Astronomy & Astrophysics

Apodized apertures for solar coronagraphy

C. Aime

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

ABSTRACT

Aims. We propose the principle of a new solar telescope that makes it possible to observe the solar corona very close to the solar limb, without the help of a Lyot coronagraph. The result is obtained using a strongly *apodized* aperture.

Methods. We obtain the theoretical form of the diffraction halo produced by the solar disk at the level of the corona for a perfect diffraction-limited telescope, for raw and apodized apertures. The problem is first solved at one dimension for which a complete set of analytical expressions can be derived, including the effect of the center-to-limb solar variation. Formal equations are written for the two-dimensional case, and it is shown that the expression may take the form of a 1D integral. Nevertheless, the problem is difficult to solve. An analytic expression can be worked out using the line spread function, which is shown to give a valid approximation of the problem, in excellent agreement with a numerical computation that uses the exact integral.

Results. We show for the raw aperture that the diffraction halo is very strong and decreases slowly as ρ^{-1}. We propose as a solution to this problem an apodized aperture based on the generalized prolate spheroidal functions (GPSF). Such an apodized aperture may reduce the diffraction halo enough to permit a direct observation of the solar corona very close to the solar limb. A signal-to-noise ratio analysis is given.

Conclusions. Different strengths of apodization may be used, but very strong apodizations are indeed mandatory. A good choice seems to be a GPSF aperture with the prolate coefficient c on the order of 10. It could reduce the halo of diffraction by a factor 10^5 (at the cost of an intensity throughput of 10% and a reduction in the classical resolution by a factor of about 1.6) and permit observation of the corona very close to the solar limb.

Key words. Sun: corona – instrumentation: high angular resolution – telescopes

1. Introduction

The coronagraph of Lyot (1932) uses an elaborated experiment to permit observation of the solar corona. It requires two diaphragms. The Lyot mask, set in the telescope focal plane, suppresses the direct image of the solar photosphere. For each point source of the solar surface, the mask induces a system of two waves: the direct wave of the instrument and the wave diffracted by the mask, as described by Aime et al. (2001). Babinet theorem shows that the two waves interfere in a destructive way in the image of the pupil, therefore reducing the direct light through the aperture image. The missing light is found outside the geometrical image of the aperture where a Lyot stop suppresses it. The solar corona can be observed in the re-imaged focal plane. To work satisfactorily, the system requires ideal observing conditions since the aberrated part of the wave cannot be corrected by the coronagraph (Aime & Soummer 2004). It also needs an excellent optical system over a wide field of view, because the Lyot mask is at least equal to the angular size of the Sun.

In the present paper, we present an alternative technique that is much simpler. An apodized aperture is used to reduce the wings of telescope point spread function (PSF) so that the light diffracted by the solar disk becomes lower than the light of the corona. No diaphragm is needed to observe the corona, and the sunlight may be entirely conserved in the experiment. It is not excluded, however, that the disk of the Sun be shielded to prevent the saturation of the detectors.

Apodized apertures have been extensively studied in optics, as can be seen in the excellent review paper of Jacquinot & Roizen-Dossier (1964). This technique has been proposed recently by Nisenson & Papaliolios (2001) for detecting exoplanets and further studied by several authors, among them Kasdin et al. (2004), Vanderbei et al. (2003), and Aime (2005a). The principal drawback of the technique is a reduction in the transmitted light, a problem that is of less importance for solar studies than for exoplanet detection.

The paper is organized as follows. After a short presentation of the problem in Sect. 2, we give a simple one-dimensional (1D) modeling of the problem in Sect. 3. The 1D simplification permits fully analytic expressions to be obtained that are very similar to the two-dimensional (2D) case treated in Sect. 4. For the 2D case, the result remains analytic if we make simplifying assumptions. The result leads to a very surprisingly simple expression for the solar diffracted light that decreases slowly as the inverse of the distance from the solar limb. This expression is used in Sect. 5 to propose prolate spheroidal apodization functions that are good enough to permit direct observation of the solar corona. A signal-to-noise ratio analysis is made in Sect. 6 and conclusions given in Sect. 7.

2. Position of the problem: the light diffracted by the solar disk

Denoting $S_\odot(\alpha, \beta)$ the geometrical image of the solar disc, the observed image $I_\odot(\alpha, \beta)$ in the focal plane of the telescope is

given by the relation of convolution:

$$
\begin{aligned}
I_\odot(\alpha,\beta) &= R(\alpha,\beta) * S_\odot(\alpha,\beta) \\
&= \int\int R(\xi,\eta) S_\odot(\alpha-\xi,\beta-\eta)\mathrm{d}\xi\mathrm{d}\eta
\end{aligned}
\tag{1}
$$

where $*$ denotes the convolution and $R(\alpha,\beta)$ is the PSF. Since the convolution is commutative, the functions $S_\odot$ and R can be exchanged in the integral, if necessary.

