

## THE FRESNEL DIFFRACTION: A STORY OF LIGHT AND DARKNESS

C. Aime<sup>1</sup>, É. Aristidi<sup>1</sup> and Y. Rabbia<sup>1</sup>

**Abstract.** In a first part of the paper we give a simple introduction to the free space propagation of light at the level of a Master degree in Physics. The presentation promotes linear filtering aspects at the expense of fundamental physics. Following the Huygens-Fresnel approach, the propagation of the wave writes as a convolution relationship, the impulse response being a quadratic phase factor. We give the corresponding filter in the Fourier plane. As an illustration, we describe the propagation of wave with a spatial sinusoidal amplitude, introduce lenses as quadratic phase transmissions, discuss their Fourier transform properties and give some properties of Soret screens. Classical diffractions of rectangular diaphragms are also given here. In a second part of the paper, the presentation turns into the use of external occulters in coronagraphy for the detection of exoplanets and the study of the solar corona. Making use of Lommel series expansions, we obtain the analytical expression for the diffraction of a circular opaque screen, giving thereby the complete formalism for the Arago-Poisson spot. We include there shaped occulters. The paper ends up with a brief application to incoherent imaging in astronomy.

### 1 Historical introduction

The question whether the light is a wave or a particle goes back to the seventeenth century during which the mechanical corpuscular theory of Newton took precedence over the wave theory of Huygens. Newton's particle theory, which explained most of the observations at that time, stood as the undisputed model for more than a century. This is not surprising since it was not easy to observe natural phenomena resulting from the wave nature of light. At that time light sources like the Sun or a candle light were used. They are incoherent extended sources while a coherent source is needed to see interference phenomenas, unquestionable signatures of the wave nature of light.

---

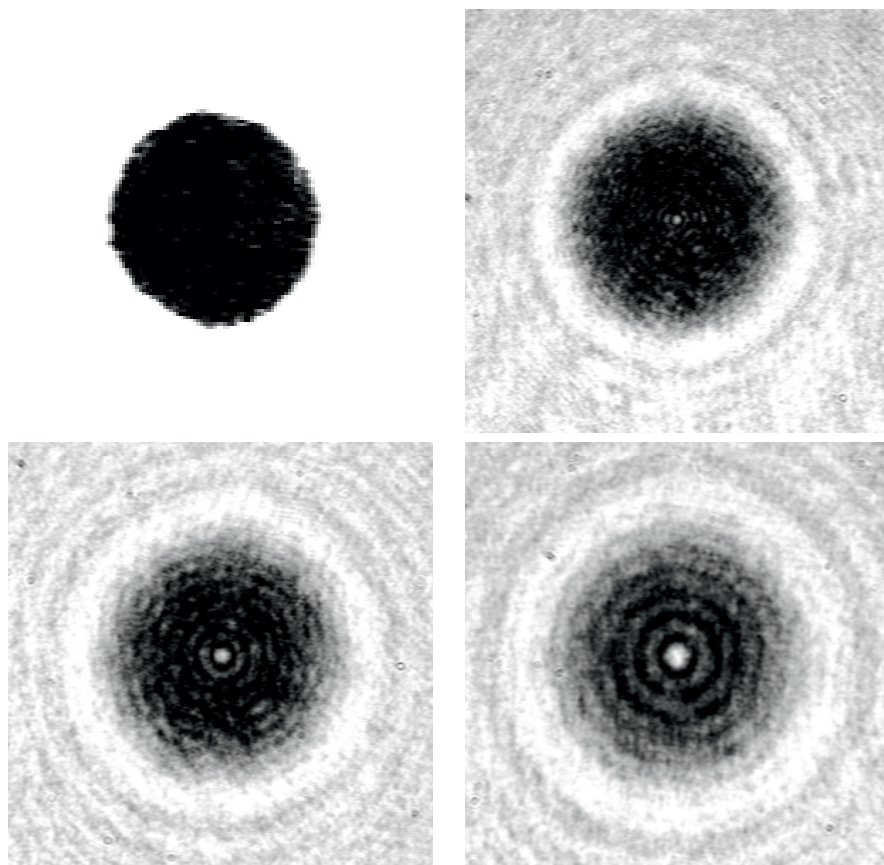
<sup>1</sup> Laboratoire Lagrange, Université de Nice Sophia-Antipolis, Centre National de la Recherche Scientifique, Observatoire de la Côte d'Azur, Parc Valrose, 06108 Nice, France

The starting point of the wave theory is undoubtedly the historical double-slits experiment of Young in 1801. The two slits were illuminated by a single slit exposed to sunlight, thin enough to produce the necessary spatially coherent light. Young could observe for the first time the interference fringes in the overlap of the light beams diffracted by the double slits, demonstrating thereby the wave nature of light against Newton's particle theory. Indeed, the darkness in the fringes cannot be explained by the sum of two particles but easily interpreted by vibrations out of phase. This argument was very strong. Nevertheless Einstein had to struggle in his turn to have the concept of photons accepted by the scientific community a century later. In astronomy, we can use the simplified semiclassical theory of photodetection, in which the light propagates as a wave and is detected as a particle (Goodman 1985).

In Young's time, Newton's prestige was so important that the wave nature of light was not at all widely accepted by the scientific community. About fifteen years later, Fresnel worked on the same problematics, at the beginning without being aware of Young's work. Starting from the Huygens approach, Fresnel proposed a mathematical model for the propagation of light. He competed in a contest proposed by the Academy of Sciences. The subject was of a quest for a mathematical description of diffraction phenomena appearing in the shade of opaque screens. Poisson, a jury's member, argued that according to Fresnel's theory, a bright spot should appear at the center of a circular object's shadow, the intensity of the spot being equal to that of the undisturbed wavefront. The experiment was soon realized by Arago, another jury's member, who indeed brilliantly confirmed Fresnel's theory. This bright spot is now called after Poisson, Arago or Fresnel.

Arago's milestone experience was reproduced during the CNRS school of June 2012, using learning material for students in Physics at the University of Nice Sophia – Antipolis. The results are given in Figure 1. A laser and a beam expander were used to produce a coherent plane wave. The occulter was a transparent slide with an opaque disk of diameter 1.5 mm, while Arago used a metallic disk of diameter 2 mm glued on a glass plate. Images obtained in planes at a distance of  $z = 150, 280$  and  $320$  mm away from the screen of observation are given in the figure. The Arago bright spot clearly appears and remains present whatever the distance  $z$  is. The concentric circular rings are not as neat as expected, because of the poor quality of the plate and of the occulting disk, a difficulty already noted by Fresnel (see for example in de Senarmont *et al.* 1866). We give the mathematical expression for the Fresnel diffraction in the last section of this paper.

The presentation we propose here is a short introduction to the relations of free space propagation of light, or Fresnel diffraction. It does not aim to be a formal course or a tutorial in optics, and remains in the theme of the school, for an interdisciplinary audience of astronomers and signal processing scientists. We restrict our presentation to the scalar theory of diffraction in the case of paraxial optics, thus leaving aside much of the work of Fresnel on polarization. We show that the propagation of light can be simply presented with the formalism of linear filtering. The reader who wishes a more academic presentation can refer to books of Goodman (2005) and Born & Wolf (2006).



**Fig. 1.** Reproduction of Arago's experience performed during the CNRS school. *Top left:* the occluder (diameter: 1.5 mm), and its Fresnel diffraction figures at the distances of 150 mm (*top right*), 280 mm (*bottom left*) and 320 mm (*bottom right*) from the screen of observation.

The paper is organized as follows. We establish the basic relations for the free space propagation in Section 2. An illustration for the propagation of a sinusoidal pattern is given in Section 3. Fourier properties of lenses are described in Section 4. Section 5 is devoted to the study of shadows produced by external occluders with application to coronagraphy. Section 6 gives a brief application to incoherent imaging in astronomy.

## 2 Basic relations for free space propagation, a simplified approach

We consider a point source  $S$  emitting a monochromatic wave of period  $T$ , and denote  $A_S(t) = A \exp(-2i\pi t/T)$  its complex amplitude. In a very simplified model

where the light propagates along a ray at the velocity  $v = c/n$  ( $n$  is the refractive index), the vibration at a point P located at a distance  $s$  from the source is:

$$A_P(s, t) = A \exp\left(-2i\pi\frac{(t - s/v)}{T}\right) = A \exp\left(-2i\pi\frac{t}{T}\right) \exp\left(2i\pi\frac{ns}{\lambda}\right) \quad (2.1)$$

where  $\lambda = cT$  is the wavelength of the light, the quantity  $ns$  is the optical path length introduced by Fermat, a contemporary of Huygens. The time dependent factor  $\exp(-2i\pi t/T)$ , common to all amplitudes, is omitted later on in the presentation. For the sake of simplicity, we moreover assume a propagation in the vacuum with  $n = 1$ .

