

Radon approach to shaped and apodized apertures for imaging exoplanets

C. Aime

UMR 6525 Laboratoire Universitaire d'Astrophysique de Nice, Université de Nice Sophia Antipolis, Parc Valrose,
06108 Nice Cedex 2, France
e-mail: Claude.Aime@unice.fr

Received 3 November 2004 / Accepted 10 December 2004

Abstract. In this paper, we present a new approach to the study of shaped and apodized apertures for the detection of exoplanets. It is based on a Radon transform of the telescope aperture and makes it possible to present the effects of shaped and apodized apertures in a unified manner for an objective comparison between them. An illustration of this approach is made for a few apertures. Our conclusion favors the apodized apertures. The approach also permits us to obtain new results. In a second part of the paper, we derive expressions for the signal-to-noise ratio (*SNR*) of an experiment using an apodized aperture and draw the corresponding curves for the example of a circular telescope apodized by a prolate spheroidal function. We found that a very marked improvement of the *SNR* can be obtained using apodization techniques. There is an apodization that optimizes the *SNR* for a given observation; this apodization is generally very strong. The analysis is made for the case of a perfect telescope operated in space.

Key words. techniques: high angular resolution – instrumentation: high angular resolution – telescopes

1. Introduction

The direct observation of an extrasolar planet is a difficult task, not because of the required angular resolution (Jupiter seen at 10 parsec is at 0.5 arcsec of the Sun), but because of the tremendous difference in flux between the planet and the star. The planet should appear over a strong luminous background, the main part of which is due to the diffraction pattern of the star. To detect the planet, this coherent background must be reduced as much as possible. Several techniques, such as the phase-mask coronagraph of Roddier & Roddier (1997) and the four-quadrants coronagraph of Rouan et al. (2000) use an interferometric process to reject the starlight outside the main part of the experiment, with the help of a Lyot stop. They are very promising solutions for exoplanet detection; detailed descriptions of these techniques can be found in several papers and will not be presented here. In the present communication, we focus our analysis on the alternative techniques that seek to detect exoplanets by strongly reducing the level of the wings of the star diffraction pattern at the planet position. In these techniques the starlight is entirely conserved in the experiment while being concentrated in the core of the diffraction pattern. These “apodization” techniques are of interest because they are simpler to implement than the coronagraphs mentioned above and are fundamentally achromatic. Our analysis is restricted to the classical techniques for the case of a perfect telescope operated in space. It does not include the non-linear approaches recently proposed by Guyon (2003) and Traub & Vanderbei (2003).

The paper is organized in two parts. The first part concerns the effects of diffraction. We will see that the Radon-based approach we propose permits a unified view of the effects of shaped and apodized apertures on telescope point spread functions (PSF). The second part of the paper is related to signal-to-noise ratios (*SNR*).

The presentation will make frequent reference to the review paper of Jacquinet & Roizen-Dossier (1964). Jacquinet (1950) was interested in the resolution of spectral lines of very large intensity differences. Assuming that a weak line could be resolved close to a strong line if its intensity was at least comparable to the envelope of the instrumental wings of the strong line, Jacquinet derived that the minimum distance of resolution increases as \sqrt{K} , where K is the contrast between the two lines. This law in \sqrt{K} results from the sinc² diffraction pattern in the spectroscopic one-dimensional geometry. Applied to the Airy pattern, this gives a resolution proportional to $\sqrt[3]{K}$. The figure of merit Q , later introduced by Brown & Burrows (1990) is similar to Jacquinet’s criterion.

Couder & Jacquinet (1939) were at the origin of the word “apodisation” that literally means feet suppression (of the PSF). They showed that this result can be obtained either by *making the rim of contour of the pupil a particular shape* or modifying the transmission of the aperture. These authors made reference to the use of square and polygonal apertures by astronomers for the observation of the companion of Sirius.

The interest in apodization has been constant in the field of optics and was renewed for laser applications. A collection

of very interesting papers can be found in the SPIE Milestone Series of Mills & Thompson (2003). In some of these studies an extensive analytical approach of the problem has been developed.

The importance of apodized apertures in astronomy for the detection of exoplanet was rediscovered by Nisenson & Papaliolios (2001). Since then, the study of various shaped and apodized apertures has been developed by several authors, such as Kasdin et al. (2003), Vanderbei et al. (2003a,b).

The term *apodized* aperture is now used for an aperture with a variable transmission, typically decreasing from the center to the edges. The efficiency of such an aperture for wing reduction follows directly from the properties of the Fourier transform: a smooth, continuous derivable function produces lower side lobes than a step like function. Shaped apertures can give a similar result; how this is obtained is less easy to understand. The Radon approach we present permits us to better understand why the two techniques may have similar effects on the PSF.

Although a PSF with strongly apodized wings is helpful for detecting exoplanets, the relevant criterion is the *SNR* at which the determination can be made. The second part of the paper is concerned with *SNR* estimations. For a perfect experiment, the fundamental limit is that of the photoelectric detection of the light (Goodman 1985). We describe a simple formulation of the *SNR* that uses equivalent surfaces and gives results similar to what could be obtained with a matched filter (Aime 2004).

Numerical examples are given for an experiment using a circular telescope apodized by a prolate spheroidal function. The principal reason for choosing prolate apodization is that it allows us to compare apertures with different strengths of apodization in a continuous way. We show that the apodization must be very strong to improve the *SNR* for exoplanet detection. However, this result may be modified by the presence of an incoherent strong background.

2. Classical effects of diffraction on the telescope PSF

2.1. Basic properties of the diffraction pattern of a telescope of variable transmission

Fourier optics shows that the complex amplitude at the focus of a telescope can be written proportional to the Fourier transform of the wave on the aperture $P(x, y)$, of the form:

$$\Psi(x, y) = \frac{A}{i\lambda F} \hat{P}\left(\frac{x}{\lambda F}, \frac{y}{\lambda F}\right) \quad (1)$$

where F is the focal length of the telescope, A the amplitude of the incident wave and λ the wavelength.

If one calibrates the focal plane in terms of angular units, $\alpha = x/F$ and $\beta = y/F$, the intensity $I(\alpha, \beta)$ can be written as:

$$I(\alpha, \beta) = \frac{I_0}{\lambda^2} \left| \hat{P}\left(\frac{\alpha}{\lambda}, \frac{\beta}{\lambda}\right) \right|^2 = I_0 \times S(\alpha, \beta) \quad (2)$$

where I_0 is the intensity of the light at the telescope aperture.

We consider for later use several properties of $S(\alpha, \beta)$, for the general case in which $P(x, y)$ presents a non-uniform complex transmission. For this we make use of the two integrated

quantities Φ_I and Φ_A that play an important role in the efficiency of a telescope:

$$\begin{aligned} \Phi_I &= \iint |P(x, y)|^2 dx dy \\ \Phi_A &= \iint P(x, y) dx dy \end{aligned} \quad (3)$$

Φ_I and Φ_A represent the flux in intensity and the flux in amplitude for a uniform incoming wave over the aperture. For a telescope with uniform transmission or with phase defaults only, $|P(x, y)|^2 = 1$, and Φ_I equals the surface \mathbb{S} of the aperture.