The PSF can be written as:

$$
R(\alpha,\beta) = \frac{1}{\lambda^2}\left|\hat{P}\left(\frac{\alpha}{\lambda},\frac{\beta}{\lambda}\right)\right|^2
\tag{2}
$$

where $P(x,y)$ is the transmission in complex amplitude of the telescope aperture. We assume in this presentation that the telescope is perfect and is operated outside the Earth's atmosphere.

Each point of the solar surface contributes to a very small amount of diffracted light outside the Sun, but the addition of all intensities produces a halo of light bright enough to overcome that of the corona by a large factor. To estimate its level, a precise computation of $I_\odot(\alpha,\beta)$ is necessary, as follows.

3. Analytic solutions for a 1D model

For the 1D model we have derived the expressions for the images of the solar disk obtained using a raw aperture and an apodized one. Drawings are limited to the part of the image near or outside the solar limb. Computations of the integrals were derived using the software *Mathematica*, and the reader can directly check them using the free "Integrator" site of *Mathematica* (Wolfram 1999), for example.

3.1. Diffraction halo for the raw aperture

The term "halo" is used here to concisely describe the spread of the image of the solar disk outside the limb due to diffraction, with no reference to any structure of a solar origin. For the 1D model, we take a top-hat function of width D for the telescope transmission $P(x)$; its PSF is the square of the sine cardinal function. We have

$$
P(x) = \Pi\left(\frac{x}{D}\right), \quad R(\alpha) = \frac{\sin^2(\pi\alpha)}{\pi^2\alpha^2}
\tag{3}
$$

where we have expressed α in units of λ/D to make the mathematical expressions simpler. The reader will note that the wings of this 1D PSF decrease as α^{-2}.

We denote $D_\odot = 2\,R_\odot$ as the diameter of the Sun and assume no center-to-limb variation, so that $S_\odot(\alpha) = \Pi(\alpha/D_\odot)$. Using Eq. (1) for the 1D case, the solar image $I_\odot(\alpha)$ becomes

$$
\begin{aligned}
I_\odot(\alpha) &= \int_\alpha^{\alpha+D_\odot}\frac{\sin^2(\pi\xi)}{\pi^2\xi^2}\mathrm{d}\xi = \frac{\sin^2(\pi\alpha)}{\pi^2\alpha} - \frac{\mathrm{Si}(2\pi\alpha)}{\pi} \\
&\quad - \frac{\sin^2(\pi(D_\odot+\alpha))}{\pi^2(D_\odot+\alpha)} + \frac{\mathrm{Si}(2\pi(D_\odot+\alpha))}{\pi}
\end{aligned}
\tag{4}
$$

where $\mathrm{Si}(\alpha)$ is the sine integral function, defined as the integral of $\sin(t)/t$ between 0 and α. For convenience, the right limb of the Sun is set to the position zero of the axes.

Equation (4) can be simplified if we assume that $D_\odot$ is very large compared to the size of the PSF, which is the case in practice. Making $D_\odot \to \infty$, we have

$$
I_\infty(\alpha) = \frac{\sin^2(\pi\alpha)}{\pi^2\alpha} - \frac{\mathrm{Si}(2\pi\alpha)}{\pi} + \frac{1}{2}.
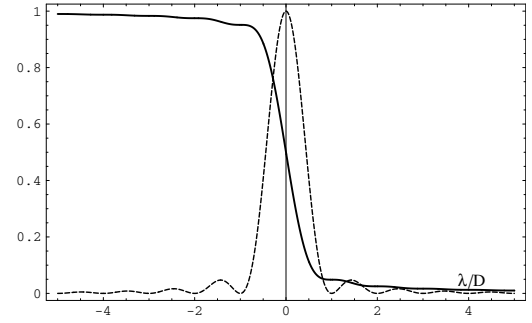\tag{5}
$$

