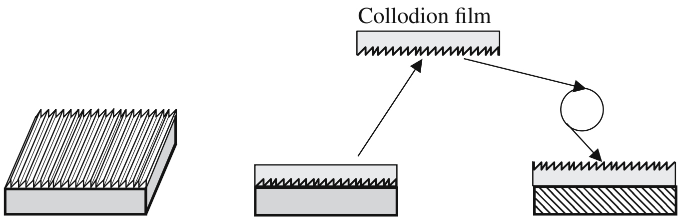


Figure 7.33. Duplication of a ruled grating. Using a diamond tool, a large number of parallel and equidistant grooves are notched on a planar glass surface, giving a master grating. A thin layer of liquid collodion is poured onto its surface. After drying, a collodion cast of the master grating is obtained, it's a solid-state film (thickness ≈ 0.1 mm) that is then glued onto a glass substrate. After metallization, an excellent replica of the initial grating is obtained.

Up to the utilization of lasers and photodyes, the gratings were obtained by engraving a glass surface with a diamond tool, the displacement of which was driven by a special ruling machine. As early as 1880, Rowland made a grating with 1700 grooves per millimeter; Michelson could produce a 30 cm long grating with 500 grooves per millimeter, which means that the diamond tool has notched 150,000 identical lines. . . . Once a first grating has been engraved, it can be replicated: Figure 7.33 shows one of the first methods that was used for duplication, it is based on the property of a varnish, called collodion.

The technology of the elaboration of gratings has largely benefited from the applied research efforts that have been devoted to the elaboration components for microelectronics, especially photolithography and cutting substrates in small pieces, using diamond saws or laser beams.

The fabrication methods rely on the polymerization of organic dyes, which can be either photosensitive (polymerization being enhanced by a suitable ultraviolet or visible radiation), or thermosensitive (polymerization being due to heating). A first grating is developed, usually 10×10 cm. Replicas are then made and cut in smaller pieces, 1×1 cm.

A dye is an organic liquid compound, made of molecules called monomers, which remains independent as long as they are not put into a situation where they polymerize, giving a solid state compound. Polymerization needs some energy that can be obtained from a light beam of a suitable color, or by a heat source of suitable temperature.

The procedure for making gratings is roughly described in Figures 7.34 and 7.35. Although there are no good reasons for that, this kind of grating is usually called a *holographic grating*.

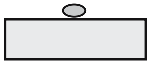
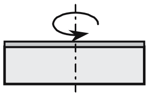
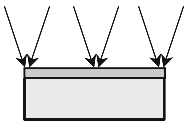
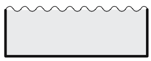
	A drop of liquid dye is poured onto a horizontal glass plate that is rapidly rotated (1000 rpm) around a vertical axis.
	The centrifugal effect spreads the dye in a thin layer that has a well-reproducible thickness (≈ 0.1–$10\,\mu$m) that can be controlled by the temperature, the speed of rotation, and the viscosity of the dye.
	A moderate heating transforms the layer in a thin film that sticks to the substrate.
	Two beams coming from the same laser interfere inside the dye. Photopolymerization occurs at the places of brilliant fringes, printing a periodic array of grooves. Exposure times are of a few minutes.
	Polymerized and nonpolymerized dyes are selectively etched, using appropriate solvents.

Figure 7.34. Modern techniques of fabrication of a master grating. The periodicity is determined by the angle of the two laser beams. The depth and shape of the grooves are reproducibly determined by the time of exposure and by the etching process.

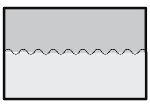
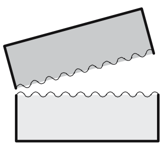
	A thick (5 mm) layer of dye is deposited onto a substrate, and then raised to a high enough temperature, so that thermopolymerization is obtained.
	After cooling, the thermopolymerized grating can be separated from the master guide that can be used again. The separation is not an easy operation.

Figure 7.35. Duplication of a master grating using a thermopolymerized dye.

7.4.3.2. *Blazed Grating*

Blazed gratings are a good illustration of *why* and *how* the energy is distributed among the different orders of diffraction. The profile of a blazed grating is shown in Figure 7.36, it's a succession of parallel narrow rectangular mirrors making an angle α with the plane of the grating. The global diffracted

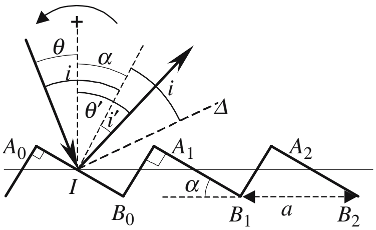


Figure 7.36. Groove profile of a blazed grating.

light is the result of the interference of the different waves diffracted by each individual mirror. The only directions in which some diffracted light can be observed are obtained from the Bragg formula.

The angular repartition of the light that is diffracted by a given mirror, A_0B_0 for example, is described by a $(\sin u/u)$ function with a maximum in the direction Δ of specular reflection on the mirror. By a suitable choice of periodicity a, angle α, angle of incidence θ, and of wavelength λ, it is possible to find conditions where Δ coincides with the direction of an order of diffraction of the grating, say the mth mode, while the directions of the other orders coincide with the zeros of the $(\sin u/u)$ function. Under such conditions the pth mode carries away all the energy

$$\cos i = \cos(\theta + \alpha) = p\frac{\lambda}{2a\sin\alpha}. \tag{7.20}$$

Numerical application: $\lambda = 0.5\,\mu\text{m}$, $N = 1000\text{ mm}^{-1}$, $a = 1\,\mu\text{m}$.

- $\alpha = 30°$. Find the angles of incidence and of refraction for the grating to be blazed in the first order: $\cos(\theta + \alpha) = 0.5 \rightarrow \theta = 0$, $\theta' = 30°$.
- Find α so that the blaze and the Littrow conditions are simultaneously fulfilled for the first order of diffraction,

$$\sin\theta = \frac{\lambda}{2a} = 0.25,$$

$$\cos(\theta + \alpha) = \frac{\lambda}{2a\sin\alpha} \quad \rightarrow \quad 2\tan\alpha\cos\theta = \frac{\lambda}{a} \quad \rightarrow \quad \alpha \approx 14.5°.$$

7.4.3.3. *The Littrow Autocollimation Mounting*

Figure 7.37 describes a simple and efficient spectrometer that uses a grating in the Littrow arrangement. For a given wavelength λ, the diffracted light comes back exactly in the opposite direction to the incident light. The entrance slit is placed in the focal plane of an achromatic lens; if the planes of the slit and of the grating are parallel, the autocollimated image just coincides with the slit, a few degrees tilt shifts the image in a place where it can be disposed as a photographic plate or an array of photodiodes.