In the Huygens-Fresnel model, the propagation occurs in a different way. First of all, instead of rays, wavelets and wavefronts are considered. A simple model for a wavefront is the locus of points having the same phase, *i.e.* where all rays originating from a coherent source have arrived at a given time. A spherical wavefront becomes a plane wavefront for a far away point source. According to Malus theorem, wavefronts and rays are orthogonal. The Huygens-Fresnel principle states that each point of any wavefront irradiates an elementary spherical wavelet, and that these secondary waves combine together to form the wavefront at any subsequent time.

We assume that all waves propagate in the  $z$  positive direction in a  $\{x, y, z\}$  coordinate system. Their complex amplitudes are described in parallel transverse planes  $\{x, y\}$ , for different  $z$  values. If we denote  $A_0(x, y)$  the complex amplitude of a wave in the plane  $z = 0$ , its expression  $A_z(x, y)$  at the distance  $z$  may be obtained by one of the following equivalent equations:

$$\begin{aligned} A_z(x, y) &= A_0(x, y) * \frac{1}{i\lambda z} \exp\left(\frac{i\pi(x^2 + y^2)}{\lambda z}\right) \\ A_z(x, y) &= \mathfrak{F}^{-1}\left[\hat{A}_0(u, v) \exp(-i\pi\lambda z(u^2 + v^2))\right] \end{aligned} \quad (2.2)$$

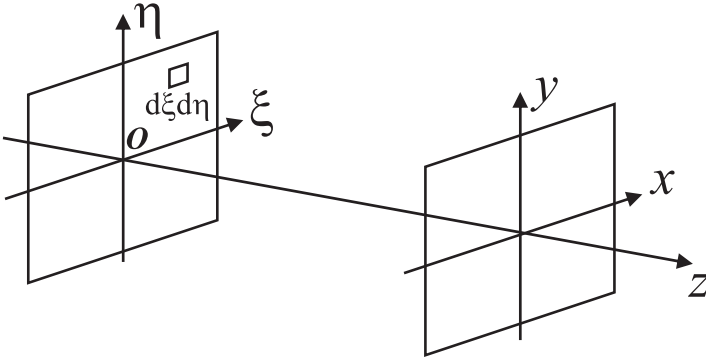
where  $\lambda$  is the wavelength of the light, the symbol  $*$  stands for the 2D convolution.  $\hat{A}_0(u, v)$  is the 2D Fourier transform of  $A_0(x, y)$  for the conjugate variables  $(u, v)$  (spatial frequencies), defined as

$$\hat{A}_0(u, v) = \iint A_0(x, y) \exp(-2i\pi(ux + vy)) dx dy. \quad (2.3)$$

The symbol  $\mathfrak{F}^{-1}$  denotes the two dimensional inverse Fourier transform. It is interesting to note the quantity  $\sqrt{\lambda z}$ , playing the role of a size factor in Equation (2.2), as we explain in the next section in which these relationships are established and their consequences analyzed.

## 2.1 The fundamental relation of convolution for complex amplitudes

The model proposed by Huygens appears as a forerunner of the convolution in Physics. In the plane  $z$ , the amplitude  $A_z(x, y)$  is the result of the addition of



**Fig. 2.** Notations for the free space propagation between the plane at  $z = 0$  of transverse coordinates  $\xi$  and  $\eta$  and the plane at the distance  $z$  of transverse coordinates  $x$  and  $y$ .

the elementary wavelets coming from all points of the plane located at  $z = 0$ . To build the relations between the two waves in the planes  $z = 0$  and  $z$ , we will have to consider the coordinates of the points of these planes. To make the notations simpler, we substitute  $\xi$  and  $\eta$  to  $x$  and  $y$  in the plane  $z = 0$ .

Let us first consider the point-like source at the origin  $O$  of coordinates  $\xi = \eta = 0$ , and the surrounding elementary wavefront of surface  $\sigma = d\xi d\eta$ . The wavelet emitted by  $O$  is an elementary spherical wave. After a propagation over the distance  $s$ , the amplitude of this spherical wave can be written as  $(\alpha/s) \sigma A_0(0, 0) \times \exp(2i\pi s/\lambda)$ , where  $\alpha$  is a coefficient to be determined. The factor  $1/s$  is required to conserve the energy.

Now we start deriving the usual simplified expression for this elementary wavelet in the plane  $\{x, y\}$  at the distance  $z$  from  $O$ . Under the assumption of paraxial optics, *i.e.*  $x$  and  $y \ll z$ , the distance  $s$  is approximated by

$$s = (x^2 + y^2 + z^2)^{1/2} \simeq z + (x^2 + y^2)/2z. \quad (2.4)$$

The elementary wavelet emitted from a small region  $\sigma = d\xi d\eta$  around  $O$  (see Fig. 2) and received in the plane  $z$  can be written:

$$A_0(0, 0) \sigma \times \frac{\alpha}{s} \exp\left(2i\pi \frac{s}{\lambda}\right) \simeq A_0(0, 0) d\xi d\eta \times \exp\left(2i\pi \frac{z}{\lambda}\right) \frac{\alpha}{z} \exp\left(i\pi \frac{x^2 + y^2}{\lambda z}\right). \quad (2.5)$$

The approximation  $1/s \simeq 1/z$  can be used, when  $s$  works as a factor for the whole amplitude, since this latter is not sensitive to a small variation of  $s$ . On the contrary the two terms in the Taylor expansion of Equation (2.4) must be kept in the argument of the complex exponential, since it expresses a phase and is very sensitive to a small variation of  $s$ . For example a variation of  $s$  as faint as  $\lambda$  induces a phase variation of  $2\pi$ .

For a point source P at the position  $(\xi, \eta)$ , the response is:

$$dA_z(x, y) = A_0(\xi, \eta) \exp\left(2i\pi\frac{z}{\lambda}\right) \frac{\alpha}{z} \exp\left(i\pi\frac{(x-\xi)^2 + (y-\eta)^2}{\lambda z}\right) d\xi d\eta. \quad (2.6)$$

According to the Huygens-Fresnel principle, we sum the wave amplitudes for all point like sources coming from the plane at  $z = 0$  to obtain the amplitude in  $z$ :

$$\begin{aligned} A_z(x, y) &= \exp\left(2i\pi\frac{z}{\lambda}\right) \iint A_0(\xi, \eta) \frac{\alpha}{z} \exp\left(i\pi\frac{(x-\xi)^2 + (y-\eta)^2}{\lambda z}\right) d\xi d\eta \\ &= \exp\left(2i\pi\frac{z}{\lambda}\right) A_0(x, y) * \frac{\alpha}{z} \exp\left(i\pi\frac{x^2 + y^2}{\lambda z}\right). \end{aligned} \quad (2.7)$$

Equation (2.7) results in the convolution of the amplitude at  $z = 0$  with the amplitude of a spherical wave. The factor  $\exp(2i\pi z/\lambda)$  corresponds to the phase shift induced by the propagation over the distance  $z$ , and will be in general omitted as not being a function of  $x$  and  $y$ . The coefficient  $\alpha$  is given by the complete theory of diffraction. We can derive it considering the propagation of a plane wave of unit amplitude  $A = 1$ . Whatever the distance  $z$  we must recover a plane wave. So we have:

$$1 * \frac{\alpha}{z} \exp\left(i\pi\frac{x^2 + y^2}{\lambda z}\right) = 1 \quad (2.8)$$

which leads to the value  $\alpha = (i\lambda)^{-1}$ , as the result of the Fresnel integral. The final expression is then:

$$A_z(x, y) = A_0(x, y) * \frac{1}{i\lambda z} \exp\left(i\pi\frac{x^2 + y^2}{\lambda z}\right) = A_0(x, y) * D_z(x, y). \quad (2.9)$$

The function  $D_z(x, y)$  behaves as the point-spread function (PSF) for the amplitudes. It is separable in  $x$  and  $y$ :

$$D_z(x, y) = D_z^0(x)D_z^0(y) = \frac{1}{\sqrt{i\lambda z}} \exp i\pi\frac{x^2}{\lambda z} \cdot \frac{1}{\sqrt{i\lambda z}} \exp i\pi\frac{y^2}{\lambda z} \quad (2.10)$$

where  $D_z^0(x)$  is normalized in the sense that  $\int D_z^0(x)dx = 1$ . It is important to note that  $D_z(x, y)$  is a complex function, essentially a quadratic phase factor, but for the normalizing value  $i\lambda z$ .