Fig. 1. One-dimensional simulation: the PSF of Eq. (3) for a top-hat aperture and the corresponding image $L_\infty(x)$ of the solar limb as given in Eq. (5). On that scale there is no difference between $L_\odot(x)$ and $L_\infty(x)$.
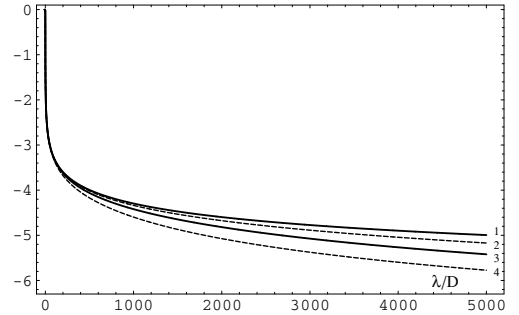

Fig. 2. Representation on a logarithmic scale of the halo profile. Curve 1 is the function $I_\infty(x)$ already drawn on a linear scale in Fig. 1. Curves 2 to 4 are computed for $D_\odot$ values of 10 000, 3000, and 1000 (units of λ/D), or equivalently for apertures in the visible of about 70, 20, and 7 centimeters, respectively. The closeness of curves 1 and 2 permits using $I_\infty(x)$ as a good approximation of $I_\odot(\alpha)$ for a large aperture.

In the linear representation of Fig. 1, there is no visible difference between $I_\infty(\alpha)$ and $I_\odot(\alpha)$. To differentiate the curves, it is necessary to plot them on a logarithmic scale, as in Fig. 2.

The approximation of $I_\odot(\alpha)$ by $I_\infty(\alpha)$ is very good for a large telescope (curve 2); for a smaller telescope (curve 4), $I_\odot(\alpha)$ is clearly lower than $I_\infty(\alpha)$. The difference between the curves corresponds to the integration of Eq. (4) from $\alpha+D_\odot$ to infinity. For a large telescope, the PSF at the distance $D_\odot$ is low enough for that contribution to be negligible; this assumption is no longer valid for a smaller telescope for which the solar diameter corresponds to a lower number in units of λ/D. The total extent of the representation (5000 units of λ/D) of the figure would correspond, for a 50-cm aperture, to a distance of about 20 arcmin at 600 nm. The interesting point is that $I_\infty(x)$ is a good approximation of $I_\odot(\alpha)$ for a large telescope, so we may conjecture that this will also be true for the 2D case treated in Sect. 4, for which no analytic expression seems to be available for the finite summation.

For high α values, $I_\infty(\alpha)$ is asymptotically equal to $1/(2\pi^2\alpha)$, or $0.05\,\alpha^{-1}$. This measure is in units of λ/D, and the larger the telescope, the lower the level of the diffraction halo at a fixed distance from the limb. However, since the halo is typically a factor 10^4 too high, the corona detection cannot come from an increase in the telescope size. The reader will note that the α^{-1} regime comes from the α^{-2} asymptotic value of the PSF and its integration. We will see that a similar result is obtained for the 2D case.
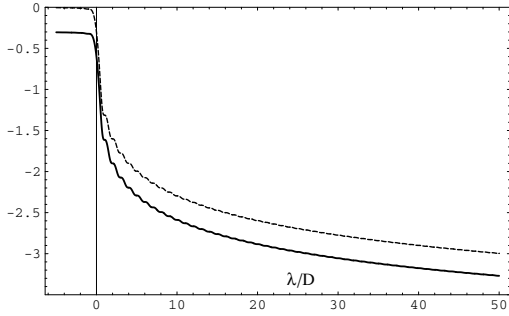
Fig. 3. Images of the solar limb obtained for a uniform Sun (upper dashed line) and a center-to-limb variation in the intensity of the form $1 - 2\alpha^2 D^{-2}$ (lower continuous line). The normalization is such that the brightness is 1 at the disk center for the two curves.

3.2. Effect of a center-to-limb variation

It is possible to obtain analytic expressions for the solar image by assuming a simple form of the center-to-limb variation of the solar brightness. The curve drawn in Fig. 3 is for a parabolic brightness intensity variation of the form $1 - 2\alpha^2 D_\odot^{-2}$, i.e. 1 at the center and 0.5 at the limb. The analytic expression is a few lines long, and we thought that it was not essential to give them here. As can be seen in Fig. 3, the diffraction figure remains similar to that of the model with no center-to-limb variation, but for a constant factor due to the level of the limb. This result permits us to be confident about the shapes derived from the simple top-hat model for the Sun. In all cases, the light diffracted by the solar disc produces a halo that decreases only as α^{-1} remains well above the 10^{-6} level expected for the corona for more than a solar radius and prevents its direct observation.

3.3. Diffraction halo for an apodized aperture

The term "apodized aperture" is used for an aperture with a transmission variable from the center to the edge, chosen so that its PSF presents the lowest possible level of diffraction feet. Depending on the transmission, computing of the integral in Eq. (1) may or may not lead to an analytical solution.