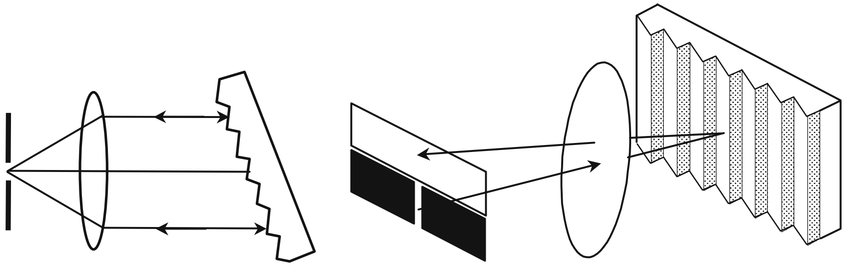


Figure 7.37. Spectrometer using the Littrow mounting.

7.4.3.4. *Monochromator*

Starting from a white source, a monochromator gives a light with a tunable wavelength λ and a narrow bandwidth. Using a grating and an optical imaging system, an entrance slit is imaged on an exit slit. The Littrow autocollimation mounting can be used, the photographic plate being replaced by a slit. The wavelength is tuned by rotating the grating around an axis in its plane. The spectral width is determined by the width of the exit slit, typically 5 to 50 μm. Chromatic aberrations should of course be avoided; this is the reason why spherical mirrors are preferred to lenses.

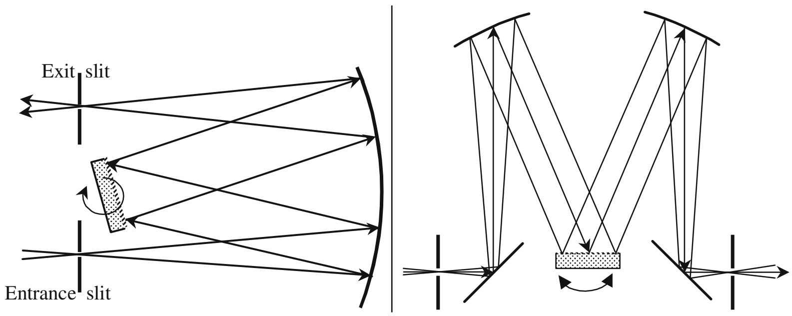


Figure 7.38. Examples of monochromators.

7.5. Holography

Denis Gabor in the middle of the twentieth century proposed the idea of holography. In the beginning, the purpose was to improve the resolving power of electron microscopes; the De Broglie waves associated to the electron beam would have been used to register the holograms that would have then been read with a monochromatic optical wave. Gabor immediately introduced the word hologram from the Greek, *holos* meaning *as a whole*, to suggest that both the amplitude and phase of the waves are registered in the photographic plate. In fact, the development of holography could only occur after the invention of lasers, the writing and reading of the hologram being made by optical waves. Denis Gabor became a Nobel Laureate in 1971.

7.5.1. *Definition of Holography*

We first consider the reason why, watching the planar picture of a landscape, we have the feeling that it is a representation of a three-dimensional space. Then we explain how different is the case of a holographic representation.

7.5.1.1. *Appreciation of the Relief on a Picture, Depth of Field*

Figure 7.39 reminds us what happens when a picture is taken of some landscape. An objective images the landscape on the photographic plate. The focusing process on some point A consists of translating the plate until it coincides exactly with the image front plate (P') associated to the object front plane (P) of point A. A landscape is not usually planar. The focusing is only valid in this couple of planes, to a point such as B is associated a tiny spot bb', instead of a dimensionless point; if the spot is small enough, we still get the impression that bb' is a good image of B. The thickness, on the inside of which must be located the points that give acceptable images, is called the depth of field of the camera.

A picture, as well as a painting made by an artist, is nothing other than a two-dimensional object; the sensation of relief, or of perspective, is purely subjective and calls up our memory.

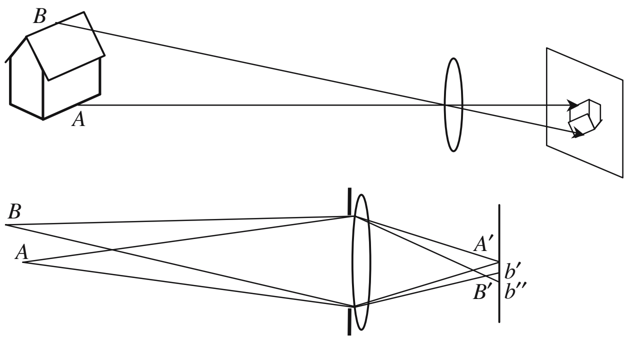


Figure 7.39. The camera is focused on the front wall of the house. The image of some point B at the top of the roof is slightly before the plate; if the tiny spot (bb'), generated by a pencil of rays coming from B, is small enough, the image of B is considered as acceptable.

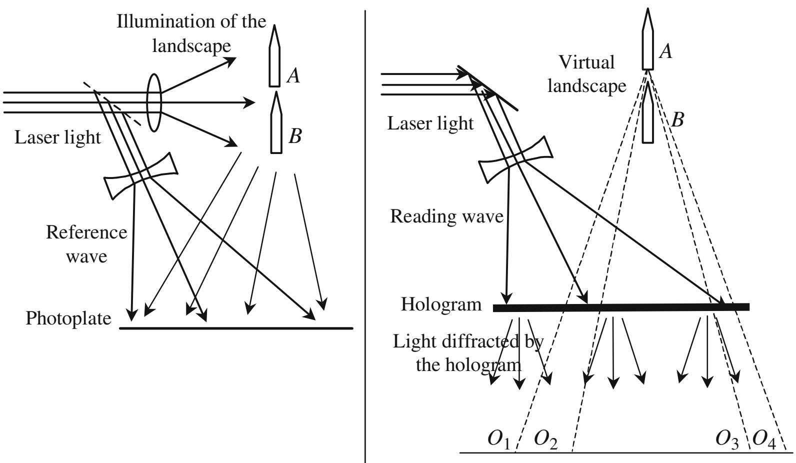
7.5.1.2. *Appreciation of the Relief on a Hologram*