The intensity at the center of the diffraction pattern is given by:

$$S(0, 0) = \frac{1}{\lambda^2} \left| \iint P(x, y) dx dy \right|^2 = \frac{\Phi_A^2}{\lambda^2}. \quad (4)$$

The term λ^2 accounts for the dependence on wavelength of the size of the PSF. For a perfect telescope with uniform transmission, $\Phi_A = \mathbb{S}$; the intensity at the center of the diffraction pattern increases as \mathbb{S}^2 and the diffraction pattern shrinks as $1/\mathbb{S}$.

In terms of signal and image processing, the PSF is a function whose integral equals 1. We obtain such a function, which we denote $S_{\text{PSF}}(\alpha, \beta)$ by dividing $S(\alpha, \beta)$ by Φ_I . If we divide $S(\alpha, \beta)$ by Φ_A^2/λ^2 we obtain a function that is equal to 1 at the origin and which we denote $S_N(\alpha, \beta)$. These three functions are related to one another by the relation:

$$S(\alpha, \beta) = \Phi_I \times S_{\text{PSF}}(\alpha, \beta) = \frac{\Phi_A^2}{\lambda^2} \times S_N(\alpha, \beta). \quad (5)$$

2.2. Resolution in terms of equivalent solid angle Δ and Strehl ratio ϖ

Rather than using an angle to determine the resolution of a telescope, one may use an angular surface Δ for characterizing the spread of the PSF on the sky. Generalizing the concept of equivalent width used in Fourier transform theory to two dimensions, the equivalent angular surface Δ (or equivalent solid angle) may be written as:

$$\Delta = \frac{1}{S(0, 0)} \iint S(\alpha, \beta) d\alpha d\beta = \frac{\lambda^2 \Phi_I}{\Phi_A^2} \quad (6)$$

where we have made use of Parseval's theorem to derive the result (Aime 2004). This makes it possible to substitute Δ to either Φ_I or Φ_A^2 in Eq. (5). For a perfect aperture with uniform transmission, $\Delta = \Delta_0 = \lambda^2/\mathbb{S}$, that is just the inverse of the surface of the telescope in units of wavelength. These quantities may be considered in relation with the conservation of the optical étendue U :

$$U = \lambda^2 = \mathbb{S} \Delta_0 = \frac{\Phi_A^2}{\Phi_I} \Delta. \quad (7)$$

Owing to the Fourier properties, the inverse of Δ is the equivalent surface (in units of angular frequencies) of the telescope modulation transfer function; it can be considered as a spectral

bandwidth for angular frequencies. The ratio ϖ of the spectral bandwidth $1/\Delta$ of the apodized aperture to that $1/\Delta_0$ of the perfect aperture with uniform transmission, equal to:

$$\varpi = \frac{\Delta_0}{\Delta} = \frac{\Phi_A^2}{\Phi_1 \mathbb{S}} \quad (8)$$

can be interpreted as a Strehl ratio of the apodized aperture. This quantity is smaller than 1; the demonstration that follows was proposed to the author by Michel (2003). Given two square integrable functions $f(x)$ and $g(x)$, the Cauchy-Schwarz inequality states that:

$$\left| \int f(x)g^*(x)dx \right|^2 \leq \int |f(x)|^2 dx \int |g(x)|^2 dx. \quad (9)$$

This relation remains valid for two dimensions if the integration is performed over the surface of the telescope aperture. Setting $f \equiv P$ and $g \equiv 1$, we obtain $|\Phi_A|^2 \leq \Phi_1 \mathbb{S}$, which demonstrates the proposition $\varpi \leq 1$. As a consequence, using an aperture with a non-uniform transmission (or with a contour) can only decrease the resolution in the usual acceptation of the term. We shall see in the following that these considerations must be modified when the objective is to resolve close-by sources with extremely large magnitude differences.

2.3. Shaped and apodized aperture for reducing the wings of the diffraction pattern

For a perfect telescope with a circular aperture, the intensity in the focal plane can be written as:

$$S(\alpha, \beta) = S(\rho) = \left(2 \frac{J_1(\pi \rho D / \lambda)}{\pi \rho D / \lambda} \right)^2 \times \frac{\mathbb{S}^2}{\lambda^2} \quad (10)$$

where $\rho = \sqrt{\alpha^2 + \beta^2}$. The first term of the product is the Airy function. Equation (10) could be further simplified if D is expressed in units of λ , but then the wavelength dependence of the diffraction is lost.

The Airy pattern presents relatively strong wings that hamper the observation of a close-by faint source like an exoplanet. The envelope of the Airy wings decreases only as the cube of the distance from the center. Rings of the diffraction pattern remain above 10^{-3} up to the 4th ring, and decrease below 10^{-4} after the 9th ring only. The diffraction pattern drops below 10^{-5} only at a distance greater than $20\lambda/D$, and would require a distance of $420\lambda/D$ to reach a value of 10^{-9} , comparable to what is expected for a terrestrial exoplanet. This effect is strong enough to consider the perturbations produced by other distant bright stars in the field.

A reduction of the strength of the Airy wings is possible, at the cost of a widening of the central part of the pattern (and therefore of Δ), modifying the pupil in shape or transmission. The first use of such apertures seems to have been published by Couder & Jacquinot (1939) who used a square aperture for the detection of faint spectral lines with a dynamic range up to 10^4 . They wrote the PSF as the following product of a function of α with a function of β of the form:

$$S_N(\alpha, \beta) = \left(\frac{\sin(\pi \alpha L / \lambda)}{\pi \alpha L / \lambda} \frac{\sin(\pi \beta L / \lambda)}{\pi \beta L / \lambda} \right)^2. \quad (11)$$

They were aware of the fact that the envelope of the diffraction pattern at 45° goes as ρ^{-4} , which is better than the ρ^{-3} decrease of the Airy pattern. They referred to the empirical use of polygonal apertures in astronomy at the same time, in particular for the companion of Sirius.

Nisenson & Papaliolios (2001), in their project of an Apodized Square Aperture (ASA), use this same shaped aperture for which the effect of wing reduction is enforced by an apodization with two separable functions of α and β . As in the Couder & Jacquinot example, the aperture is utilized at 45° of the axes.

2.3.1. A unified presentation of diffraction of shaped and apodized apertures based on the Radon transform

A unified presentation of diffraction patterns of shaped and apodized apertures can be presented using the Radon transform. This can be obtained by expressing the focal plane intensity in radial coordinates. For that, we make use of well known properties of two-dimensional Fourier transforms, in particular the so-called central slice theorem. This theorem allows us to write the diffraction pattern in the direction θ as the one-dimensional Fourier transform of the Radon transform of the aperture. This can be demonstrated as follows. The Fourier transform $\hat{P}(u, v)$ of the aperture transmission function $P(x, y)$ can be written as:

$$\hat{P}(u, v) = \int \int P(x, y) \exp -2i\pi(ux + vy) dx dy. \quad (12)$$

We express $\hat{P}(u, v)$ as the function $\hat{R}(\rho, \theta)$ of the radial coordinates ρ and θ , such that $u = \rho \cos(\theta)$ and $v = \rho \sin(\theta)$. In the aperture plane, we write the integration as a function of X and Y , coordinates of the axes rotated by the angle θ from x and y . We have $x = X \cos(\theta) - Y \sin(\theta)$ and $y = X \sin(\theta) + Y \cos(\theta)$. Substituting X, Y, ρ and θ to x, y, u and v we obtain:

$$\hat{R}(\rho, \theta) = \int R(X, \theta) e^{-2i\pi\rho X} dX \quad (13)$$

where $R(X, \theta)$ is the Radon transform of the aperture transmission function:

$$R(X, \theta) = \int P(X \cos(\theta) - Y \sin(\theta), X \sin(\theta) + Y \cos(\theta)) dY. \quad (14)$$

Note that the result for $\theta = 0$ is straightforwardly obtained from Eq. (12) by setting $v = 0$. The function $R(X, \theta)$ is generally represented for positive and negative values of X , and a range of θ values restricted to $\{0, \pi\}$, or to $\{-\pi/2, \pi/2\}$.