### 2.1.1 The Fresnel transform

Another form for the equation of free space propagation of the light can be obtained by developing Equation (2.9) as follows

$$\begin{aligned} A_z(x, y) &= \frac{1}{i\lambda z} \exp\left(i\pi\frac{x^2 + y^2}{\lambda z}\right) \times \\ &\quad \iint \left\{ A_0(\xi, \eta) \exp\left(i\pi\frac{\xi^2 + \eta^2}{\lambda z}\right) \right\} \exp\left(-2i\pi\left(\xi\frac{x}{\lambda z} + \eta\frac{y}{\lambda z}\right)\right) d\xi d\eta. \end{aligned} \quad (2.11)$$

The integral clearly describes the Fourier transform of the function between brackets for the spatial frequencies  $x/\lambda z$  and  $y/\lambda z$ . It is usually noted as:

$$A_z(x, y) = \frac{1}{i\lambda z} \exp\left(i\pi \frac{x^2 + y^2}{\lambda z}\right) \mathcal{F}_z \left[ A_0(x, y) \exp\left(i\pi \frac{x^2 + y^2}{\lambda z}\right) \right]. \quad (2.12)$$

Following Nazarathy & Shamir (1980), it is worth noting that the symbol  $\mathcal{F}_z$  can be interpreted as an operator that applies on the function itself, keeping the original variables  $x$  and  $y$ , followed by a scaling that transform  $x$  and  $y$  into  $x/\lambda z$  and  $y/\lambda z$ . Although it may be of interest, the operator approach implies the establishment of a complete algebra, and does not present, at least for the authors of this note, a decisive advantage for most problems encountered in optics.

The Fresnel transform and the convolution relationship are strictly equivalent, but when multiple propagations are considered, it is often advisable to write the convolution first, and then apply the Fresnel transform to put in evidence the Fourier transform of a product of convolution.

## 2.2 Filtering in the Fourier space

The convolution relationship in the direct plane corresponds to a linear filtering in the Fourier plane. If we denote  $u$  and  $v$  the spatial frequencies associated with  $x$  and  $y$ , the Fourier transform of Equation (2.9) becomes

$$\hat{A}_z(u, v) = \hat{A}_0(u, v) \cdot \hat{D}_z(u, v) \quad (2.13)$$

where:

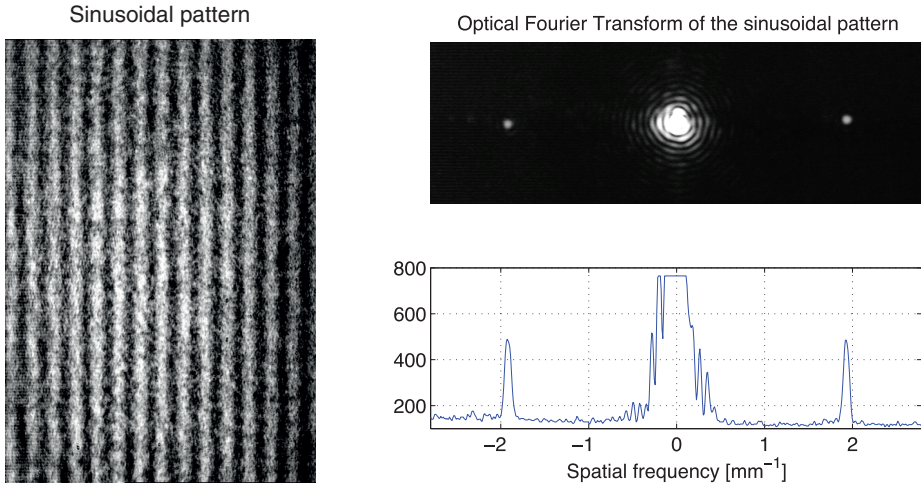
$$\hat{D}_z(u, v) = \exp(-i\pi\lambda z(u^2 + v^2)) \quad (2.14)$$

is the amplitude transfer function for the free space propagation over the distance  $z$ . Each spatial frequency is affected by a phase factor proportional to the square modulus of the frequency. For the sake of simplicity, we will still denote this function a modulation transfer function (MTF), although it is quite different from the usual Hermitian MTFs encountered in incoherent imagery.

The use of the spatial filtering is particularly useful for a numerical computation of the Fresnel diffraction. Starting with a discrete version of  $A_0(x, y)$ , we compute its 2D Fast Fourier Transform (FFT)  $\hat{A}_0(u, v)$ , multiply it by  $\hat{D}_z(u, v)$  and take the inverse 2D FFT to recover  $A_z(x, y)$ . We used this approach to derive the Fresnel diffraction of the petaled occulter given in the last section of this paper.

Before ending this section, we can check that the approximations used there do not alter basic physical properties of the wave propagation. To obtain the coefficient  $\alpha$ , we have used the fact that a plane wave remains a plane wave along the propagation. The reader will also verify that a spherical wave remains also a spherical wave along the propagation. This is easily done using the filtering in the Fourier space. A last verification is the conservation of energy, *i.e.* the fact that the flux of the intensity does not depend on  $z$ . That derives from the fact that the MTF is a pure phase filter and is easily verified making use of Parseval theorem.





**Fig. 3.** *Left:* image of a two-dimensional sinusoidal pattern (Eq. (3.1)). The spatial period is  $1/m = 0.52$  mm. *Top right:* optical Fourier transform of the sinusoidal pattern in the  $(u, v)$  plane. *Bottom right:* plot of the intensity of the optical Fourier transform as a function of the spatial frequency  $u$  for  $v = 0$ .

### 3 Fresnel diffraction from a sinusoidal transmission

The particular spatial filtering properties of the Fresnel diffraction can be illustrated observing how a spatial frequency is modified in the free space propagation. The experiment was presented at the CNRS school observing the diffraction of a plate of transmission in amplitude of the form:

$$f_1(x, y) = (1 - \epsilon) + \epsilon \cos(2\pi(m_x x + m_y y)). \quad (3.1)$$

The plate is a slide of a set of fringes. This transmission in fact bears three elementary spatial frequencies at the positions  $\{u, v\}$  respectively equal to  $\{0, 0\}$ ,  $\{m_x, m_y\}$  and  $\{-m_x, -m_y\}$ . For the simplicity of notations we assume in the following that the fringes are rotated so as to make  $m_y = 0$  and  $m_x = m$ , and we assume  $(1 - \epsilon) \sim 1$ . The fringes and their optical Fourier transform are shown in Figure 3. We describes further in the paper how the operation of Fourier transform can be made optically.

As one increases the distance  $z$ , the fringes in the images almost disappear and appear again periodically with the same original contrast. A careful observation makes it possible to observe an inversion of the fringes in two successive appearances. This phenomenon is a consequence of the filtering by  $\hat{D}_z(u, v)$ . The frequencies at  $u = \pm m$  are affected by the same phase factor  $\exp(-i\pi\lambda z m^2)$ , while the zero frequency is unchanged. At a distance  $z$  behind the screen the complex amplitude therefore expresses as

$$U_z(x, y) \sim 1 + \epsilon \cos(2\pi m x) \exp(-i\pi\lambda z m^2). \quad (3.2)$$



When  $\lambda zm^2$  is equal to an integer number  $k$ , the amplitude is purely real and equal to  $1 \pm \epsilon \cos(2\pi mx)$ . When  $\lambda zm^2 = 1/2 + k$ , the amplitude modulation is an imaginary term, and  $U_z \simeq 1 \pm i\epsilon \cos(2\pi mx)$ . For  $\epsilon$  very small, the wavefront is then almost a pure phase factor of uniform amplitude. It can be represented as an undulated wavefront, with advances and delays of the optical path compared to the plane wave. The wave propagates towards the  $z$  direction, continuously transforming itself from amplitude to phase modulations, as illustrated in Figure 4.

The observed intensity is

$$I_z(x, y) \sim 1 + 2\epsilon \cos(2\pi mx) \cos(\pi \lambda zm^2). \quad (3.3)$$

The fringes almost disappear for  $z = (k + 1/2)/(\lambda m^2)$ , with  $k$  integer. They are visible with a contrast maximum for  $z = k/(\lambda m^2)$ , and the image is inverted between two successive values of  $k$ .