There is an analytical solution for an aperture transmission following the triangle function $P(x) = (1 - 2|x|/D)$, an apodization function known as the Bartlett function in signal processing. The PSF is given by the square of the PSF of the raw aperture, because the triangle can be written as the autocorrelation function of the top-hat function. The throughput of such an apodized aperture is reduced to $1/3$ of the un-apodized one. There is a loss of resolution that we can fix to $3/4$, in terms of the Strehl ratio ϖ as defined by Aime (2005a) as the ratio of the spectral angular bandwidth of the apodized aperture divided by that of the aperture with a uniform transmission. These losses are the price to pay to obtain reduction of the diffraction wings. We have

$$P(x) = \Lambda\left(\frac{x}{D}\right), \quad R(\alpha) = \frac{4\sin^4(\pi\alpha/2)}{\pi^4\alpha^4}. \quad (6)$$

The wings of the apodized PSF decrease now as the function α^{-4}. The image of the solar limb, in the limiting case where $D_\odot \to \infty$, can be written as

$$I_\infty(\alpha) = \frac{\cos(\pi\alpha) - \cos(2\pi\alpha)}{\pi^2\alpha} + \frac{-4\cos(\pi\alpha) + \cos(2\pi\alpha) + 3}{2\pi^4\alpha^3}$$
$$+ \frac{2\sin(\pi\alpha) - \sin(2\pi\alpha)}{2\pi^3\alpha^2} + \frac{Si(\pi\alpha) - 2Si(2\pi\alpha)}{\pi} + \frac{1}{2}. \quad (7)$$
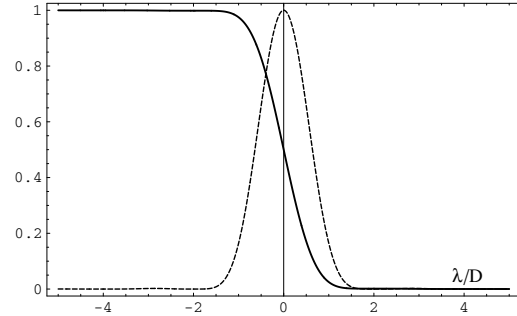

Fig. 4. The PSF of Eq. (6) for a transmission aperture with the form of a triangle function and the corresponding image $I_\infty(x)$ of the solar limb, as given in Eq. (7). The x-axis is in units of λ/D.
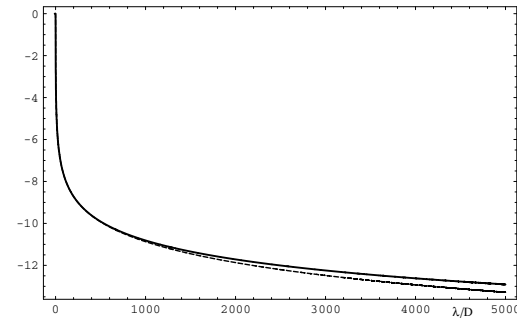

Fig. 5. Comparison on a logarithmic scale between $I_\infty(\alpha)$ (top curve) and $I_d(\alpha)$. To make the effect visible, $D_\odot$ is taken equal to the very low value of 1000 units of λ/D, corresponding to a very small telescope.
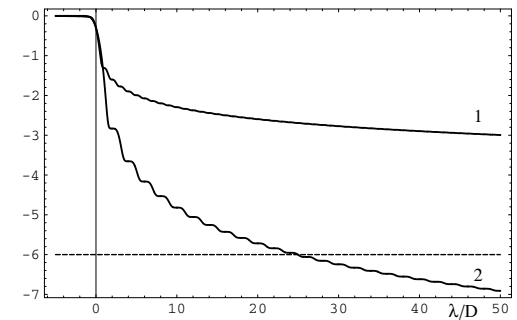

Fig. 6. Representation on a logarithmic scale of the images of the solar limb given by (1) a raw aperture and (2) an apodized aperture in the form a triangle function.

This halo profile is given in Fig. 4. The limiting expression for large α is $\frac{3}{2\pi^4}\alpha^{-3}$. The α^{-3} behavior is the result of the integration of the PSF in α^{-4}. The analytical expression for $I_\odot(\alpha)$ can be obtained easily, but it presents a lot of terms and is of little interest because $I_\odot(\alpha)$ and $I_\infty(\alpha)$ are almost identical for the regions of interest. The difference between the curves can only be seen at very long distances for a rather small telescope, which is what we show in Fig. 5 where the exact curves are drawn for the two models. The reason is that wings of the PSF of Eq. (6) in α^{-4} decrease fast enough so that far away point sources do not contribute to the diffracted light.