An illustration of the computation of the Radon transform $R(X, \theta)$ is given in Fig. 1. It corresponds, for given X and θ to the integral of the aperture transmission function along a line (segment AB in Fig. 1) perpendicular to the direction of diffraction. The result that would be obtained for a clear elliptic aperture is drawn in Fig. 2. In medical tomography, this integral is called a projection; in this domain, the interest is in the inversion of the Radon transform to retrieve $P(x, y)$ knowing $R(X, \theta)$. This is mainly done numerically, using filtered back-projection.

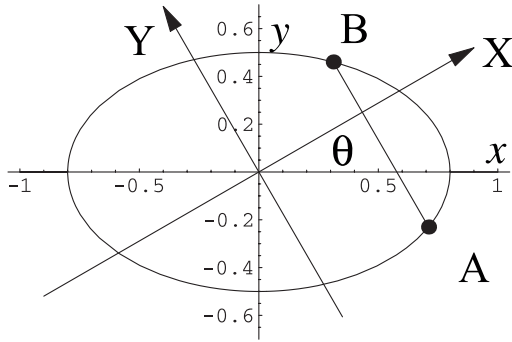


Fig. 1. Illustration of the computation of the aperture Radon transform $R(X, \theta)$ described by Eq. (14). The figure is drawn for an elliptical aperture, with axes proportional to 1 and 1.4. $R(X, \theta)$ corresponds to the sum of the amplitude transmission along AB . For a clear aperture (transmission equal to 1), it simply equals the length of the segment AB .

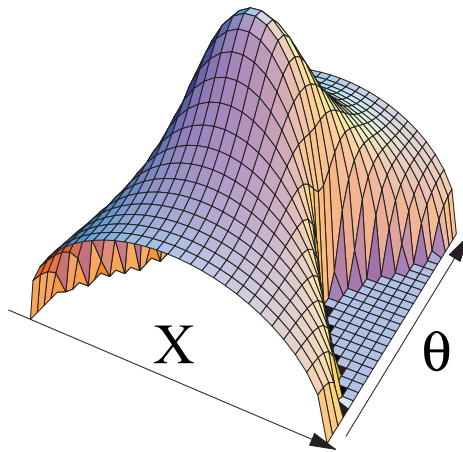


Fig. 2. Radon transform $R(X, \theta)$ of the elliptical clear aperture drawn in Fig. 1. For each θ value the projection takes the form of an ellipse that follows the law $\alpha \sqrt{-X^2 + a^2 \cos^2(\theta) + b^2 \sin^2(\theta)}$, where α is a factor that keeps the integral of the curve equal to the area πab of the ellipse, a and b being the semi-axes of the ellipse.

It might be of interest for astronomy if we seek to find the aperture that produces a given PSF, but this delicate inverse problem (which does not necessarily have a solution) will not be treated here.

2.3.2. Illustration of the Radon approach for simple examples

Let us illustrate the Radon approach for the examples of a circular aperture, a square aperture and a Gaussian shaped aperture.

A clear circular aperture can be taken as the basis for an un-shaped, un-apodized aperture. Its Radon transform is independent of θ , and simply equal to the cord of the circle:

$$R_C(X, \theta) = 2 \sqrt{D^2 - 4 X^2} \Pi\left(\frac{X}{D}\right) \tag{15}$$

where D is the diameter of the aperture, and $\Pi(x)$ a function equal to 1 for $|x| \leq 1/2$ and 0 otherwise.

The one-dimensional Fourier transform of $R_C(X, \theta)$ gives the amplitude of the Airy function (Eq. (10)). If we apply to this aperture a circular-symmetric apodization function $t(r)$, its Radon transform $R_{Ca}(X, \theta)$ will be an even function of X , independent of θ . The circular diffraction pattern can be computed using either Eq. (13) or by the Hankel transform of $P(r)t(r)$:

$$\begin{aligned} \hat{R}(\rho, 0) &= \int_{-\frac{D}{2}}^{\frac{D}{2}} R_{Ca}(X, \theta) \cos(2\pi\rho X) dX \\ &= 2\pi \int_0^{\frac{D}{2}} X P(X) t(X) J_0(2\pi\rho X) dX. \end{aligned} \tag{16}$$

There is an unlimited number of possible two-dimensional transmissions functions for $t(r)$. Some of them have analytical expressions for both the Radon transform and the diffraction pattern; a few of them lead to very simple expressions. For example, with $t(r) = \sqrt{1 - (\frac{2r}{D})^2}$, the Radon transform is proportional to $(1 - (\frac{2X}{D})^2)$, which is known as the Welch apodization window. Using the latter function for $t(r)$ gives a value of $R_{Ca}(X, \theta)$ proportional to $(1 - (\frac{2X}{D})^2)^{\frac{3}{2}}$. The corresponding diffraction pattern is $8J_2(\pi D\rho)/(\pi D\rho)^2$, whose wings decrease as r^{-5} . The diffraction pattern goes below 10^{-9} at a distance 10 times smaller than for the un-apodized aperture. This result is obtained with the drawback of a low throughput of value $1/3$. We discuss the importance of this quantity later in the paper.

The treatment of an elliptical aperture (our example in Fig. 1) could be done as a generalization of that of a circular aperture. The circular symmetry is obviously lost, but the value of $R(X, \theta)$, given in the caption of Fig. 2, resembles that obtained for a circle; most of the above results remain valid after a geometrical transformation that consists of a similitude in the direction of one of the axes of the ellipse.

A square aperture can be considered as the simplest shaped aperture. Its throughput compared with the circular aperture is less by $\pi/2$ (square inscribed in the circular aperture). After a few computations, the formula below can be derived to give the Radon transform of a square:

$$\begin{aligned} R_S(X, 0) &= L\Pi\left(\frac{X}{L}\right) \\ R_S\left(X, \frac{\pi}{4}\right) &= L\sqrt{2}\Lambda\left(\frac{X}{L\sqrt{2}}\right) \\ R_S(X, \theta) &= L\frac{a}{\sin 2\theta} \Lambda\left(\frac{2X}{a}\right) - L\frac{b}{\sin 2\theta} \Lambda\left(\frac{2X}{b}\right) \end{aligned} \tag{17}$$

where $a = \cos \theta + \sin \theta$, $b = \cos \theta - \sin \theta$, L is the side of the square and $\Lambda(x)$ is the triangle function. The expression is valid for θ between 0 and $\pi/4$; the entire Radon transform can be derived using symmetries at θ equal to $\pi/4$ and $\pi/2$. A representation of $R(X, \theta)$ is shown in Fig. 3. The square aperture gives its worst possible values at θ equals 0 and $\pi/2$, where it corresponds to the un-apodized window function. At an angle of $\pi/4$, $R(X, \theta)$ equals the triangle function, an apodization function known as the Bartlett function. In between these two directions, $R(X, \theta)$ has the shape of an isosceles trapezoid.