Holography is a technique that memorizes the phase of a wave at the instant of exposure; in fact, it's the relative phase shifts from one point to the next that are stored. The phase storage is obtained thanks to the interference of the incident wave to be recorded with a reference wave that has a well-known and well-reproducible phase repartition. The photographic plate registers interference fringes, the blackening being more intense at the place of clear fringes, where both incident and reference waves are in phase. After impression and development the plate becomes a *hologram.*

The hologram is illuminated with a reading wave that has the same phase distribution as the reference wave that was used when recording the hologram; each point of the hologram becomes a source of coherent optical wavelets. It will be shown that the repartition of the complex amplitudes along the plane of the hologram is identical with the repartition that was created by the initial wave at the level of the photographic plate of Figure 7.40(a): the Huygens-Fresnel principle says that, along the line $O_1O_2O_3O_4$, the optical fields are identical, respectively:

- In Figure 7.40(a), in the absence of the plate.
- Beyond the hologram of Figure 7.40(b), the three dimensions of the landscape have actually been reconstituted.

It is interesting to say that, if the hologram is broken in several pieces, each piece contains all the information about the landscape and can be used as the entire initial hologram.



(a) Recording of the hologram.

(b) Reading of the hologram.

Figure 7.40. Recording and reading a hologram. An observer moves from O_1 to O_4. Between O_1 to O_2 the two objects are simultaneously visible; from O_2 to O_3, A is hidden by B and is not visible; from O_3 to O_4, A and B can be seen again.

7.5.2. *Principle of Holography*

Let us consider an initial wave A of complex amplitude $a_{1(x,y,z)}e^{j\varphi(x,y,z)}$. The plate where the hologram is to be recorded is taken as the xOy reference plane of coordinate ($z = 0$); in this plane the complex amplitude is simplified as

$$a_{1(x,y)}e^{j\phi(x,y)} = a_{1(x,y,0)}e^{j\varphi(x,y,0)}.$$

For the sake of simplicity, the reference wave will be a planar wave of amplitude a_2 and parallel to the plate, a_2 is supposed to be real. The amplitudes of vibration and intensity at point $P_{(x,y,0)}$ are given by

$$a_{(x,y)} = \left(a_{1(x,y)}e^{j\phi(x,y)} + a_2\right),$$
$$I_{(x,y)} = \left(a_{1(x,y)}e^{j\phi(x,y)} + a_2\right)\left(a^*_{1(x,y)}e^{-j\phi(x,y)} + a_2\right),$$
$$I_{(x,y)} = a_{1(x,y)}a^*_{1(x,y)} + a_2^2 + a_2\left(a_{1(x,y)}e^{jj\phi(x,y)} + a^*_{1(x,y)}e^{-j\phi(x,y)}\right).$$

We admit that the conditions of exposure and of development of the photographic plate have been chosen so that the variation of the darkening varies linearly with the light intensity. Under such conditions, when the hologram is illuminated by the reference wave, the repartition of the amplitude of the Huygens-Fresnel sources is described by the following formula:

$$\alpha_{(x,y)} + \beta_{(x,y)}e^{j\phi(x,y)} + \beta^*_{(x,y)}e^{-j\phi(x,y)}. \tag{7.21}$$

The second term of (7.21) reproduces exactly the complex amplitude repartition of the initial wave.

7.5.3. *Examples of Holograms*

We are now going to study several examples to illustrate how the initial wave can be regenerated.

7.5.3.1. *Hologram of a Planar Wave (see Figure* 7.41*)*

The reference and wave to be recorded are both planar and have equal amplitudes. The intersection of the plane of the hologram with the plane of the two wave vectors is taken as the Ox axis. The interference pattern of the two waves is extremely simple; the variation of transparency follows a sine law, versus x, and is proportional to the following expression:

$$\left(1+e^{-j\frac{2\pi}{\lambda}x\sin\theta}\right)\left(1+e^{+j\frac{2\pi}{\lambda}x\sin\theta}\right) = 2+e^{-j\frac{2\pi}{\lambda}x\sin\theta}+e^{+j\frac{2\pi}{\lambda}x\sin\theta}. \tag{7.22}$$

We recognize the three terms of the more general formula (7.16). The hologram is made parallel with equidistant lines (periodicity $a = \lambda/\sin\theta$). Illumi-

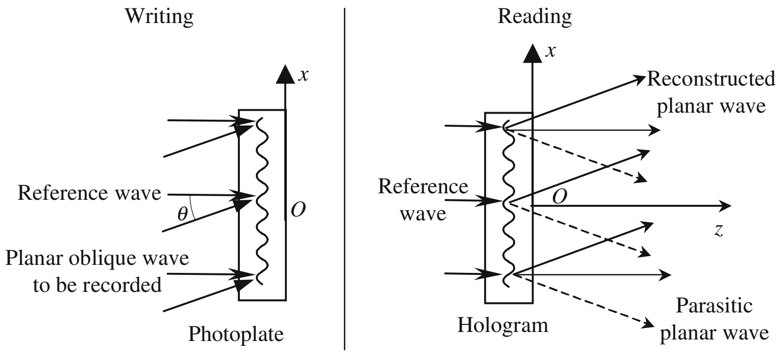


Figure 7.41. Writing and reading the hologram of a planar wave.

nated at normal incidence by the reference beam, the grating generates diffracted beams in the directions defined by the Bragg formula:

$$\sin\theta_{\text{diff}} = \sin\theta + p\frac{\lambda}{a} = (1+p)\sin\theta.$$

When the hologram is illuminated by the reference wave of Figure 7.41, its surface is covered with secondary radiating sources. The diffracted field is the Fourier transform of the amplitude repartition of formula (7.22), to each term of which is associated a parallel beam propagating in one of the following directions:

- $+\theta$, this is the reconstructed wave;
- Oz axis, corresponding to a partial transmission of the reference wave;
- $-\theta$, this is a parasitic wave.