A comparison between the apodized and non apodized cases is shown in Fig. 6. The diffracted light for the apodized aperture goes below 10^{-6} for α on the order of 25 λ/D, i.e. very close to the solar limb (a few arc seconds for most telescopes). Of course this assumes perfect optics, a condition that is also mandatory for Lyot coronagraphy.

4. The 2D real case

Now we go back to function $I_\odot(\alpha, \beta)$ of two variables defined by Eq. (1). Since $S_\odot(\alpha, \beta)$ and $R(\alpha, \beta)$ are radial functions, $I_\odot(\alpha, \beta)$ is also a radial function that only depends on the distance $\rho = \sqrt{\alpha^2 + \beta^2}$. It is therefore enough to compute $I_\odot(\alpha, \beta)$ along a central slice such as $\alpha = \rho$ and $\beta = 0$, with the form:

$$I_\odot(\rho) \equiv I_\odot(\alpha = \rho, 0) = \int \int R(\xi, \eta) S_\odot(\alpha - \xi, -\eta) d\xi d\eta \quad (8)$$

where, for convenience, we use the same symbol to define the 2D function and its radial cut.

4.1. The formal problem for the Airy pattern

For a perfect telescope with a circular aperture, the PSF is the Airy pattern, of the form:

$$R(\rho) = \frac{J_1^2(\pi\rho)}{\pi\rho^2} \quad (9)$$

where ρ is in units of λ/D and where we make use of a normalization of the PSF so that its integral equals 1.

For large ρ, $R(\rho)$ decreases as ρ^{-3}, which is much better than the ρ^{-2} decrease of the equivalent 1D example of the previous section, and one may think that the effect of diffraction will be less for the 2D image than for the 1D one. In fact, as we will see, final results are very similar because for the 2D case a 2D integration is needed.

A Sun of uniform brightness corresponds to the function $\Pi(\sqrt{\xi^2 + \eta^2}/D_\odot)$. The 2D integral of Eq. (8) takes the form:

$$I_\odot(\alpha) = \int_{\xi=\alpha}^{\alpha+D_\odot} \left\{ \int_{\eta=-h}^{+h} \frac{J_1^2(\pi\sqrt{\xi^2+\eta^2})}{\pi(\xi^2+\eta^2)} d\eta \right\} d\xi \quad (10)$$

where $h = \sqrt{(\xi - \alpha)(D_\odot + \alpha - \xi)}$, assuming the origins of the axes on the solar limb. An illustration of this integral is schematized in Fig. 7 where the circle, not at all to scale, represents the Sun and the domain of integration. The real size of the Sun can be estimated from a limb contour (almost linear) drawn in the figure.

It is possible to take advantage of the radial shape of the PSF to reduce the 2D integral to a 1D one. Indeed the values of the PSF are the same along circular arcs, such as AMB of Fig. 7. Denoting ρ the distance OM, the length of the arc becomes

$$\widehat{AMB} = 2\rho \arcsin\left(\frac{h'}{\rho}\right) \quad (11)$$

where h' is equal to

$$h' = \frac{\sqrt{4(R_\odot + \alpha)^2\rho^2 - \left((R_\odot + \alpha)^2 - R_\odot^2 + \rho^2\right)^2}}{2(R_\odot + \alpha)}. \quad (12)$$

The reader should note that the quantities h and h' indeed correspond to the same length HB of Fig. 7, one written as a function of ξ in Cartesian coordinates, the other as a function of ρ in polar coordinates.

The double integral then reduces to a simple integral:

$$I_\odot(\alpha) = \frac{2}{\pi} \int_\alpha^{\alpha+D_\odot} \frac{J_1^2(\pi\rho)}{\rho} \arcsin\left(\frac{h'}{\rho}\right) d\rho. \quad (13)$$
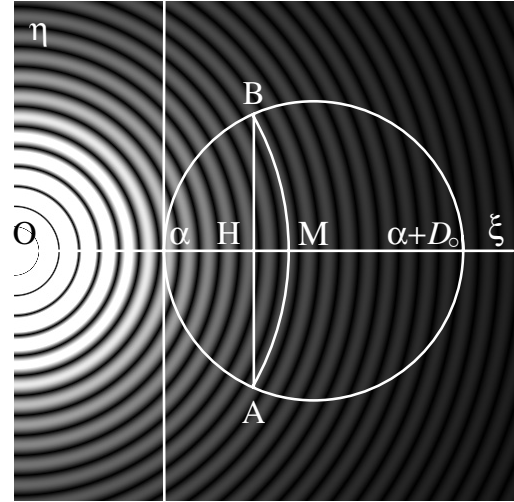