Various kinds of apodization can be used together with a rectangular aperture. Nisenson and Papaliolios for ASA proposed to use Sonine apodizations, of the form $(1 - x^2)^{\frac{m-1}{2}}$. These

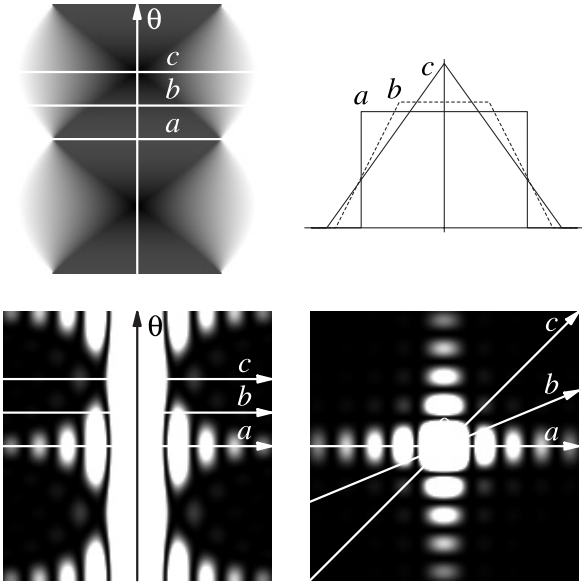


Fig. 3. Illustration of the Radon approach for a square aperture. *Top left:* gray level representation of the Radon transform $R(X, \theta)$ of a square for θ varying between $-\pi/2$ and $\pi/2$. *Top right:* cuts of $R(X, \theta)$ for $a = 0$ (function $L\Pi(\frac{x}{L})$) and $c = \pi/4$ (triangle function $L\sqrt{2}\Lambda(\frac{x}{L\sqrt{2}})$). For intermediate values (such as $b = \pi/8$), the shape is that of an isosceles trapezoid as given in Eq. (17). Curves are such that their integrals equal the area of the square. Because of the regular shape of the square, the Radon transform presents symmetries that makes it possible to restrict its computation to a range of θ values between 0 and $\pi/4$. *Bottom left:* gray level representation of $|\hat{R}(\rho, \theta)|^2$, corresponding to squared modulus of the horizontal one-dimensional Fourier transform of the top left figure. *Bottom right:* usual PSF of the square aperture. Lines a , b and c are the corresponding cuts or transforms for all these figures.

apodizations were compared to prolate spheroidal apodizations by Soummer et al. (2002), who also give expressions for various apodizations and corresponding PSFs.

A shaped aperture can be constructed using simple or very complex contours. The elliptic aperture given as an example in Fig. 1 is a simple modification of the circular aperture, as already discussed. To the contrary, the aperture proposed by Kasdin et al. (2003) uses masks with 6 to 8 elongated transparent zones. In such a case, it is difficult to find an analytical expression for the Radon transform, and the computation must be made numerically. This is already the case for the aperture drawn in Fig. 4 whose contour is defined by two truncated Gaussian curves of the form $\pm \exp(-y^2)$. The resulting figure is not convex; for some values of θ and X , the integration line (a line such as AB in Fig. 1) crosses the aperture in 4 points. The corresponding value for $R(X, \theta)$ is double peaked, and gives strong diffraction arms outside the region where θ is close to 0.

2.3.3. Comparison between shaped and apodized apertures

As already indicated, several recent studies have been made on various shaped apertures (Kasdin et al. 2003; Vanderbei et al. 2003b,a). These authors emphasize two advantages of shaped

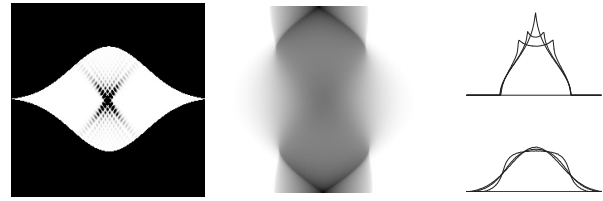


Fig. 4. *Left:* shaped aperture (Gaussian contour) with its corresponding diffraction pattern inside, in a representation similar to that of Jacquinot (1950). *Middle:* Radon transform $R(X, \theta)$ of the aperture for θ varying between $-\pi/2$ and $\pi/2$ (representation similar to that of Fig. 3). *Right:* cuts of the Radon transform for θ values close to 0 (*bottom*), where smooth shapes are obtained, and $\pi/2$ (*top*) where the projections can display double peaks, already visible in the gray level image.

apertures compared to apodized ones. The first is the simplicity of fabrication, which is obvious. The second is that, for a similar result, shaped apertures provide a better intensity throughput than apodized apertures, because the term $|P(x, y)|^2$ makes the intensity flux Φ_I to be smaller than Φ_A for an apodized aperture, while $\Phi_I = \Phi_A$ for a shaped aperture. With the Radon approach of Eqs. (13) and (14), it is clear that different shaped or apodized apertures can lead to the same value of $R(X, \theta)$ for a given direction θ . But different apertures cannot give the same $R(X, \theta)$ for all θ values, unless they are identical. This derives from the inverse properties of both Fourier and Radon transforms. A shaped aperture cannot wholly replace an apodized aperture and vice versa.

An illustration of this is made in Fig. 5, which compares the diffraction patterns of two apertures included in a rectangle of surface S . One aperture is apodized by a linear function in one direction (of the form $1 - 2|x|/L$, L being the length of the rectangle); the other aperture is shaped in the form of a rhombus. Both apertures give the same triangle function for $R(X, 0)$. The value $I(0, 0)$ of the diffraction pattern at the center is the same for these two apertures and equals $1/4$ (times the constant factor S/λ^2). The throughput favors the shaped aperture, as claimed by the authors using these techniques: it is of $1/2$ for the shaped aperture (shaping reduces the aperture area by a factor 2), against only $1/3$ for the apodized aperture (result of the integration of $(1 - 2|x|/L)^2$). But this apparent gain is misleading, and expresses only the fact that the intensity is uselessly spread in the other directions by the shaped aperture. This is clearly visible in Fig. 5, where the diffraction takes the shape of an X elongated in the vertical direction, preventing any useful planet detection in this region. The apodized aperture is much more efficient for the whole plane on average. It makes it possible to discover an exoplanet in a wider region. This can be quantified using the Strehl factor ϖ that strongly favors the apodized aperture for which $\varpi = 3/4$ against only $1/2$ for the shaped aperture.