The reason why there are only three orders of diffraction comes from the sinusoidal variation of the transparency; this effect is very analogous to the blaze effect in the case of a triangular profile.

The previous results are general: when reading a hologram, besides the reconstruction of the initial wave, parasitic waves are also obtained.

7.5.3.2. *Hologram of a Spherical Wave*

We are now going to holograph a diverging spherical wave that is supposed to have an aperture small enough to allow the same type of development that was used in Section 2.6.3. A parasitic spherical wave is also generated; an oblique reference wave efficiently separates the reconstructed wave from the parasitic wave.

7.5.3.3. *The Reference Writing and Reading Waves Are Both Orthogonal to the Hologram*

We refer to Figure 7.42; the reference wave and spherical wave have the same amplitudes a_0. ρ being the distance of a point P of the hologram to Oz and D the distance $D = A_0O$ of the center of the spherical wave to the plate, the spherical wave is described by $a_0 e^{-j\pi(\rho^2/\lambda D)}$. The amplitude $a(\rho)$ and the intensity $I(\rho)$ of the optical vibration at point P are, respectively, given by

$$a(\rho) = a_0^2\left(1 + e^{-j\frac{\pi\rho^2}{\lambda D}}\right),$$

$$I(\rho) = a_0^2\left(1 + e^{-j\frac{\pi\rho^2}{\lambda D}}\right)\left(1 + e^{+j\frac{\pi\rho^2}{\lambda D}}\right) = a_0^2\left(2 + e^{+j\frac{\pi\rho^2}{\lambda D}} + e^{-j\frac{\pi\rho^2}{\lambda D}}\right), \tag{7.23.a}$$

$$I(\rho) = 2a_0^2\left(1 + \cos\frac{\pi\rho^2}{\lambda D}\right). \tag{7.23.b}$$

The hologram, as is shown by formula (7.23.b), is made of alternate clear and dark rings, their radii increasing as the square root of the successive integers. It looks like a Fresnel zone plane (see Section 7.3), except that the transparency is not equal to zero or to one, but varies sinusoidally. A Fresnel zone plate has many focal points (see Figure 7.23); because of a kind of blaze effect, in the case of a sinusoidal variation of the transparency, only two focuses are illuminated, they correspond to the points A_0 and A_0' of Figure 7.42. The spherical wave centered at point A_0 is the reconstituted wave; it corresponds to the term $e^{j\pi\rho^2/\lambda D}$ of formula (7.18.a). The wave centered at point A_0 is a parasitic wave; it should be associated to the term $e^{-j\pi\rho^2/\lambda D}$.

An observer facing the hologram has the impression that he sees a real point A_0 that would be behind the hologram. If the hologram was registered

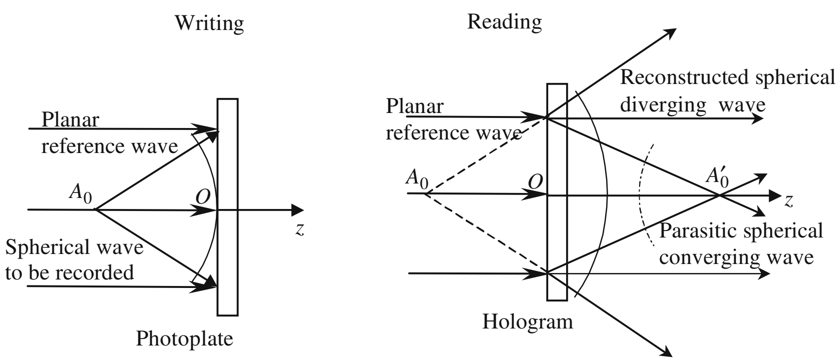


Figure 7.42. Hologram of a spherical wave with a reference wave orthogonal to the plate. Interference fringes are concentric circles.

with some extended object, the same treatment can be repeated for every point of the object. The parasitic wave also reproduces the initial object, which appears as a real object located in front of the hologram.

7.5.3.4. *The Planar Reading Reference Wave Is Oblique*

In the arrangement of Figure 7.42 the observer's eye also receives the reference wave and the parasitic wave, this drawback can be avoided by using an oblique reference wave, as is the case in Figure 7.43.

If the angle between the direction of the reference wave and the normal to the hologram, the intensity of the light vibration when recording the hologram, is equal to

$$I_{(x,y)} = a_0^2\left(e^{-j\frac{2\pi}{\lambda}x\sin\theta} + e^{-j\frac{\pi}{\lambda D}\left(x^2+y^2\right)}\right)\left(e^{j\frac{2\pi}{\lambda}x\sin\theta} + e^{j\frac{\pi}{\lambda D}\left(x^2+y^2\right)}\right),$$
$$I_{(x,y)} = a_0^2\left(2 + e^{-j\frac{2\pi}{\lambda}x\sin\theta}e^{j\frac{\pi}{\lambda D}\left(x^2+y^2\right)} + e^{j\frac{2\pi}{\lambda}x\sin\theta}e^{-j\frac{\pi}{\lambda D}\left(x^2+y^2\right)}\right). \tag{7.24.a}$$

When the hologram is illuminated by the reference wave, the repartition of the complex amplitudes, $a(\rho)$, is obtained by multiplying the amplitude of the reference wave by the transparency of the hologram, which is given by (7.24.a):

$$\begin{aligned} a(\rho) &= I_{(x,y)}e^{-j\frac{2\pi}{\lambda}x\sin\theta} \\ &= a_0^2\left(2e^{-j\frac{2\pi}{\lambda}x\sin\theta} + e^{-j\frac{\pi}{\lambda D}\left(x^2+y^2\right)} + e^{-j\frac{4\pi}{\lambda}x\sin\theta}e^{j\frac{\pi}{\lambda D}\left(x^2+y^2\right)}\right). \end{aligned} \tag{7.24.b}$$

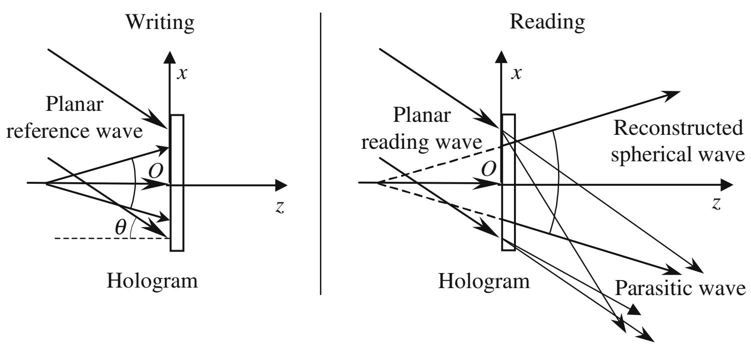


Figure 7.43. With an oblique reference wave, the reconstructed wave is well separated; the observer is not dazzled by the transmitted reference wave or by the parasitic wave.