At the CNRS school we have also shown the Fresnel diffraction of a Ronchi pattern, a two dimensional square wave  $f_R(x, y)$  made of alternate opaque and transparent parallel stripes of equal width, as illustrated in Figure 5. Making use of the Fourier series decomposition, we can write the square wave as a simple addition of sinusoidal terms of the form:

$$f_R(x, y) = \frac{1}{2} + \frac{2}{\pi} \sum_{n=0}^{\infty} \frac{1}{2n+1} \sin(2\pi(2n+1)mx). \quad (3.4)$$

The complex amplitude  $U_z(x, y)$  at a distance  $z$  behind the Ronchi pattern is simply obtained by the sum of the sine terms modified by the transfer function. We have:

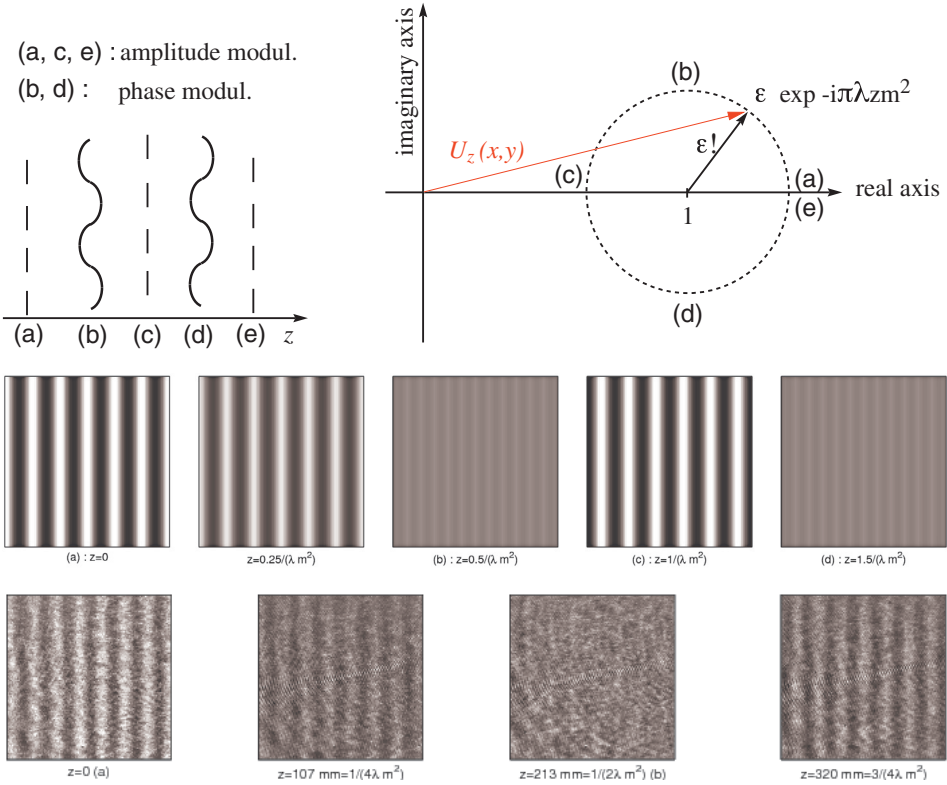
$$U_z(x, y) = \frac{1}{2} + \frac{2}{\pi} \sum_{n=0}^{\infty} \frac{1}{2n+1} \sin(2\pi(2n+1)mx) \exp(-i\pi\lambda z(2n+1)^2 m^2). \quad (3.5)$$

As the wave propagates, each sinusoidal component experiences a phase modulation depending on its spatial frequency. One obtains an image identical to the Ronchi pattern when all the spatial frequencies in  $U_z(x, y)$  are phase-shifted by a multiple of  $2\pi$ . The occurrences of identical images are obtained for  $\lambda zm^2 = 2k$ , as for the single sine term. This property is known as the Talbot effect.

#### 4 Focusing screens and Fourier transform properties of lenses

The Fresnel transform makes easy to introduce the converging lens and its properties relative to the Fourier transform. To make the notations simpler, we assume that the wavefront  $A_0(x, y)$  is simply of the form  $A \times f(x, y)$ , where  $A$  stands for an incident plane wave and  $f(x, y)$  is the transmission of a screen. Let us consider that we can manufacture a phase screen with the following transmission:

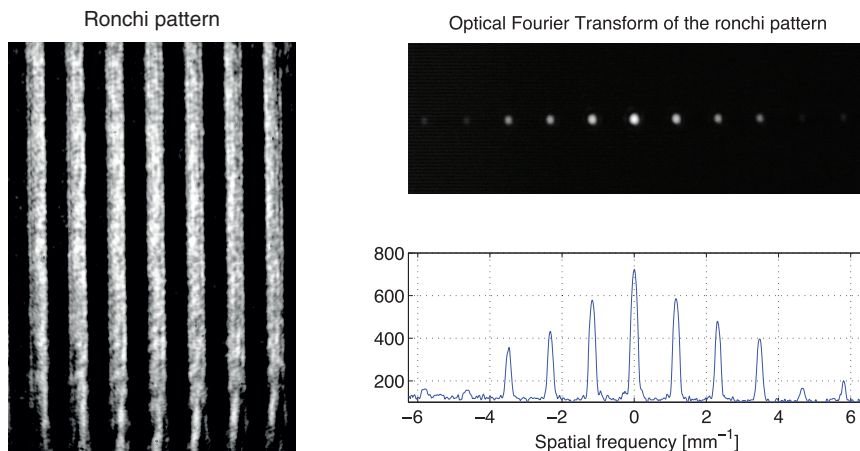
$$L_\phi(x, y) = \exp\left(-i\frac{\pi(x^2 + y^2)}{\lambda\phi}\right) \quad (4.1)$$



**Fig. 4.** Illustration of the Fresnel diffraction of the sinusoidal mask (Eq. (3.1)) using Equation (3.2). *Top left:*  $z$  positions (a), (c), (e) where the amplitudes become again identical to the mask. Corresponding values for the distances are  $z_a = \frac{k}{\lambda m^2}$ ,  $z_c = \frac{k+1}{\lambda m^2}$ ,  $z_e = \frac{k+2}{\lambda m^2}$  ( $k$  integer). Positions (b) and (d) correspond to almost pure phase modulation (uniform intensity):  $z_b = \frac{k+1/2}{\lambda m^2}$ ,  $z_d = \frac{k+3/2}{\lambda m^2}$ . *Top right:* illustration of Equation (3.2) in the complex plane. *Middle row:* simulated images as seen at distances  $z = z_a$ ,  $z = \frac{1/4}{\lambda m^2}$ ,  $z = z_b$ ,  $z = z_c$  and  $z = z_d$ . Notice the contrast inversion between positions  $z_a$  and  $z_c$ . *Bottom row:* experimental images obtained with a sinusoidal grid of frequency  $m = 1/0.52 \text{ mm}^{-1}$ . From left to right: positions  $z = 0$ ,  $z = \frac{1/4}{\lambda m^2}$ ,  $z = z_b$  and  $z = \frac{3/4}{\lambda m^2}$ . Here again, the contrast inversion between the first and last images is visible.

that we affix to  $f(x, y)$ . At the distance  $z = \phi$ , Equation (2.12) shows that the amplitude becomes  $\exp(i\pi \frac{x^2+y^2}{\lambda\phi}) \hat{f}(\frac{x}{\lambda\phi}, \frac{y}{\lambda\phi})$ , and the intensity appears here as a scaled Fourier transform of  $f(x, y)$ .

In the absence of screen (or  $f(x, y) = 1$ ), the diffracted amplitude is proportional to a Dirac function  $\delta(x, y)$ , which explicits the focusing effect of a perfect lens on the axis. Such a phase screen is a converging lens (a thin piece of glass formed between a plane and a sphere gives the desired transmission), or a parabolic mirror of focal length  $\phi$ .



**Fig. 5.** *Left:* image of a Ronchi pattern. The period is  $1/m = 0.86$  mm (see Eq. (3.4)). *Top right:* optical Fourier transform of the sinusoidal pattern in the  $(u, v)$  plane. *Bottom right:* intensity of the optical Fourier transform as a function of the spatial frequency  $u$  for  $v = 0$ . Note that the even harmonics of the frequency  $m$  are present, while they should not with a perfect Ronchi pattern with identical widths of white and black strips.

The phase factor  $\exp(i\pi \frac{x^2+y^2}{\lambda\phi})$  which remains in this focal plane corresponds to a diverging lens  $L_{-\phi}(x, y)$ . It can be cancelled adding here a converging lens of focal length  $\phi$ . So, a system made of two identical converging lenses of focal length  $\phi$  separated by a distance  $\phi$  optically performs the exact Fourier transform of the transmission in amplitude of a screen. Such a device is called an optical Fourier transform system. This property becomes obvious if we re-write the Fresnel transform of Equation (2.12) making explicit the expression corresponding to diverging lenses:

$$A_z(x, y) = \frac{1}{i\lambda z} L_{-z}(x, y) \mathcal{F}_z[A_0(x, y) L_{-z}(x, y)]. \quad (4.2)$$

It is clear here that for the optical Fourier transform system the two converging lenses cancel the diverging terms of propagation. Another similar Fourier transform device can be obtained with a single converging lens of focal length  $\phi$ , setting the transmission  $f(x, y)$  in front of it at the distance  $\phi$  and observing in its focal plane. Such systems have been used to perform image processing, as described by Françon (1979).

It is important to note that phase factors disappear also when the quantity of interest is the intensity, as for example in incoherent imagery (see Sect. 6). Optical Fourier transforms were actually used in the past to analyse speckle patterns at the focus of large telescopes (Labeyrie 1970).