Similar conclusions can be drawn for the other shaped apertures recently proposed in the literature. In fact, to give better useful throughput than an apodized aperture and the same diffraction pattern in a given direction θ_0 a shaped aperture should be able to produce a value $kR(X, \theta_0)$, with k greater than 1. This is not possible since the maximal value cannot

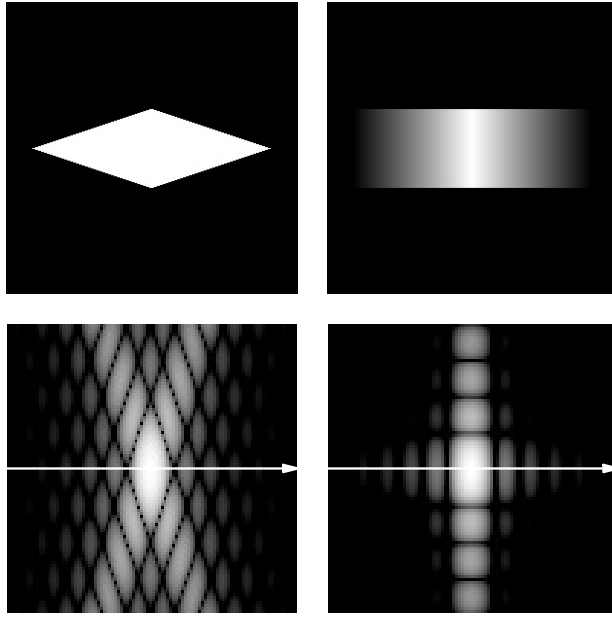


Fig. 5. *Top:* two aperture transmissions included in a rectangle and giving exactly the same triangle function for $R(X, \theta)$ for $\theta = 0$. Left, the rectangular aperture is diaphragmed to a rhombus, right the amplitude transmission of the rectangular aperture linearly decreases from 1 to 0 in one dimension, from center to limb. *Bottom:* corresponding PSFs. Along the horizontal axis (arrows), the two apertures give exactly the same diffraction pattern; these patterns are otherwise different elsewhere.

exceed the length of the aperture in the transverse direction. The interest of a shaped aperture that remains is its ease of fabrication.

Now let us consider the more general case of apodization. If we seek to have a point spread function with reduced side-lobes in a given direction θ , then we want a smooth value for $R(X, \theta)$ in that direction. Since the work of Duffieux (1946), it is known that the diffracted amplitude in the far wings decreases as a power series of the form $f^{(n-1)}x^{-n}$, where $f^{(n)}$ is the value of the n th derivative of the transmission at the edge of the aperture (for a full aperture with no central obscuration). This result is derived from repeated integration by parts of the diffraction integral written as a Fourier transform. The same result applies for the Fourier transform of $R(X, \theta)$. A square aperture gives a decreasing amplitude in $1/x$ along the axes because the Radon transform $R(X, \theta)$ is not zero at $X = L/2$ (Eq. (11)). With this idea in mind, one would propose apodizing functions equal to zero at the edge of the aperture with the first non-zero derivative as high as possible. Jacquinot & Roizen-Dossier (1964) pointed out that this goal is difficult to realize in practice, because the optical density of an absorbing medium cannot rise from 0 to infinity from the center to the edge of the aperture. As a consequence, the transmission at the margin of the aperture may be very low, but not zero. However, the overall shape of the aperture may compensate that effect. Indeed, because of the integration in Eq. (9), a strictly convex two-dimensional aperture (boundary containing no line segment) gives a value of zero for $R(X, \theta)$ at the edge for any θ value.

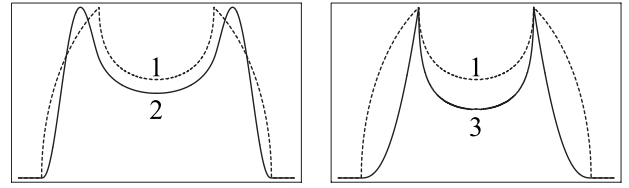


Fig. 6. Example of projections $R(X, \theta)$ for a ring aperture (circular aperture with central circular obscuration) apodized with two different functions. The curve corresponding to the raw aperture (curve 1) is reproduced in the two figures. It simply corresponds to the difference between two functions of the form $\sqrt{1-x^2}$. The other curves correspond to: (2 – left) an apodization that decreases the transmission both toward the limb and the center of the aperture and (3 – right) an apodization where the transmission decreases only towards the limb. In the latter case it is clear that the result may even be worse than the original aperture. All curves are normalized to their maximum.

Several apertures recently proposed, such as the checkerboard aperture of Vanderbei et al. (2004), do not obey that requirement and present discontinuities because they are made of disjoint transmission regions. These discontinuities induce step-like variations in $R(X, \theta)$ that produce ghosts toward some directions, or diffracted amplitudes for circular concentric rings (Vanderbei et al. 2003a,b). In the latter case, $R(X, \theta)$ can be written as a weighted sum of functions of the form given in Eq. (15). An elementary representation of such an aperture is the classical circular aperture with central obstruction that can be written as $2(\sqrt{D^2 - 4X^2} \Pi(\frac{X}{D}) - \sqrt{d^2 - 4X^2} \Pi(\frac{X}{d}))$, where D and d are the outer and inner diameters of the aperture. This function is continuous, but not its derivative at the points $|X| = d/2$, as it can be seen for the dashed curves of Fig. 6. To reduce the side lobes, Jacquinot & Roizen-Dossier (1964) proposed to use pupil transmissions that decrease both toward the edge *and* toward the center of the aperture; for that kind of apodization they used a function of the form $J_2(r)$. In Fig. 6 we have drawn for comparison the results on the projections for the two cases of apodization (we used a simple polynomial function in this example). The projection corresponding to a double apodization presents a smooth structure, while that corresponding only to a single apodization presents a structure with unwanted peaks.

A difficulty that remains is the practical implementation of these continuous apodizations. Good results seem to have been obtained in the past by Jacquinot who used a special apparatus. Recent developments have been made that use interferometric apodizations, as proposed by Aime et al. (2001) and Martinache (2003). The reader will find several other techniques in the selection of papers by Mills & Thompson (2003) already quoted. Some of the techniques proposed are very surprising, such as the apodization using frustrated total reflection proposed by Diels (1975). Nevertheless, it might be interesting to use discrete pupil masks because they appear to be easy to realize from an engineering point of view. In that case we may try to overcome the problem of discontinuities. As a line of investigation one might imagine a discrete aperture drawn continuously in the plane. An example of that is the one in the form of a spiral drawn with a pencil of variable width in Fig. 7.

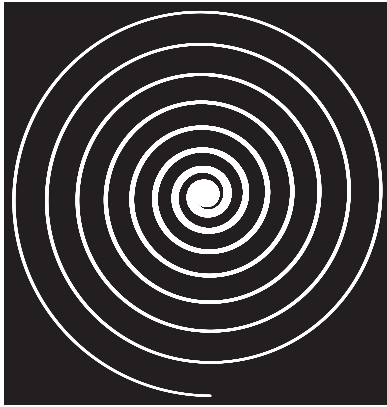


Fig. 7. Principle of a spiral aperture whose projections $R(X, \theta)$ (and first derivative) may not present discontinuities. As described in the body of the paper this aperture is shown for illustration only; a more realistic aperture would be obtained by using a spiral with a larger number of revolutions.

Note that this figure is given only for illustration, and no theory was developed for it by the author. More complex Hilbert plane-filling curves might also be used for the same purpose. We do not intend to develop their study here, and come back to apertures with variable transmission.