The first term in equation (7.24.b) reconstitutes the reference wave that propagates in the direction θ. The second term reconstitutes a diverging spherical wave, which seems to have been emitted by point A_0. The last term is a spherical wave that converges at point A_0' in front of the hologram and in a direction making an angle 2θ with the normal.

7.5.3.5. *Volume Holography*

In the previous sections the holograms were planar, we are now going to use photographic emulsions that are no longer very thin. We thus consider the interference of a planar reference light wave (R) with some other wave (W) inside a volume of finite size. If the two waves are both planar, a family of parallel, alternately bright and dark, planar layers is printed inside the emulsion.

If the second wave is not a planar wave, it can always be considered as the superposition of planar waves with suitable wave vectors and intensities; each component of the planar wave decomposition of W, when interfering with the reference wave R, creates a family of parallel layers. The set of printed layers is called a *volume hologram.*

Let $\boldsymbol{k}_0$ and $\boldsymbol{k}_r$ be the respective wave vectors of W and R that will be supposed to have the same amplitude a_0; the interference pattern is described by

$$a(\boldsymbol{r}) = a_0\left(e^{-j\boldsymbol{k}_r\boldsymbol{r}} + e^{-j\boldsymbol{k}_0\boldsymbol{r}}\right) \quad \rightarrow \quad I(\boldsymbol{r}) = 2a_0^2[1+\cos((\boldsymbol{k_0}-\boldsymbol{k}_r)\boldsymbol{r})],$$

$$I(\boldsymbol{r}) = 2a_0^2[1+\cos \boldsymbol{k_G r}] = 2a_0^2\left[1+\cos\frac{2\pi}{\Lambda}\boldsymbol{u_G r}\right],$$

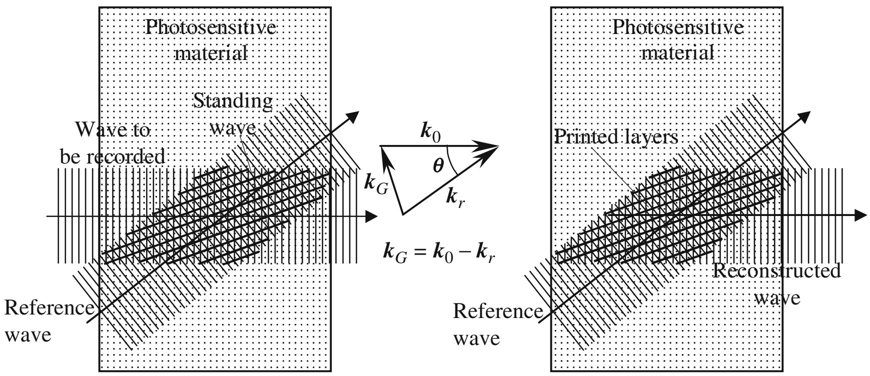


Figure 7.44. Recording and reading of a volume hologram. The Bragg reflection on the layers printed in the plate reconstructs the initial wave.

where

$\boldsymbol{k}_G = (\boldsymbol{k}_0 - \boldsymbol{k}_r)$ is called the grating vector;
$\boldsymbol{u}_G = \boldsymbol{k}_G/kG$ is a unit vector parallel to one of the bisectors of the angle $(\boldsymbol{k}_0, \boldsymbol{k}_r)$; and
$\Lambda = \lambda/(\sin\theta/2)$ is the periodicity of the printed grating.

We now examine what happens when the reference beam illuminates the hologram. The situation can be compared to the diffraction by the reticular planes of a crystal lattice; the parallel planes of a grating reflect the light if a kind of Bragg condition is satisfied. Let λ' be the wavelength of the reference beam and let α be the angle between the reference wave vector and the layers of the hologram:

$$\sin\alpha = \frac{\lambda'}{2\Lambda}. \tag{7.25}$$

The initial Gabor proposition was to make a hologram with an electron beam, and then to read it with an optical beam; λ would have been the wavelength of the De Broglie wave associated to the electrons and λ' an optical wavelength. For many reasons no realization has ever been obtained. Let us come back to the case where the writing and reading reference waves are optical waves of the same frequency (see Figure 7.44): $\lambda' = \lambda$, which implies that $\alpha = \theta/2$. The set of bright and dark layers plays the same role as the reticular planes in the diffraction by a crystal: the reading beam is selectively reflected in the direction of the second bisector of the angle $(\boldsymbol{k}_0, \boldsymbol{k}_r)$, and reconstructs the initial planar beam.

Because information is stored in a three-dimensional space, volume holography is, in principle, more powerful. Several, eventually many, different holograms can be registered inside the same volume, either changing the direction or the frequency of the reference wave.

Color holograms: A volume hologram can be read by a white reference light beam, Bragg reflection occurs only for that part of the spectrum which coincides with the color that was used for recording. Using the principle of the trichromatic system (see Section 7.8), "colored holograms" can be made; three holograms are registered in the same plate with three different wavelengths, when the plate is illuminated with a white light, the initial object appears in color and in three dimensions.

7.5.3.6. *Interference Holography*

Very small deformations (μm) of large size objects can be revealed thanks to holography, for example, the stress deformations of mechanical structures such as a bridge, a shaft, or the body of a car. Two holograms are consecutively recorded in the same plate, the object being successively in the two situations to be compared. The reference beam then simultaneously illuminates the two holograms and an objective makes a projection on a screen. The image reproduces the object, but it is striped with interference fringes, the deformation remaining constant along a given fringe.