### 4.1 Focusing screens with a real transmission

It is possible to make screens of real transmission (between 0 and 1) acting as converging lenses. For that, the transmission of the screen must contain a term similar to  $L_\phi(x, y)$ . To make its transmission real, we can add the transmission of a diverging lens  $L_{-\phi}(x, y)$ . Doing so we get a cosine term. It is then necessary to add a constant term and use the right coefficients to make the transmission of the screen between 0 and 1. We result in:

$$s_\phi(x, y) = \frac{1}{2} + \frac{1}{4}\{L_\phi(x, y) + L_{-\phi}(x, y)\} = \frac{1}{2} + \frac{1}{2} \cos\left(\frac{\pi(x^2 + y^2)}{\lambda\phi}\right). \quad (4.3)$$

Such a screen acts as a converging lens of focal length  $\phi$ , but with a poor transmission (1/4 in amplitude, 1/16 in intensity). It will also act as a diverging lens and as a simple screen of uniform transmission. A different combination of lenses leads to a transmission with a sine term.

The variable transmission of such screens is very difficult to manufacture with precision. It is easier to make a screen of binary transmission (1 or 0). This can be done for example by the following transmission:

$$\begin{aligned} S_\phi(x, y) &= \frac{1}{2} + \frac{2}{\pi} \sum_{n=0}^{\infty} \frac{1}{2n+1} \sin\left(\pi(2n+1) \frac{x^2 + y^2}{\lambda\phi}\right) = \frac{1}{2} + \\ &\frac{1}{i\pi} \sum_{n=0}^{\infty} \frac{1}{2n+1} \left\{ \exp\left(i\pi(2n+1) \frac{x^2 + y^2}{\lambda\phi}\right) - \exp\left(-i\pi(2n+1) \frac{x^2 + y^2}{\lambda\phi}\right) \right\} \\ &= \frac{1}{2} + \frac{i}{\pi} \sum_{n=0}^{\infty} \frac{1}{2n+1} \{L_{\phi/(2n+1)}(x, y) - (L_{-\phi/(2n+1)}(x, y))\}. \end{aligned} \quad (4.4)$$

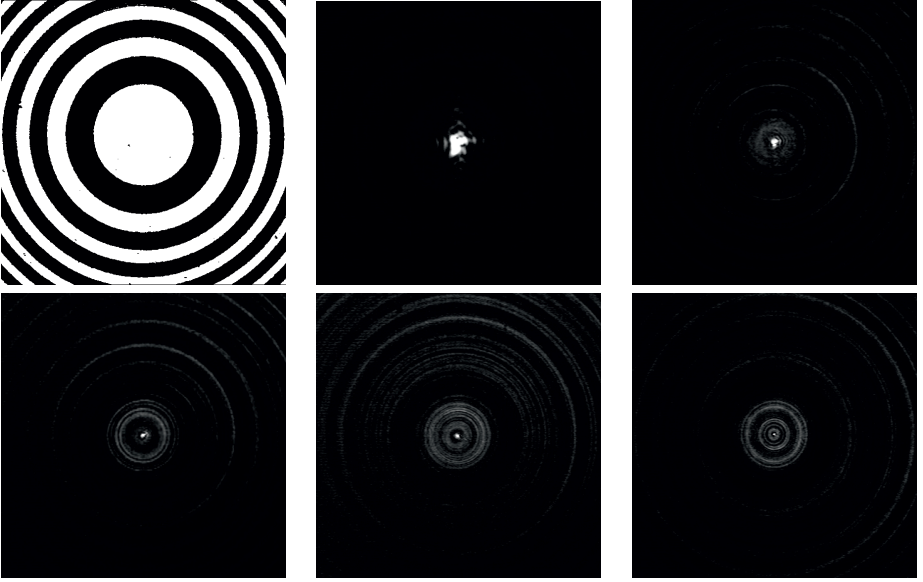
The transmission of such a screen is given in Figure 6 (top left). Its efficiency to focus in the plane  $z = \phi$  is slightly improved (from 1/4 to 1/ $\pi$ ) at the expense of an infinite number of converging and diverging lenses (of focal lengths  $\phi/(2n+1)$ ). A few of these ghost focal planes are shown in Figure 6 (experimental results).

These systems may find interesting applications at wavelengths for which it is difficult to manufacture classical lenses or mirrors. It is interesting to note that screens based on this principle have been proposed also for astronomical applications in the visible domain by Koechlin *et al.* (2009).

## 5 Fresnel diffraction and shadowing in astronomy: Application to coronagraphy

### 5.1 Fresnel diffraction with complementary screens

Let us consider two complementary screens of the form  $t(x, y)$  and  $1 - t(x, y)$ . The amplitude diffracted by the complementary screen is 1 minus the diffracted



**Fig. 6.** *Top row*, from *left to right*: image of a Soret screen, intensity distribution in the plane  $z = \phi$  and  $z = \phi/3$  (corresponding to the focal planes of the lenses  $L_{\phi/(2n+1)}$  for  $n=0,1$  in Eq. (4.4)). *Bottom row*: intensity distribution at  $z = \phi/5$ ,  $z = \phi/7$  and  $z = \phi/9$  (terms  $n=2,3,4$  of the sum). The intensity of the central spot decreases with  $n$  as predicted.

amplitude from  $t(x, y)$ . Indeed, at a distance  $z$ , we have for an incident plane wave of unit amplitude:

$$(1 - t(x, y)) * D_z(x, y) = 1 - t(x, y) * D_z(x, y) \quad (5.1)$$

a property which is sometimes confused with Babinet's principle in the literature (see Cash 2011, for example).

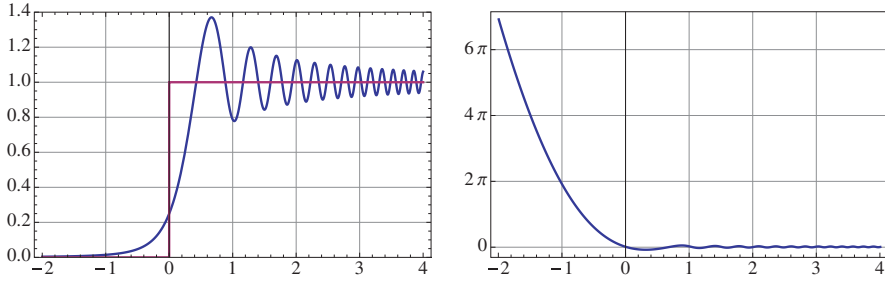
## 5.2 Diffraction with rectangular apertures

The diffraction of rectangular diaphragms (infinite edge, slit, square or rectangle) can be easily computed making use of the separability in  $x$  and  $y$  of these functions and the corresponding properties of the convolution. Indeed, if the transmission  $t(x, y)$  can be written as  $t_x(x) \times t_y(y)$ , then:

$$D_z(x, y) * t(x, y) = D_z^0(x) * t_x(x) \times D_z^0(y) * t_y(y). \quad (5.2)$$

In these cases, many problems find a solution using the Fresnel integrals  $C(x)$  and  $S(x)$ , that can be defined as:

$$F(x) = C(x) + iS(x) = \int_0^x \exp\left(i\frac{t^2}{2}\right) dt. \quad (5.3)$$



**Fig. 7.** *Left:* normalized intensity and *right:* phase (unwrapped) of the Fresnel diffraction of an infinite edge  $H(x)$ , outlined in the left figure. The observing plane is at 1 m from the screen, the wavelength is  $0.6 \mu\text{m}$ . The x-axis is in mm.

The complex amplitude diffracted by an edge is obtained computing the convolution of  $D_z(x, y)$  with the Heaviside function  $H(x)$  for all  $x$  and  $y$  (we may denote its transmission as  $t(x, y) = H(x) \times 1(y)$  for clarity). We have:

$$A_H(x, y) = D_z(x, y) * H(x) = D_z^0(x) * H(x) = \frac{1}{2} + \frac{1}{\sqrt{2i}} F\left(x\sqrt{\frac{2}{z\lambda}}\right). \quad (5.4)$$

The intensity and the phase of the wave are given in Figure 7. The intensity is very often represented in Fresnel diffraction, but this is not the case for the phase. The rapid increase of phase in the geometrical dark zone may be heuristically interpreted as a tilted wavefront, the light coming there originating from the bright zone.

Similarly, the free space propagation of the light for a slit of width  $L$  can be directly derived from the above relation assuming that the transmission is  $t(x, y) = H(x + L/2) - H(x - L/2)$ . We have:

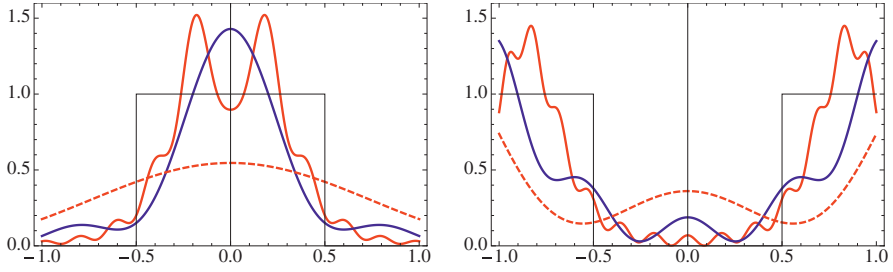
$$A_{L+}(x, y) = \frac{1}{\sqrt{2i}} \left\{ F\left(\frac{2x + L}{\sqrt{2\lambda z}}\right) - F\left(\frac{2x - L}{\sqrt{2\lambda z}}\right) \right\} \quad (5.5)$$

where the subscript  $L+$  stands for a clear slit of width  $L$ . Graphical representations of the corresponding intensity and phase are displayed in Figure 8.