For an even pattern, $R(X, \theta)$ must be even. In that case, the modulus squared of the real and imaginary parts of the transform add independently, and there is no advantage for $R(X, \theta)$ not to be real. This result was already obtained by Dossier et al. (1954) using a different reasoning. It might in principle present negative parts corresponding to phase π , but this is unlikely to be realized because of the difficulty of obtaining achromatic phase shifters. We therefore come to the conclusion that the apodizing function should have a real transmission between 0 and 1. It should be noted moreover that all of the classical apodizing functions proposed in the literature of signal processing (Bartlett, Blackman, Cosine, Gaussian, Hamming, Hanning, Welch or others) described for example in Harris (1978) are positive-only functions.

Jacquinet & Roizen-Dossier (1964) consider several techniques for a systematic search for pupil functions with given apodizing properties, such as to have a dark region in the diffraction pattern, an idea further envisaged by Malbet et al. (1995), or to consider several criteria, among them the rate of decrease of energy already discussed, the spreading factor, or the maximum encircled energy. For the latter case, they failed to describe the prolate spheroidal functions that were discovered at that time by Slepian (1964) and Slepian & Pollak (1961) and whose application to optics was later reviewed by Frieden (1971).

An example of prolate spheroidal function and corresponding PSF is given in Fig. 8. For it, we used a special program written by P.E. Falloon (Falloon et al. 2003) in *Mathematica* (Wolfram 1999) to compute prolate circular spheroidal functions. The behavior of prolate apodization is quite different to the other apodizations. Indeed, the rate of attenuation of the wings remains r^{-3} as for the Airy pattern, but starts at a much lower level. This is not surprising because prolate apodizations

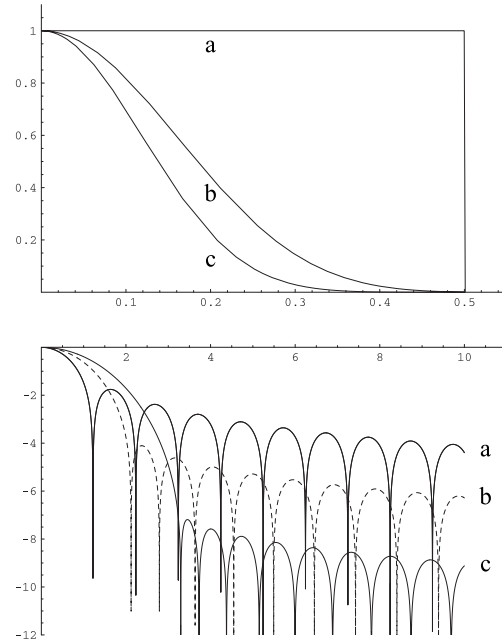


Fig. 8. Examples of prolate apodization functions for a circular aperture. *Top:* radial cuts of the transmission in amplitude, for a telescope of diameter 1 (radius 0.5). *Bottom:* corresponding PSFs, normalized to 1 at the origin ($S_{N,c}(\rho)$); the x -axis is in units of λ/D , the y -axis is on a logarithmic scale. The curves are drawn for the prolate parameter c equal to 6 (curve **b**) and 10 (curve **c**). The un-apodized aperture and corresponding Airy function are also drawn for comparison (curve **a**, prolate parameter $c = 0$). Note that the wings of the prolate PSFs show the same r^{-3} decrease as the Airy pattern, to a strong factor of attenuation.

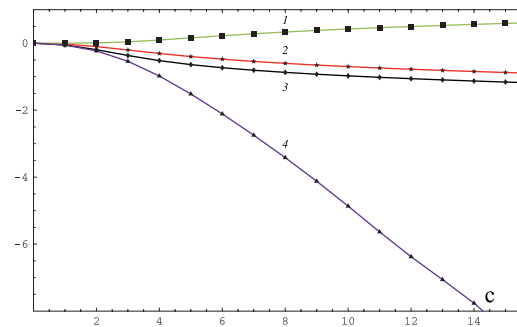


Fig. 9. Representation on a semi-logarithmic scale of (1) the aperture equivalent resolving solid angle Δ , (2) the flux in amplitude Φ_A , (3) the flux in intensity Φ_I and (4) the level of the wings of the apodized aperture. The curves are drawn for a circular prolate apodized aperture; the x -axis corresponds to the parameter c that describes the strength of the apodization. All curves are normalized to 1 at the origin, which corresponds to the raw un-apodized circular aperture.

do not end with zero at the edge of the aperture, and we refer the reader to the reasoning conducted above on the behavior of apodizations and strictly convex apertures. For a circular prolate spheroidal function, the values of Φ_I , Φ_A and Δ are represented in Fig. 9 as a function of the parameter c that defines the strength of the apodization (see for example Frieden (1971) for description of this parameter). In the same graph, we have plotted the decrease of the level of the wings compared to that of the Airy function. For some aspects, the prolate functions may

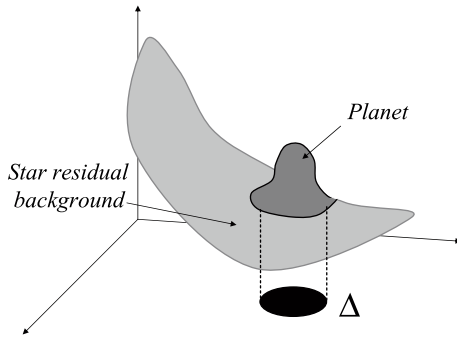


Fig. 10. Schematic representation of a planet over a strong background due for example to the wings of the diffraction pattern of the star. We assume in this model that all the photons from the planet gather in a surface Δ corresponding to the resolution of the aperture. Within this area, the number of photons diffracted by the star is proportional to the volume of a cylinder of basis Δ and height the level of residual light. The noise comes from fluctuations of the overall collected photons within Δ .

be considered as the best apodizers (they maximize the encircled energy). However, their importance for apodization is not so fundamental as in coronagraphy (Soummer et al. 2002).

3. Signal-to-noise ratio considerations

The detection of a signal embedded in noise is a classical problem of signal detection theory (Michel & Ferrari 2003). For large difference between the planet and the background and a large number of collected photons a classical signal-to-noise analysis can be utilized.

The intensity produced in the focal plane by the star is $I_0 S(\alpha, \beta)$, as described by Eq. (2). The constant term I_0 depends on the brightness of the star, and must be expressed in number of photons. If ε is the intensity ratio between the planet and the star, the function corresponding to a planet at the position (α_P, β_P) is $\varepsilon \times I_0 S(\alpha - \alpha_P, \beta - \beta_P)$. To obtain a simple expression for the SNR of the experiment, we make two simplifying assumptions. We assume that the residual diffraction wings of the star can be approximated as a local constant background of value $I_0 \times \bar{S}(\alpha_P, \beta_P)$, where $\bar{S}(\alpha_P, \beta_P)$ corresponds to a local mean of $S(\alpha_P, \beta_P)$, integrated over a region of size Δ . Moreover, we make the optimistic assumption that all the light of the planet can be collected in a pixel, or a group of pixels, of angular equivalent surface Δ . For simplicity of notations, the overall efficiency of the optical system and detector is assumed to be 1; if not, this would change only the value of I_0 .