Fig. 7. Schematic illustration of the integral of Eq. (10) giving $I_\odot(\alpha)$. The Airy pattern must be integrated over a circle of the Sun surface. Of course the circle drawn in the figure is not at all to scale (diameter $15\lambda/D$). The strong, almost vertical line corresponds to the edge of a circle with size ($3000\lambda/D$). The length of the arc $\widehat{AMB}$ appears in Eq. (11).

Unfortunately, not one of these integrals leads to an analytical solution, as far as the author has tried, even using its approximate cosine form for the Bessel function. Nevertheless the two expressions may be useful. As we see below, we can derive an approximate analytical solution for the diffraction pattern from the expression of the 2D integral, which is in perfect agreement with a numerical approach to the problem made using the 1D integration.

4.2. Approximate solution via the line spread function

It is possible to obtain an approximation for the integral between braces of Eq. (10) for the limit h and $D_\odot \to \infty$, as the line spread function (LSF) of the circular aperture. The computations made for the 1D case have shown that this approximate result is a good approximation of the exact one. This can be done as follows.

We recall that the telescope modulation transfer function (MTF) is defined as the Fourier transform of the PSF, of the form:

$$M(u, v) = \int \int R(\alpha, \beta) \exp(-2\iota\pi(\alpha u + \beta v)) d\alpha d\beta. \quad (14)$$

Making $v = 0$, we have:

$$M(u, 0) = \int \left\{ \int R(\alpha, \beta) d\beta \right\} \exp(-2\iota\pi\alpha u) d\alpha, \quad \text{or} \quad (15)$$

$$\int R(\alpha, \beta) d\beta = \int M(u, 0) \exp(2\iota\pi\alpha u) du \quad (16)$$

so that $L_\infty(\xi)$ can be computed as the Fourier transform of the cut of the telescope MTF. This function is also the aperture autocorrelation function. For the circular aperture, it is a radial function of the form

$$M(u) = \frac{2}{\pi} \left(\arccos(u) - u\sqrt{1 - u^2} \right). \quad (17)$$

Since $M(u)$ is an even function, the Fourier transform reduces to its cosine term, and we can write

$$L(\xi) = 2 \int_0^1 M(u) \cos(2\pi\xi u) du = \frac{H_1(2\pi\xi)}{2\pi^2\xi^2} \quad (18)$$
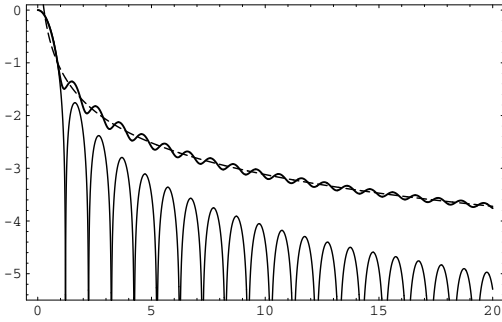
Fig. 8. Comparison between the Airy pattern (bottom curve) and the line spread function of Eq. (18). The curve drawn in dashed lines corresponds to the approximation in ξ^{-2} of Eq. (20).
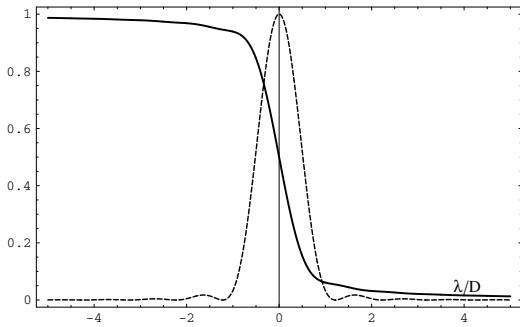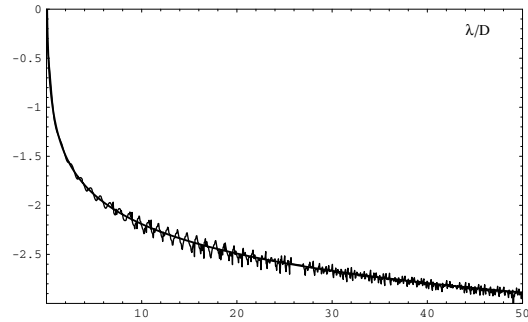