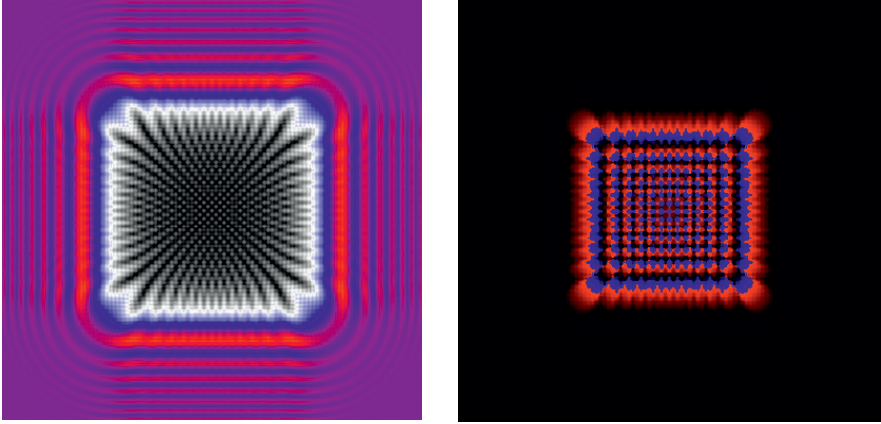
As already written, the Fresnel diffracted amplitude from the complementary screen can be obtained as 1 minus the diffracted amplitude from the slit. It can be also be written as the sum of the diffraction of two bright edges  $H(x - L/2)$  and  $H(-x - L/2)$ . Then it is clear that ripples visible in the shadow of the slit are due to phase terms produced by the edges. We have:

$$\begin{aligned} A_{L-}(x, y) &= 1 - \frac{1}{\sqrt{2i}} \left\{ F\left(\frac{2x + L}{\sqrt{2\lambda z}}\right) + F\left(\frac{2x - L}{\sqrt{2\lambda z}}\right) \right\} \\ &= \frac{1}{\sqrt{2i}} F\left(\frac{-2x - L}{\sqrt{2\lambda z}}\right) + \frac{1}{\sqrt{2i}} F\left(\frac{2x - L}{\sqrt{2\lambda z}}\right) \end{aligned} \quad (5.6)$$

where  $L-$  stands for an opaque strip of width  $L$ .



**Fig. 8.** Fresnel diffractions (in intensity) of a transmitting slit (*left*) and an opaque slit (*right*) of width 1 mm (slits are outlined in the figures). Observing planes are at 0.3 m (red), 1 m (blue) and 3 m (dashed) from the screen, the wavelength is  $0.6 \mu\text{m}$ .



**Fig. 9.** Fresnel diffraction (*left*: amplitude, *right*: phase) of a square occulter of side 50 m at 80 000 km, with  $\lambda = 0.6 \mu\text{m}$ . The region represented in the figures is  $100 \text{ m} \times 100 \text{ m}$ . The color scale for the amplitude (black, white, blue, red) is chosen so as to highlight the structures in the dark zone of the screen. The color scale for the phase is blue for  $-\pi$ , black for 0 and red for  $+\pi$  (the phase is not unwrapped here).

A transmitting square (or rectangle) aperture can be written as the product of two orthogonal slits. Therefore the Fresnel diffraction of the open square  $A_{L^2+}(x, y)$  is the combination of two Fresnel diffractions in  $x$  and  $y$ . This property of separability is no longer verified for the diffraction of the opaque square  $A_{L^2-}(x, y)$ , which transmission must be written as 1 minus the transmission of the open square. We have:

$$\begin{aligned} A_{L^2+}(x, y) &= A_{Lx+}(x, y) \times A_{Ly+}(x, y) \\ A_{L^2-}(x, y) &= 1 - A_{L^2+}(x, y). \end{aligned} \quad (5.7)$$

We give in Figure 9 an example of the amplitude and phase of the wave in the shadow of a square occulter of  $50 \times 50$  meters at a distance of 80 000 km, and that



could possibly be used for exoplanet detection. These parameters are compatible with the observation of a planet at about 0.1 arcsec from the star (a Solar – Earth system at 10 parsec) with a 4-m telescope. It is however interesting to note the strong phase perturbation in the center of the shadow, while it is almost zero outside. In such an experiment, the telescope is set in the center of the pattern to block the direct starlight, and the planet is observable beyond the angular dark zone of the occulter. The level of intensity in the central zone is of the order of  $10^{-4}$  of that of the direct light. This is still too bright to perform direct detection of exoplanets: the required value is  $10^{-6}$  or less. To make the shadow darker it would be necessary to increase the size of the occulter and the distance between the occulter and the telescope.

### 5.3 Fresnel diffraction with a circular occulter: The Arago-Poisson spot

The transmission of a circular occulter of diameter  $D$  can be written as  $1 - \Pi(r/D)$ , where  $r = \sqrt{x^2 + y^2}$ , and  $\Pi(r)$  is the rectangle function of transmission 1 for  $|r| < 1/2$  and 0 elsewhere. Since the occulter is a radial function, its Fresnel diffraction is also a radial function that can be written as:

$$A_D(r) = 1 - \frac{1}{i\lambda z} \exp\left(i\pi \frac{r^2}{\lambda z}\right) \int_0^{D/2} 2\pi\xi \exp\left(i\pi \frac{\xi^2}{\lambda z}\right) J_0\left(2\pi \frac{\xi r}{\lambda z}\right) d\xi \quad (5.8)$$

where  $J_0(r)$  is the Bessel function of the first kind. Here again, the Fresnel diffraction from the occulter writes as 1 minus the Fresnel diffraction of the hole. At the center of the shadow we have  $A_D(0) = \exp[i\pi D^2/(4\lambda z)]$  and we recover the value of 1 for the intensity.

Obtaining the complete expression of the wave for any  $r$  value is somewhat tricky. The integral of Equation (5.8) is a Hankel transform that does not have a simple analytic solution. A similar problem (the wave amplitude near the focus of a lens) has been solved by Lommel, as described by Born & Wolf (2006). It is possible to transpose their approach to obtain the Fresnel diffraction from a circular occulter.

After a lot of calculations, we obtain the result in the form of alternating Lommel series for the real and imaginary parts of the amplitude. The result can be represented in a concise form as:

$$\begin{aligned} \Psi(r) = \\ r < D/2 : & \quad A \exp\left(i \frac{\pi r^2}{\lambda z}\right) \exp\left(i \frac{\pi D^2}{4\lambda z}\right) \times \sum_{k=0}^{\infty} (-i)^k \left(\frac{2r}{D}\right)^k J_k\left(\frac{\pi D r}{\lambda z}\right) \\ r = D/2 : & \quad \frac{A}{2} \left[ 1 + \exp\left(i \frac{\pi D^2}{2\lambda z}\right) J_0\left(\frac{\pi D^2}{2\lambda z}\right) \right] \\ r > D/2 : & \quad A - A \exp\left(i \frac{\pi r^2}{\lambda z}\right) \exp\left(i \frac{\pi D^2}{4\lambda z}\right) \times \sum_{k=1}^{\infty} (-i)^k \left(\frac{D}{2r}\right)^k J_k\left(\frac{\pi D r}{\lambda z}\right). \end{aligned} \quad (5.9)$$

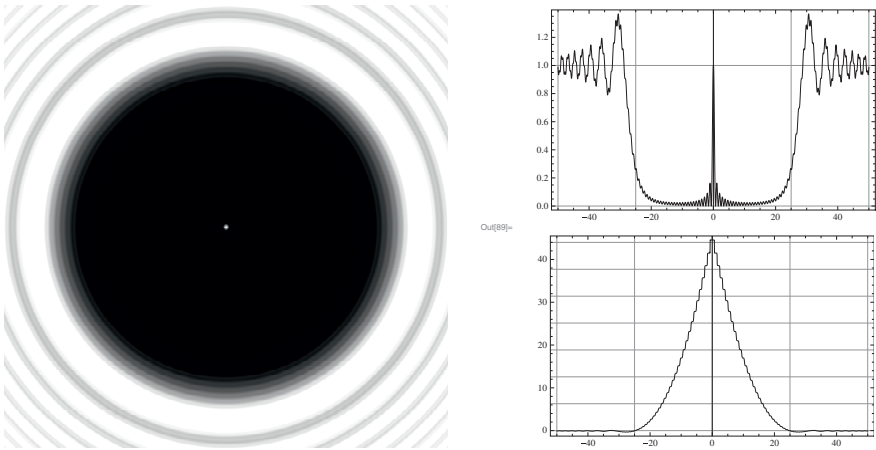
Two expressions are needed to ensure the convergence of the sum depending on the value of  $2r/D$  compared to 1. The convergence is fast except for the transition zone around  $r \sim D/2$ , and luckily there is a simple analytical form there. An upper bound of the series limited to  $n$  terms is given by the absolute value of the  $n + 1$  term, according to Leibniz' estimate.