The method can work under real time conditions. A hologram of the unstressed object is first registered and the virtual image that is obtained when illuminated by the reference beam is put into exact coincidence with the object that is also illuminated. On the screen the fringes due to deformation appear as if they were painted on the object.

7.6. Diffraction and Image Processing

The Fourier transform plays a preeminent role in diffraction problems and in the formation of images. Starting from some object, described as a repartition of amplitudes $f(x, y)$ in a front plane, a lens L_1 elaborates the Fourier transform $F(X, Y)$ of $f(x, y)$, which is available in its image focal plane. A second lens L_2 then makes a second Fourier transform $f'(x', y')$ of $F(X, Y)$. $f'(x', y')$, which is available in the image focal plane of L_2, is the image of the initial object and, except for a multiplication by a "magnification coefficient," is similar to $f(x, y)$. The intermediate Fourier transform $F(X, Y)$ can be modified by placing a slide with a controlled transparency on the focal plane (Π) of L_1, (Π) is also called the Fourier plane of the arrangement.

7.6.1. *Fourier Plane*

The arrangement shown in Figure 7.45, which is analogous to that of Figure 7.11, will allow us to go deeper into the insight of the mechanism of the formation of an image by a lens. An object is placed in the focal plane (π) of a lens (focal length f). We consider an elementary area $dx\, dy$ around some point $M(x, y)$, a spherical wavelet emitted by this point is collimated by the lens into a planar wave propagating in the direction defined by the coefficients

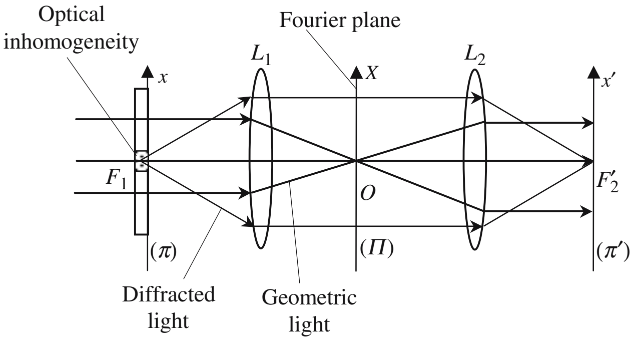


Figure 7.45. The image focal plane of L_1 coincides with the object focal plane of L_1. (π) and (π') are the antiprincipal planes of the centered system.

$u = x/f$ and y/f; its amplitude is proportional to $dx\ dy$ and is written as $f(x, y)\ dx\ dy$, $f(x, y)$ varies from one point of the object to another and is a complex number. Let us evaluate the complex amplitude $F(X, Y)$ of the optical vibration at a point $P(X, Y)$ of the second focal plane (Π) of the lens. $F(X, Y)$ results from the superposition of all the planar waves coming from the different points of the object. If M doesn't belong to the lens axis Oz, the planar wave coming from M is not parallel to Oz; the corresponding amplitude repartition is described by

$$e^{-j\frac{2\pi}{\lambda}\frac{xX+yY}{f}} f(x,y)\ dx\,dy.$$

The total amplitude at point $P(X, Y)$ is given by

$$F(X,Y)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)e^{-j\frac{2\pi}{\lambda}\frac{xX+Yy}{f}}\ dx\,dy. \tag{7.26}$$

We have considered that the size of the lens is large enough so that the diffraction effects by the corresponding aperture are negligible, and so, the integrals are extended to infinity. Formula (7.21) shows that the amplitude distribution in the plane (Π) is the Fourier transform of the function $f(x, y)$,

$$F(X,Y)=\mathrm{FT}[f(x,y)].$$

The plane (Π) is then considered as an object playing, for lens L_2, the same role as (π) for L_1. The amplitude repartition, $f'(x', y')$, in the focal plane (π') of the second lens is the Fourier transform of $F(X, Y)$:

$$F'(x',y')=\mathrm{FT}[F(X,Y)]=\mathrm{FT}\{\mathrm{FT}[f(x,y)]\} \quad\rightarrow\quad f'(x',y')=f(-x,-y).$$

The intermediate plane (Π) is called the *Fourier plane* of the arrangement.

The previous method is nothing other than a sophisticated method for establishing that the two planes (π) and (π') are the two conjugate planes of the centered system for which the magnification is equal to −1; such planes are also called "antiprincipal planes," by analogy with the principal planes of Section 3.5.2. The method is however very interesting, since it opens the way to very powerful methods of image processing. If a plate, with a variable transparency $t(X, Y)$, is placed on the Fourier plane, the planar wave decomposition of the function $f(x, y)$ can be modified at will, as well as the appearance of the image that the second wave will reconstruct, by the inverse Fourier transform.

7.6.2. *Image Processing*

We are now going to examine different examples where optical signals are processed by modification of their planar wave decompositions. To each component of the planar wave decomposition is associated a point of the Fourier plane; if a tiny opaque obstacle is placed at this point, the corresponding component will be filtered out.

7.6.2.1. *Strioscopic Arrangement—Phase Contrast Microscopy*

We refer again to Figure 7.45, the optical inhomogeneity concerns the phase, the refractive index being slightly different from the surrounding index, this is, for example, the case of bacteria floating in water of a microscopic preparation. The amplitude repartition in the plane (π) is described by

$$a(\rho) = a_0 e^{-j\varphi(\rho)} \quad \text{with} \quad \varphi(\rho) = \frac{2\pi}{\lambda}[1 + \Delta n \delta(\rho)],$$

where ρ is the distance to the axis and $\delta(\rho)$ is a Dirac pulse. The index difference Δn between the bacteria and water is very small. Changing the phase origin and considering that $e\Delta n$ is small, as compared to the wavelength, we can write

$$\text{Plane } (\pi): \quad a(\rho) = a_0 e^{-j\frac{2\pi\Delta ne}{\lambda}\delta(\rho)} \approx a_0\left[1 - j\frac{2\pi\Delta ne}{\lambda}\delta(\rho)\right].$$

Since the Fourier transform of a Dirac pulse is a constant and vice versa, the amplitude repartition in the Fourier plane is equal to

$$\text{Plane } (\Pi): \quad A(\rho) = a_0\left[\delta(\rho) - j\frac{2\pi\Delta ne}{\lambda}\right]. \tag{7.27.a}$$