Fig. 10. Numerical vs. analytical curves of the halo. The analytic curve is drawn using Eq. (22). The noisy curve corresponds to a numerical computation of Eq. (13).
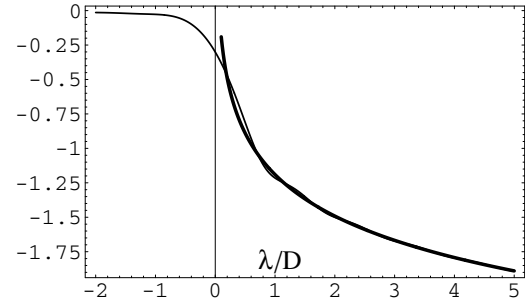

Fig. 9. Cuts of the Airy pattern and the halo given by Eq. (21).

Fig. 11. Representation of the diffracted halo given in Eq. (21) with its simple approximation in α^{-1} of Eq. (23).

where $H_1(z)$ is the Struve function. Using for this function the approximation (Aarts & Weisstein 2004):

$$H_1(z) \sim \frac{\left(12-\frac{36}{\pi}\right)(1-\cos(z))}{z^2} + \frac{\left(\frac{16}{\pi}-5\right)\sin(z)}{z} + \frac{2}{\pi} - J_0(z) \quad (19)$$

we can obtain an asymptotic form for $L(\xi)$ for high ξ values:

$$L(\xi) \sim \frac{3}{4\pi^2\xi^2}. \quad (20)$$

The reader will note that the behavior of the LSF in α^{-2} is that of the simple 1D model of Eq. (3). A representation of $L(\xi)$ in Eq. (18) and its approximation in Eq. (20) is given in Fig. 8, where $L(\xi)$ is smoother and higher than the Airy pattern that is also represented in the figure.

Substituting the expression of $L(\xi)$ of Eq. (18) into the integral between braces of Eq. (10), we have:

$$I_\odot(\alpha) = \int_\alpha^{\alpha+d} \frac{H_1(2\pi\xi)}{2\pi^2\xi^2}d\xi$$
$$= \frac{1}{2} - \frac{8d}{3\pi}\,_2F_3\left(\frac{1}{2},1;\frac{3}{2},\frac{3}{2},\frac{5}{2};-\alpha^2\pi^2\right) \quad (21)$$

where the function $_pF_q(a_1,...,a_p;b_1,...,b_q;x)$ is the generalized hypergeometric function, which is given by a hypergeometric series (see for example Weisstein 2003). It can be numerically computed using *Mathematica*. The halo corresponding to Eq. (21) is represented in Fig. 9, together with a cut of the Airy pattern.

For α values higher than 6 or so, the hypergeometric function $_2F_3\left(\frac{1}{2},1;\frac{3}{2},\frac{3}{2},\frac{5}{2};-\alpha^2\pi^2\right)$ is difficult to compute with enough precision. In that case it is easier to use the expression derived

from the approximation of the Struve function given in Eq. (19) for $I_\odot(\alpha)$. We obtain

$$I_\odot(\alpha) \sim \frac{2\pi^2\alpha^2 + \pi - 3 - (\pi-3)\,_1F_2\left(-\frac{3}{2};-\frac{1}{2},\frac{1}{2};-\pi^2\alpha^2\right)}{\pi^5\alpha^3}$$
$$- \frac{_1F_2\left(-\frac{1}{2};\frac{1}{2},1;-\pi^2\alpha^2\right)}{\pi^2\alpha}$$
$$+ \frac{(16-5\pi)\,_1F_2\left(-\frac{1}{2};\frac{1}{2},\frac{3}{2};-\pi^2\alpha^2\right)}{\pi^3\alpha} + \frac{1}{2} \quad (22)$$

which involves only hypergeometric functions with $p = 1$ and $q = 2$, which can in turn be computed for high values of α. In the range were the two expressions can both be computed numerically (i.e. for example between -5 and $+5$), there is no difference between Eqs. (21) and (22).

Moreover, it is possible to compare the analytic model with the numerical computation of the 1D integral of Eq. (13). The result is given in Fig. 10. Although the numerical expression remains noisy, there is excellent agreement between the two approaches. In fact, the halo is very soon (a few λ/D) represented well by a surprisingly simple expression

$$I_\odot(\alpha) \sim \frac{2}{\pi^3\rho}, \quad (23)$$

which takes the same form as the halo in the 1D model (to a factor $4/\pi$). This expression can be derived as the limit of Eq. (22) for α large. The comparison given in Fig. 11 between this simple model and the exact one of Eq. (21) shows the closeness of the approximation. The approximation, which is fair even for α values of a few λ/D, becomes very good for high values.