An illustration of this formula is given in Figure 10 for an occulter of diameter 50 m, observed at a distance of 80 000 km, at  $\lambda = 0.55 \mu\text{m}$ , which corresponds to data for the exoplanet case. For this figure, we computed the series for 100 terms, which can be rapidly done using *Mathematica* (Wolfram 2012) and gives a sufficient precision everywhere. The Arago spot is clearly visible at the center of the diffraction zone. For  $r \ll D$ , the amplitude is fairly described by the only non-zero term of the Lommel series that is the Bessel function  $J_0(\pi r D/(\lambda z))$ . Its diameter is approximately  $1.53\lambda z/D$ .

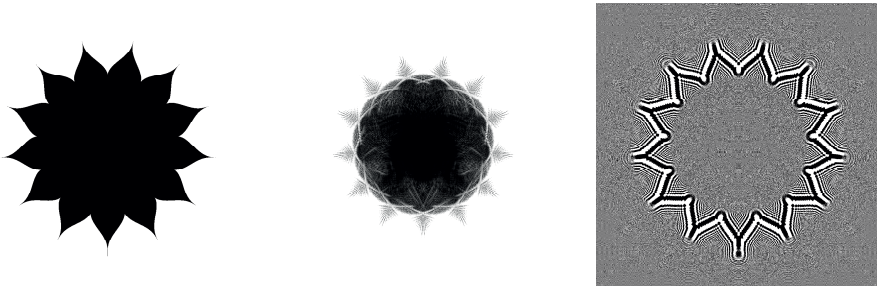
As mentioned in the introduction, Arago's experience was reproduced during the CNRS school of June 2012. Fresnel diffraction patterns (intensity) of a small occulter, reproduced in Figure 1 show the Arago spot at their center.

Thus a circular screen is not a good occulter.

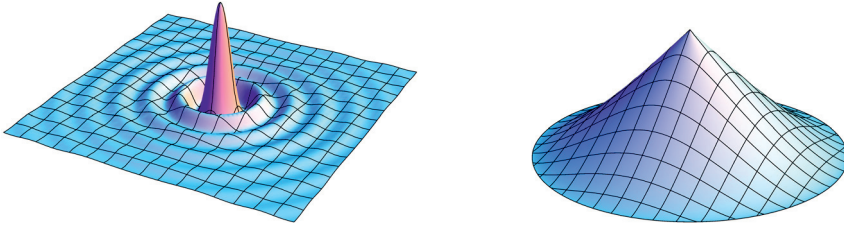
For the detection of exoplanets, several projects envisage petaled occulters (Arenberg *et al.* 2007; Cash 2011) and we give an illustration of the performances in Figure 11. The analytic study of circular occulters remains however of interest for solar applications. Indeed, because of the extended nature of the solar disk, it seems difficult to use shaped occulters there, even if serrated edge occulters have been envisaged for that application (Koutchmy 1988).



**Fig. 10.** Fresnel diffraction of an occulter of diameter 50 m, observed at a distance of 80 000 km, at  $\lambda = 0.55 \mu\text{m}$ . *Left:* 2D intensity, *top right:* central cut of the intensity, *bottom right:* central cut of the unwrapped phase. Notice the strong Arago spot at the center of the shadow and the important phase variation.



**Fig. 11.** Fresnel diffraction of an occulter with a petal shape. From *left to right*: the occulter, the intensity ( $\times 10$ ) in the shadow and the phase. The parameters are the same as in Figure 10  $D = 50$  m,  $z = 80\,000$  km,  $\lambda = 0.55$   $\mu\text{m}$ . The shadow at the center of the screen is much darker (no Arago spot) and the phase variation is weak there.



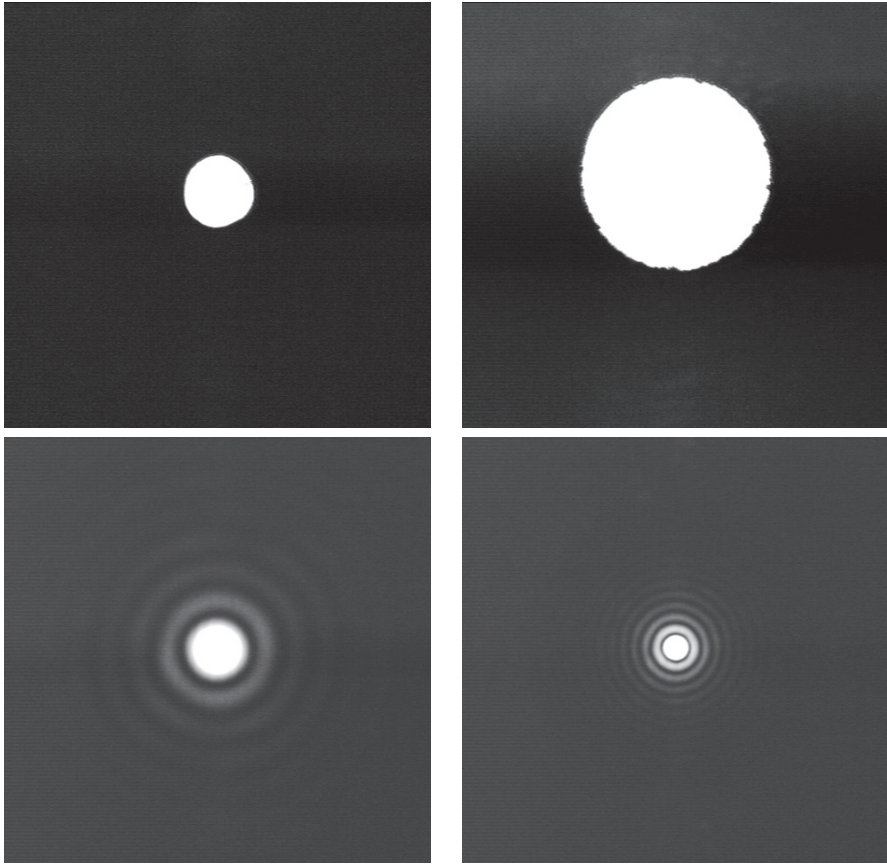
**Fig. 12.** Numerical 3D representation of the PSF (left), here an Airy function, and the corresponding OTF (right) of a perfect telescope with a circular entrance aperture.

## 6 Application to incoherent imaging in astronomy

The formation of an image at the focus of a telescope in astronomy can be divided into two steps, one corresponding to a coherent process leading to the point spread function (PSF) and the other corresponding to a sum of intensities, *e.g.* an incoherent process. Equation (4.2) makes it possible to write the PSF observed in the focal plane of the telescope as a function of the spatial  $(x, y)$  or angular  $(\alpha = x/\phi, \beta = y/\phi)$  coordinates. For an on-axis point-source of unit intensity, we have:

$$\begin{aligned}
 R_\phi(x, y) &= \frac{1}{S\lambda^2\phi^2} \left| \hat{P} \left( \frac{x}{\lambda\phi}, \frac{y}{\lambda\phi} \right) \right|^2 \\
 R(\alpha, \beta) &= \frac{1}{S\lambda^2} \left| \hat{P} \left( \frac{\alpha}{\lambda}, \frac{\beta}{\lambda} \right) \right|^2
 \end{aligned} \tag{6.1}$$

where  $\phi$  is the telescope focal length and  $P(x, y)$  is the function that defines the telescope transmission. Aberrations or other phase defaults due to atmospheric turbulence can be included in the term  $P(x, y)$ . The division by the surface area