After a second Fourier transform, made by the second lens, we finally obtain the amplitude and intensity repartitions in the image plane,

$$\text{Plane } (\pi): \quad a'(\rho) = a_0\left[1 - j\frac{2\pi\Delta ne}{\lambda}\delta(\rho)\right], \tag{7.27.b}$$

$$a'a'^* = a_0^2\left[1 + \left(\frac{2\pi\Delta ne}{\lambda}\right)^2 \delta(\rho)\right]. \tag{7.27.c}$$

According to formula (7.27.c), plane (π') is homogeneously illuminated with, in the middle, a clearer spot corresponding to the presence of the bacteria. Figure 7.45 gives an interpretation of what happens, in terms of optical light rays. The planar incident wave is focused at the center of the Fourier plane and gives the term $a_0\delta(\rho)$ in formula (7.27), it then diverges and is collimated by the second lens to give a constant illumination on (π'). The second term, $-ja_0(2\pi\Delta ne/\lambda)\delta(\rho)$ in (7.27), should be associated to the light that is diffracted by the bacteria.

It's not difficult to take into account the fact that the lenses have a finite size and that the diameter of the bacteria is not equal to zero. Let R and r be the respective radii of the lens and of the bacteria. In the Fourier plane (Π) the amplitude repartition of the "geometric light" on the one hand, and of the "diffracted light" on the other are described by Airy functions ($J(\rho)/\rho$). The spot inside which the geometric light is concentrated has a diameter equal to $1.22\lambda f/R$ and goes to zero if $R \to \infty$. On the contrary the diffracted light spreads over a large area ($1.22\lambda f/r$).

Strioscopic Arrangement

In the case of the arrangement of Figure 7.45, the variation of intensity at the place of the image of the bacteria is very small, $\approx (2\pi\Delta ne/\lambda)^2$, and will hardly be perceived. In the spectroscopic arrangement the "geometric light" is removed by covering the focal point O with a tiny opaque stop. The plane (π') remains completely dark in the absence of bacteria, and the image of the bacteria is more easily revealed.

Phase Contrast

The strioscopic arrangement is not very luminous, and its ability to exhibit small details is quite limited by the unavoidable existence of parasitic stray light. In the phase contrast method, the tiny stop is replaced by a transparent plate, which has the same size, and a thickness such that the optical path of the geometric light is increased by an odd number of half-wavelengths, with regard to the diffracted light. Formulas (7.27.a) and (7.27.b), giving respectively, the amplitude repartitions in the Fourier plane and in the image plane, are now replaced by

$$A(\rho) = a_0\left[e^{\pm j\pi/2}\delta(\rho) - j\frac{2\pi\Delta ne}{\lambda}\right] = ja_0\left[\pm\delta(\rho) - \frac{2\pi\Delta ne}{\lambda}\right], \tag{7.28.a}$$

$$a'(\rho) = ja_0\left[\pm 1 - \frac{2\pi\Delta ne}{\lambda}\delta(\rho)\right]. \tag{7.28.b}$$

Neglecting second-order terms, the intensity repartition in the image plane is

$$E(\rho) = a'_{(\rho)}a'^{*}_{(\rho)} \cong a_0^2\left[1 \pm \frac{4\pi\Delta ne}{\lambda}\delta_{(\rho)}\right]. \tag{7.29}$$

The intensity variation is larger, since it's now of the order of $(2\pi\Delta ne/\lambda)$. According to the fact that the geometric and diffracted beams are in phase, or in opposition, the variation is positive (clear image on a dark background) or negative (dark image on a clear background).

The phase contrast method is still handicapped by the existence of stray light. It's finally the aptitude of the observer to perceive a small intensity contrast that fixes the smallest size of detectable details. Using a plate, which at the same time, changes the phase and partially absorbs the geometric light, can accumulate the advantage of strioscopy and of phase contrast. Let us suppose that the phase difference is still $\pm\pi/2$, while the intensity is multiplied by a coefficient α smaller than unity, (7.24) then becomes

$$E(\rho) \cong a_0^2\left[\alpha^2 \pm 4\alpha\frac{\pi\Delta ne}{\lambda}\delta(\rho)\right]. \tag{7.30}$$

The geometric light is more attenuated than the diffracted light and the contrast more favorable. With a proper choice of α, a difference of optical length as low as a few nanometers can be perceived. Phase contrast microscopy is much appreciated by biologists who generally observe transparent objects immersed inside a liquid having almost the same index of refraction. A typical example is the observation of living cells; phase contrast allows the observation in vivo, without using a coloring dye that is selectively fixed by a given component of the cell.

7.6.2.2. *Optical Image Processing*

Spatial Filtering

The planar waves that are focused near the center of the Fourier plane are said to be *low-frequency* spatial components. The spatial frequency is so high that the wave vector makes a greater angle with the axis, the corresponding focus being farther from the center. A tiny opaque spot, suitably placed in the Fourier plane, will filter out a given component.

A tiny opaque stop, at the center of the Fourier plane, is a *high-frequency* filter, which blocks the low-frequency component and especially the zero frequency; this is the case of strioscopy. On the contrary, a small hole is a *low-frequency* filter, transmitting only the low frequencies. With a pair of scissors and a sheet of black paper, various filters can easily be made. A horizontal slit, with broadness D, will only transmit the so-called vertical frequencies, with a bandwidth of the order of $\Delta k = D/2\lambda f$. In the same way a vertical slit only transmits the horizontal frequencies.

A function that is rapidly varying versus x and/or y, has many high-frequency components. If a high-frequency filter is used the areas, inside of which the intensity variations are smooth, will appear dark, while the zone of more rapid variation will be more brightly illuminated (see Figure 7.48).

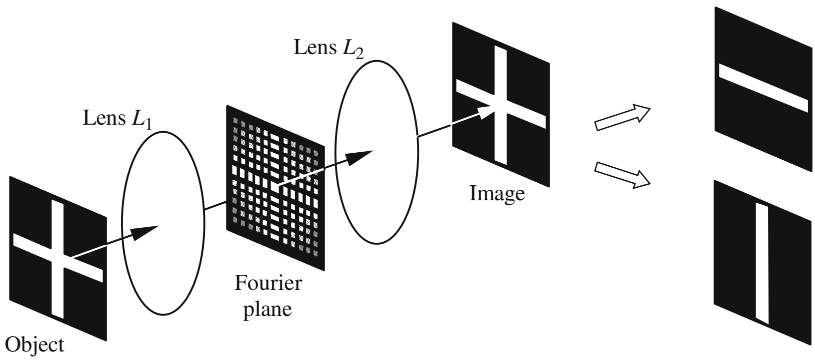


Figure 7.46. Formation of the image of a cross. At first with no filtering at all, and then with filters that select either the “horizontal” or “vertical” frequencies.