With this model the expected number of photons collected for the planet is (optimistically) estimated to be $\varepsilon I_0 \times \Phi_1$. Within the resolution surface Δ , the number of photons due to the diffraction of the star is given by the volume of the cylinder below the planet as schematized in Fig. 10:

$$I_0 \bar{S}(\alpha_P, \beta_P) \times \Delta = I_0 \frac{\Delta \Phi_\Delta^2}{\lambda^2} \bar{S}_N(\alpha_P, \beta_P) = I_0 \Phi_1 \bar{S}_N(\alpha_P, \beta_P) \quad (18)$$

where we have used the relations obtained in Eqs. (5) and (6); $\bar{S}_N(\alpha_P, \beta_P)$ corresponds to the same local integration as $\bar{S}(\alpha_P, \beta_P)$. If, moreover, there is a uniform background due for

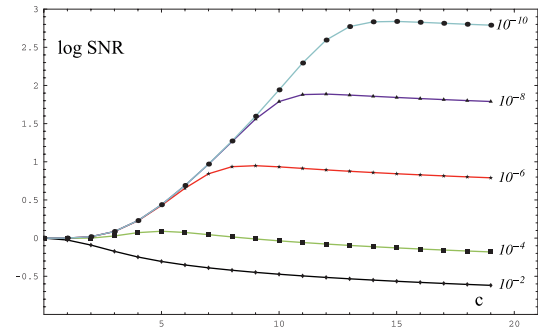


Fig. 11. Representation of the logarithm of the SNR as a function of the strength of apodization (parameter c) for a prolate apodized circular telescope, as given by Eq. (20), for the case of a background free observation ($\gamma = 0$). The parameters entering the expression of the SNR are represented in Fig. 9 for c up to 15. The planet is assumed to be at about $7\lambda/D$. All curves are normalized to 1 at the origin, with reference to the un-apodized case. The values of ε (from 10^{-10} to 10^{-2}) are given at the right of the curves. Very strong improvement of SNR can be obtained for low ε values while apodization is useless for low contrast sources.

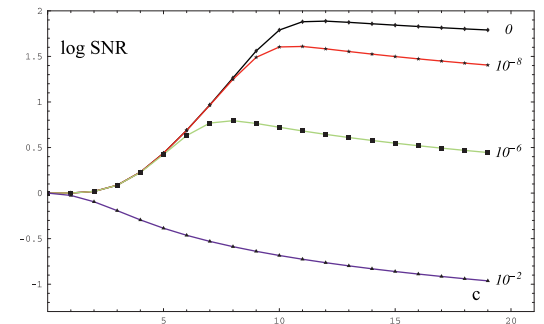


Fig. 12. Similar to Fig. 11, for a fixed value $\varepsilon = 10^{-8}$ and variable values of the background factor γ equal to 0, 10^{-8} , 10^{-6} and 10^{-2} . For a large value of γ , the combined effects of the decrease of transmission and the increase of the equivalent surface Δ make the effect of apodization inefficient. It remains very efficient otherwise (the y -axis is a logarithmic scale).

example to exozodiacal light, we must add to Eq. (18) a term of the form $I_B \Phi_1 \Delta$, where I_B is a factor that measure the light energy per steradian. For simplicity, we write $I_B = \gamma I_0$. The fluctuations of this background produces a noise, to which we must add the fluctuations of the number of photons of the planet itself. For a Poisson process, the variance equals the mean, and the resulting overall variance is simply the sum of all the contributions, regardless of origin:

$$\sigma^2 = I_0 \Phi_1 \bar{S}_N(\alpha_P, \beta_P) + \varepsilon I_0 \Phi_1 + \gamma I_0 \Phi_1 \Delta. \quad (19)$$

Writing the SNR as the ratio of the expected total number of photons due to the planet divided by σ , we have:

$$SNR = \varepsilon \sqrt{\frac{I_0 \Phi_1}{\bar{S}_N(\alpha_P, \beta_P) + \varepsilon + \gamma \Delta}}. \quad (20)$$

An illustration of this result is given in Figs. 11 and 12 for a circular aperture apodized by a circular prolate spheroidal function. The SNR is given as a function of the parameter c , and all curves are normalized to 1 for the unapodized case. Figure 11

is drawn for $\gamma = 0$ (no background) and ε varying from 10^{-2} to 10^{-10} . In this case the relevant parameters are the telescope throughput and the level of the diffraction wings. For low ε values, the loss of photons as the apodization increases is compensated in an advantageous way by the collapse of the wings of the PSF, as it can be seen in Fig. 9.

A strong apodization is more effective than a weak one, and permits considerable improvement of the *SNR*. For low c values (up to 4 or so), the apodization is not efficient since Φ_I decreases as fast as the wings. For very large values of c , the *SNR* decreases when the number of photons in the wings under the planet is comparable to that in the planet, leading to an optimal c value. For a small dynamical range between the sources, apodization is not efficient. This is the case for example for $\varepsilon = 10^{-2}$ (double star of 5 mag difference), at least for the separation chosen ($7\lambda/D$) in Fig. 11. From this example we may conclude that in general the larger the magnitude difference between the star and the planet, the stronger the apodization to use.

The effect of the background may reduce the interest of apodization because it reintroduces the effect of the equivalent area Δ . We show it in an example in Fig. 12, where we have represented the *SNR* for $\varepsilon = 10^{-8}$ and different values of the background coefficient γ . In that case the *SNR* is both sensitive to the throughput and to the equivalent surface of resolution Δ . The optimum is obtained for a lower value of the parameter c ; it still corresponds to a strong apodization. A very strong background (large γ values) makes apodization useless.

It is possible to obtain a simplified form for the *SNR* if we make the assumption that both ε and $\gamma\Delta$ remain small compared to the wings of the PSF. In that case, Eq. (20) reduces to a function $SNR_c(\rho)$ that only depends, for the circular prolate apodization, on the parameter c and the distance star to planet ρ (ρ here stands for α_P, β_P):

$$SNR_c(\rho) \sim \varepsilon \sqrt{I_0} \sqrt{\frac{\Phi_c}{S_{N,c}(\rho)}}. \quad (21)$$

A representation of this relation is given in Fig. 13. To draw the figure, the PSF has been approximated to get rid of the zeroes of the prolate function; the central part of the prolate function is conserved and an envelope of the form κQ^{-3} is substituted for the wings of the functions. As indicated in the comments to the figure, the region of low *SNR* corresponds to the central core of the PSF. A planet, to be observed, should lie outside that region, but fairly close to it (one or two Δ). Within the approximation of Eq. (21), it appears clearly that there is an optimal apodization that depends on the planet position. In fact it corresponds to the strongest apodization allowing the observation of the planet. For example, if a planet is around $3\lambda/D$, the apodization corresponding to $c \sim 5$ seems to be optimal; a stronger apodized aperture ($c \sim 10$) will provide a better *SNR* for a planet around $6\lambda/D$. Increasing the size of the aperture will be particularly efficient because it will permit using stronger apodizations for a given planet position.