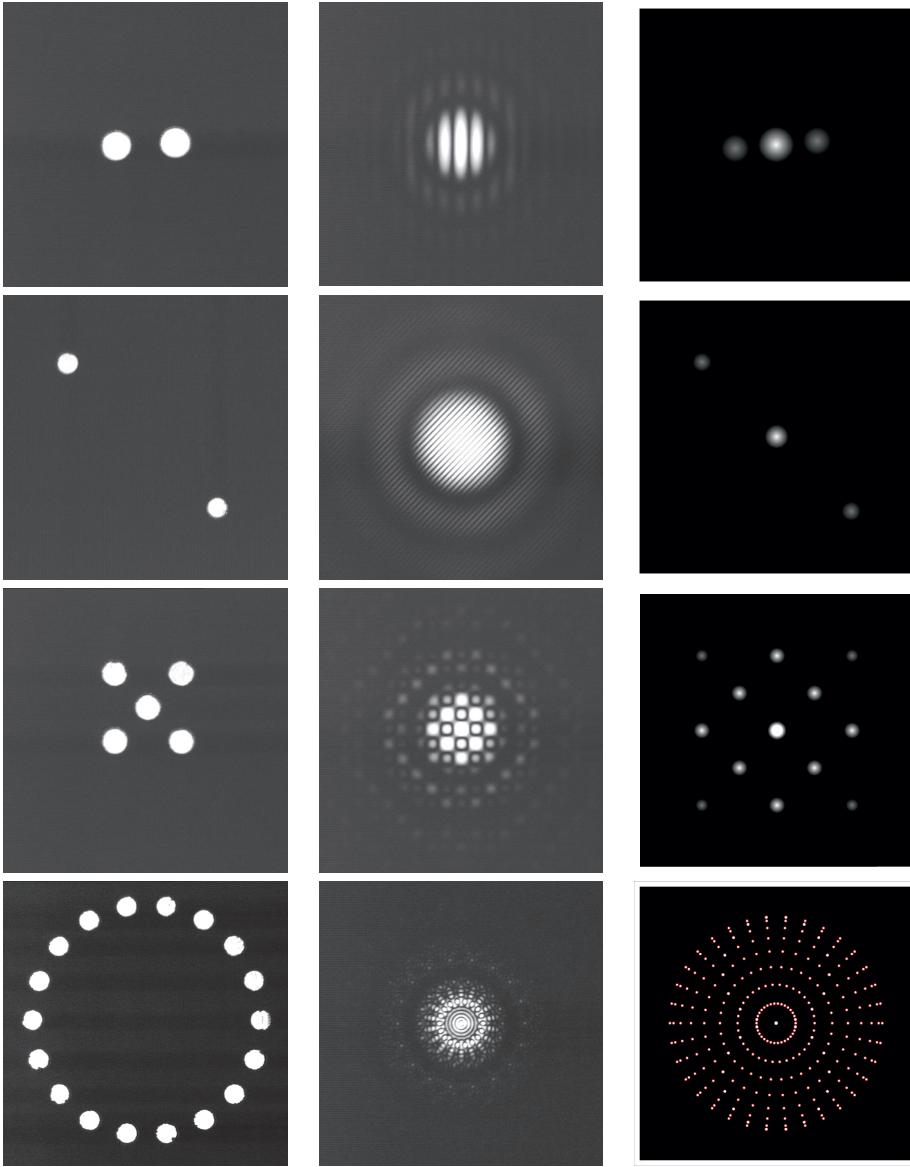
**Fig. 13.** Example of PSFs shown during the CNRS school using a simple optical setup. *Top*, circular apertures, *bottom*, corresponding PSFs. Note the inverse relationship between the size of the PSF and the aperture diameter.

$S$  of the telescope allows the calculation of a normalized PSF. The normalizing coefficients ensure the energy conservation of the form:

$$\iint R_\phi(x, y) dx dy = \iint R(\alpha, \beta) d\alpha d\beta = \frac{1}{S} \iint |P(\xi, \eta)|^2 d\xi d\eta = 1. \quad (6.2)$$

We have made use of Parseval theorem to write the last equality. For a telescope of variable transmission, see Aime (2005).

It is convenient to consider angular coordinates independent of the focal length of the instrument. Each point of the object forms its own response in intensity shifted at the position corresponding to its angular location. This leads to a convolution relationship. The focal plane image is reversed compared to the object sense. By orienting the axes in the focal plane in the same direction as in the sky,



**Fig. 14.** From *left to right*: aperture, PSF and MTF. For the sake of clarity, the MTF corresponding to the 18-aperture interferometer is drawn for smaller elementary apertures than those of the left figure.

we obtain:

$$I(\alpha, \beta) = O(\alpha, \beta) * R(-\alpha, -\beta) \quad (6.3)$$

where  $O(\alpha, \beta)$  is the irradiance of the astronomical object. The Fourier transform of  $R(-\alpha, -\beta)$  gives the optical transfer function (OTF)  $T(u, v)$ :

$$T(u, v) = \mathcal{F}[R(-\alpha, -\beta)] = \frac{1}{S} \iint P(x, y) P^*(x - \lambda u, y - \lambda v) dx dy \quad (6.4)$$

where  $u$  and  $v$  are the angular spatial frequencies.

For a perfect circular aperture of diameter  $D$  operated at the wavelength  $\lambda$ , the PSF becomes the following radial function of  $\gamma$ :

$$R(\alpha, \beta) = R(\gamma) = \left( 2 \frac{J_1(\pi \gamma D / \lambda)}{\pi \gamma D / \lambda} \right)^2 \frac{S^2}{\lambda^2} \quad (6.5)$$

where  $\gamma = \sqrt{\alpha^2 + \beta^2}$ , and the OTF is the radial function of  $w = \sqrt{u^2 + v^2}$ :

$$T(u, v) = T(w) = \frac{2}{\pi} \left( \arccos \left( \frac{\lambda w}{D} \right) - \frac{\lambda w}{D} \sqrt{1 - \left( \frac{\lambda w}{D} \right)^2} \right). \quad (6.6)$$

This expression is obtained computing the surface common to two shifted discs. The OTF looks like a Chinese-hat, with a high frequency cutoff  $w_c = D/\lambda$ .

Examples of PSFs for various apertures presented during the CNRS school are given in Figure 14. The corresponding MTFs shown in the same figure are computed numerically.

## 7 Conclusion

This presentation aimed at introducing the formalism for Fresnel's diffraction theory, widely used in optics and astronomy.

Besides analytical derivation of basic relationships involving instrumental parameters, visual illustrations using laboratory demonstrations are given, as was presented during the CNRS school. Most of these are basic in the field of image formation and are frequently met in astronomy. A few of them concerning the shadows produced by the screens are seldomly addressed in the astronomical literature up to now, though they presently are emerging topics. Demonstrations are made using laboratory material for students in Physics: a laser and a beam expander, various transmitting or opaque screens and a detector.

The paper begins with a historical background leading to the current context. Then analytical derivations, based on the Huyghens-Fresnel principle, using wavefronts and complex amplitudes are presented, providing expressions for the free space propagation of light. Plenty use is made of convolution relationships and filtering aspects.

Fresnel's diffraction is illustrated through some situations, such as the propagation after a screen with sinusoidal transmission function, or such as shadowing produced by occulters set on the pointing direction of a telescope for coronagraphy. Here are met such effects as the so-called Poisson-Arago spot, and diffraction

by sharp edges (rectangular or circular screens). The use of focusing screens have been considered as well. Along that way, expressions of diffracted amplitudes are given for various shapes of apertures. Then, the Fourier transform properties of lenses and binary screens (made of transparent and opaque zones, *i.e.* transmission function being 0 or 1 accordingly) are presented.

The paper ends with a section describing incoherent imaging in astronomy and dealing with PSFs (intensity response of the instrument to a point-like source) and MTFs (a link with linear filtering). Images of PSFs obtained with the demonstration set-up, are presented for various shapes and configurations of collecting apertures: from single disk to diluted apertures (several sub-pupils) as used in aperture synthesis with several telescopes. Besides, illustrations for associated MTFs are obtained by computation.

The paper could hopefully be used either as a reminder or as an introduction to the basics of the image formation process in the context of diffraction theory.

The authors wish to thank Dr S. Robbe-Dubois for her critical reading of the manuscript.

## References

- Aime, C., 2005, A&A, 434, 785  
Arenberg, J.W., Lo, A.S., Glassman, T.M., & Cash, W., 2007, C.R. Physique, 8, 438  
Born, M., & Wolf, E., 2006, Principles of Optics, 7th Ed. (Cambridge University Press), 484  
Cash, W., 2011, ApJ, 738, 76  
Koechlin, L., Serre, D., & Deba, P., 2009, Ap&SS, 320, 225  
Françon, M., 1979, Optical image formation and processing (New York: Academic Press)  
Goodman, J.W., 1985, Statistical Optics (New York, NY: John Wiley and Sons)  
Goodman, J.W., 2005, Introduction to Fourier Optics (Roberts and Company Publishers)  
Koutchmy, S., 1988, Space Sci. Rev., 47, 95  
Labeyrie, A., 1970, A&A, 6, 85  
Lamy, P., Damé, L., Vivès, S., & Zhukov, A., 2010, SPIE, 7731, 18  
Nazarathy, M., & Shamir, J., 1980, J. Opt. Soc. Am., 70, 150  
de Senarmont, M., Verdet, E., & Fresnel, L., 1866, Oeuvres complètes d'Augustin Fresnel (Paris, Imprimerie Impériale)  
Wolfram Mathematica, 2012, Wolfram Research, Inc., Champaign, IL