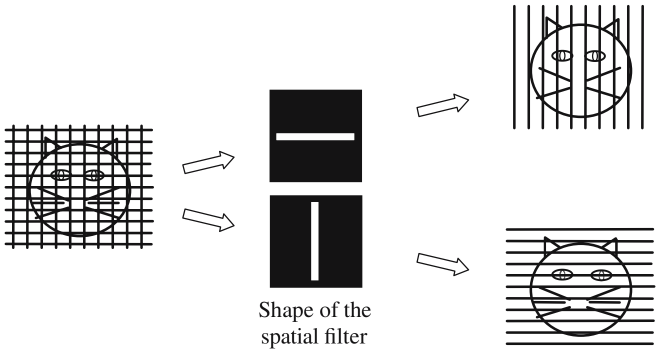


Figure 7.47. Filtering of the image of a cat seen through a square mesh grid.

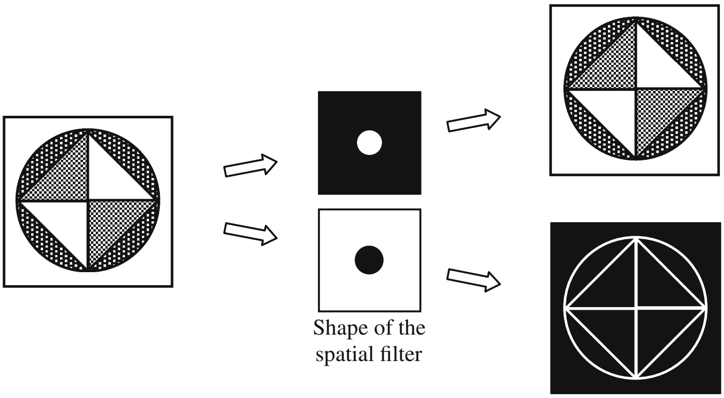


Figure 7.48. With a high-frequency filter, only the zones where the gradient of intensity is important will be illuminated. The image contour will appear as bright lines drawn on a dark background.

Analog Calculation of a Two-Dimensional Fourier Transform

The calculation of a two-dimensional Fourier transform needs a powerful computer. A simple lens achieves this operation in an analog way and can, in principle, do it in real time. This is the reason why, between 1960 and 1980, an important research effort was devoted to optically computing the Fourier transform. A drawback of this method is due to the noise that artificially introduces high-frequency components. Hopefully, or not, at the same time, the important development of microelectronics has drastically increased the storage capacity of computers and simultaneously decreased computation times, Fourier transforms are exclusively made by computers.

Computer-Assisted Image Processing

The continuous variation of the light intensity along some image can easily be sampled, using small adjacent surfaces that are called *pixels*. Each pixel is characterized by one real number (or by two real numbers if the amplitude is complex). These numbers, stored in good order in the memory of a computer, constitute a numerical image. Once the image has been replaced by a set of numbers, it's easy to imagine many operations, including the Fourier transform.

A first Fourier transform will give the repartition of complex amplitudes $A(X, Y)$ in the Fourier plane. Spatial filtering becomes very easy; it is just a multiplication by a number $\alpha(X, Y)$. A new set of numbers is obtained from which, thanks to a cathode ray tube or a printer, the new image can be visualized.

A Fourier transform followed by a spatial filtering was the first operation to be thought of. Since then, engineers have invented many other mathematical manipulations which, on one hand, are more adequately fitted to computer calculations and, on the other hand, give images of better quality. Such methods have considerably improved the numerical images that are transmitted from telescopes installed in satellites rotating around the Earth. The numerization of the image is done with a matrix of photodiode places in the focal plane of the telescope. Radio or microwaves make the transmission of the set of numbers back to Earth. This can be images of the Earth or of the stars and planets. In the last case, the main interest comes from the fact that the perturbations due to crossing the atmosphere are avoided. A difficulty with these indirect methods comes from the fact that an image is always obtained, the problem is to be sure that it has something to do with the initial object.

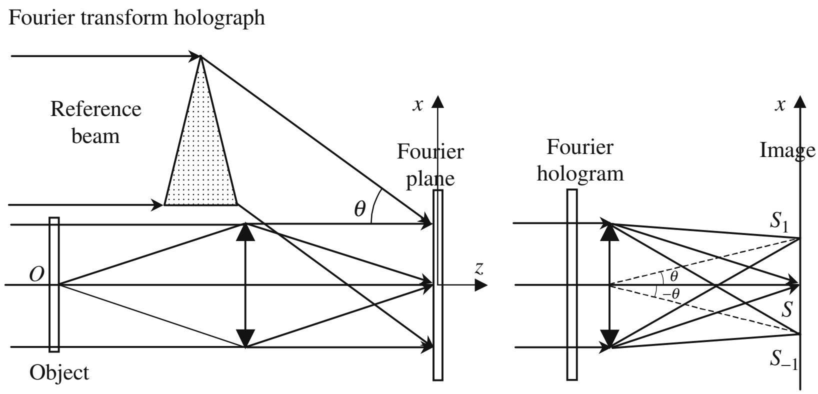


Figure 7.49. Recording and reading a Fourier hologram.

Fourier Transform Holography

Thanks to holography it is possible to store the repartition of complex amplitudes existing in a given plane. The arrangement of Figure 7.49 shows how to store the amplitude repartition of the Fourier plane; the hologram is then called a Fourier hologram. If the object is a point, the wave after the lens is a planar wave propagating parallel to Oz; the interference with the oblique reference wave gives horizontal fringes parallel to Oy: the transparency of the hologram is a sine function of x. A parallel beam illuminating the hologram is diffracted in three directions that are focused by the lens at S, S_1, and S_{-1}, the two last points can be considered as the images of the initial object.