The assumptions we have used to derive Eq. (20) are optimistic (all the flux of the planet within Δ) and cannot be verified in practice. It is possible however to use a matched

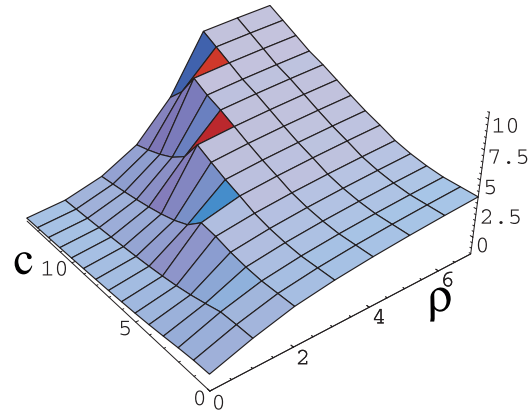


Fig. 13. Representation on a logarithmic scale of the term $\Phi_c/S_{N,c}(\rho)$ that appears in the *SNR* expression of Eq. (21), as a function of the spatial position ρ in units of λ/D and the prolate parameter c . The curve at $c = 0$ corresponds to the un-apodized circular aperture. The *SNR* increases with ρ , as expected. For a large ρ value, the *SNR* increases with c , because the gain obtained by lowering the PSF $S_{N,c}(\rho)$ dominates the throughput reduction Φ_c . The depletion zone on the left (small ρ and large c) corresponds to the wide central core of the PSF that increases with the apodization strength.

filter, convolving the image with the PSF, or an estimate of the PSF (Aime 2004). This will not change the level of the background. It will slightly modify the maximum collected flux for the planet. As described by Aime & Soummer (2003), a new quantity Φ'_I should be substituted to Φ_I , of the form:

$$\Phi'_I = \int \int S^2(\alpha, \beta) d\alpha d\beta = \int \int |T(u, v)|^2 du dv \quad (22)$$

where $T(u, v)$ is the telescope modulation transfer function (MTF). We recall that Φ_I corresponds to the integral of $S(\alpha, \beta)$, instead of that of $S^2(\alpha, \beta)$ for Φ'_I . For circular and rectangular apertures, Aime & Soummer (2003) find that Φ'_I was about half the value of Φ_I . This ratio is about the same for prolate apodizations and is almost independent of the value of the parameter c . We do not give the new values because the curves drawn in Figs. 11–13 remain almost unchanged.

4. Conclusion

The results presented in this paper can be divided in two parts, the first being a Radon presentation of aperture diffraction effects, and the second a presentation of *SNR* for apodized apertures.

We have shown that the use of the Radon transform permits a better understanding of diffraction patterns of shaped and apodized apertures, mainly because it makes it possible to reduce the two-dimensional problem to an ensemble of one-dimensional projections. Not all the possibilities allowed by this new approach have been exploited in this paper. We used it for a comparison between shaped, discrete and apodized apertures. Our conclusion favors the apertures with continuous variable transmission, in contradiction with recent publications on this topic. This assumes, of course, that apertures with perfectly controlled transmission can be realized in practice.

Simplified expressions for the *SNR* of the detection of an exoplanet have also been given. Illustrations for the optimal case of a circular aperture apodized by a prolate spheroidal function have been drawn. Aside from the fact that these functions are optimal in a particular sense for apodization, they make it possible to modulate the strength of the apodization in a continuous way. Several remarks can be made from this study. One is that the apodization must be very strong to be efficient for faint exoplanet detection. The *SNR* improvement can then be very large. Moreover, the apodization must be adapted to the star to planet distance. As a simple rule-of-thumb, the optimal apodization is the strongest that permits geometric observation of the planet. This conclusion greatly favors the use of the largest possible telescopes.

Acknowledgements. The author would like to thank Peter Falloon for his Mathematica program, Andrea Ferrari, Henri Lantéri and Olivier Michel for stimulating discussions. Thanks are also due to the referee Wesley Traub for very constructive comments, and in particular for his suggestion to use the étendue in Sect. 2.2.

References

- Aime, C. 2004, *EAS Publ. Ser.*, 12, 65
 Aime, C., Soummer, R., & Ferrari, A. 2001, *A&A*, 379, 697
 Aime, C., & Soummer 2003, *EAS Publ. Ser.*, 8, 353
 Brown, R. A., & Burrows, C. J. 1990, *Icarus*, 87, 484
 Couder, A., & Jacquinot, P. 1939, *C R Acad Sci. Paris*, 208, 1639
 Diels, J.-C. 1975, *Appl. Opt.*, 14, 2810
 Dossier, B., Boughon, P., & Jacquinot, P. 1950, *J Recherches, CNRS*, No. 11
 Duffieux, P. M., *L'intégrale de Fourier et ses applications à l'optique*, Université de Besançon, France
 Falloon, P. E., Abbott, P. C., & Wang, J. B. 2003, *J. Phys. A*, 36, 5477
 Frieden, B. R. 1971, *Prog. Opt.*, 9, 311
 Goodman, J. W. 1968, *Introduction to Fourier Optics* (McGraw-Hill)
 Goodman, J. W. 1985, *Statistical Optics* (New York: Wiley & Sons)
 Guyon, O. 2003, *A&A*, 404, 379
 Harris, J. H. 1978, *Proc. IIIE*, 66, 51
 Jacquinot, P. 1950, *Proc. Phys. Soc. B*, 63, 969
 Jacquinot, P., & Roizen-Dossier, B. 1964, *Prog. Opt.*, 3, 31
 Kasdin, N. J., Vanderbei, R. J., Spergel, D. N., & Littman, M. G. 2003, *ApJ*, 582, 1147
 Malbet, F., Yu, J. W., & Shao, M. 1995, *PASP*, 107, 386
 Martinache, F. 2004, *EAS Publ. Ser.*, 12, 311
 Mehta, C. L. 1970, *Prog. Opt.*, 8, 374
 Michel, O. 2003, private communication
 Michel, O., & Ferrari, A. 2003, *EAS Publ. Ser.*, 8, 129
 Mills, J. P., & Thompson, B. J. 1999, *SPIE Milestone Series*, MS119
 Nisenson, P., & Papaliolios, C. 2001, *ApJ*, 548, L201
 Roddier, F., & Roddier, C. 1997, *PASP*, 109, 815
 Rouan, D., Riaud, P., Boccaletti, A., Clénet, Y., & Labeyrie, A. 2000, *PASP*, 112, 1479
 Slepian, D. 1964, *Bell Syst. Tech. J.*, 43, 3009
 Slepian, D., & Pollak, H. O. 1961, *Bell Syst. Tech. J.*, 40, 43
 Soummer, R., Aime, C., Ferrari, A., & Falloon, P. E. 2002, *SPIE*, 4860, 211
 Soummer, R., Aime, C., & Falloon, P. E. 2003, *A&A*, 397, 1161
 Traub, W. A., & Vanderbei, R. J. 2003, *ApJ*, 599, 695
 Vanderbei, R. J., Spergel, D. N., & Kasdin, N. J. 2003a, *ApJ*, 590, 593
 Vanderbei, R. J., Spergel, D. N., & Kasdin, N. J. 2003b, *ApJ*, 599, 686
 Vanderbei, R. J., Kasdin, N. J., & Spergel, D. N. 2004, *ApJ*, 615, 555
 Wolfram, S. 1999, *The Mathematica Book*, Fourth Edition (Cambridge University Press)