The conclusion for the 2D case is the same as for the 1D one. The diffraction halo due to the solar disk decreases very slowly and prevents detection of the corona. To give numbers, it requires
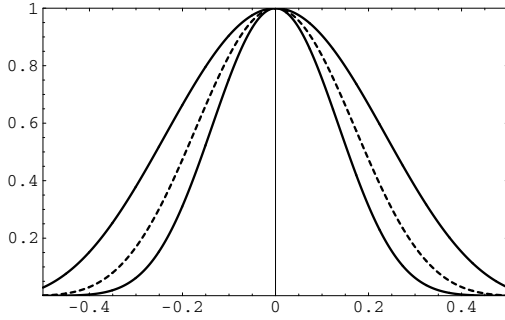
Fig. 12. Examples of generalized prolate apodization functions for circular apertures: radial cuts of the transmission in amplitude, for a telescope with a diameter of unity. The curves are drawn for the prolate parameters c equal to 6 (broadest curve), 10, and 15 (narrowest curve).

a distance of more than 6.4×10^4 units of λ/D to fall below 10^{-6}, or more than 8 solar radii for a 1-m telescope. A way to solve the problem is to use an apodized aperture, which we describe in the next section.

5. Apodization functions for a circular aperture

Many functions can be utilized for the apodization of a circular aperture, some of which have been studied in the literature (see for example the review paper of Jacquinot & Roizen-Dossier 1964). Most of the apodizations tend to change the ρ^{-3} behavior of the Airy pattern into a deeper decrease of the wings, in a way similar to the example shown for the 1D case. For example, the transmission $1 - \rho^2$, known as the Welsh apodization, will produce a PSF with a ρ^{-5} behavior. A transmission function with the form of the 2D aperture autocorrelation function of Eq. (11) is similar to the triangle function in the 1 D case and, mutatis mutandis, this will lead to a ρ^{-6} behavior.

In fact we will propose a very particular kind of transmission curves for the aperture: the generalized prolate spheroidal function (GPSF). Prolate functions and GPSFs were discovered by Slepian & Pollak (1961) and Slepian (1964); the term "generalized" applies to the 2D form of these functions that were first derived for 1D signals. In fact we could have used them in our 1D analysis. The application of GPSFs to optics was reviewed by Frieden (1971), and their particular interest for stellar coronagraphy was shown by Aime et al. (2001 and 2002) for rectangular and circular apertures. In the present analysis the GPSFs of interest are the prolate functions $\Phi_{0,0}(c, \rho)$.

Prolate functions present many interesting properties. They may be considered as the best apodizers since they maximize the encircled energy. They are the only functions invariant to a finite Fourier transform (Gaussian curves have this property in the infinite case). They allow a family of apertures with different apodization strengths depending of the prolate parameter c in a continuous way. The larger c, the more apodized the function.

An example of GPSFs for the aperture transmission in amplitude is given in Fig. 12. The curves are drawn for the prolate parameter c equal to 6, 10, and 15. The GPSFs do not go to zero at the edge of the aperture (which would be the case of many others apodizing functions), even for the strongest apodizations. The transmission of the curves of Fig. 12 at the edge of the aperture are 0.028, 0.00069, and 5.7×10^{-6}. This is also an interesting point, since Jacquinot & Roizen-Dossier (1964) pointed out in another context that a variable transmission going continuously to a zero transmission at some point of the diaphragm cannot be realized in practice by means of absorbing screens.
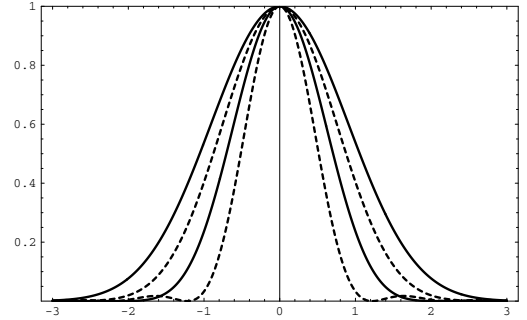

Fig. 13. Central part of the PSFs for the prolate transmissions of Fig. 12 and the raw aperture. From the narrowest curve to the broadest: the Airy pattern, the prolate apodized curves with parameters c of 6 (narrowest curve), 10, and 15 (broadest curve). Due to basic Fourier properties, widths are in the inverse order of the transmissions in Fig. 12.
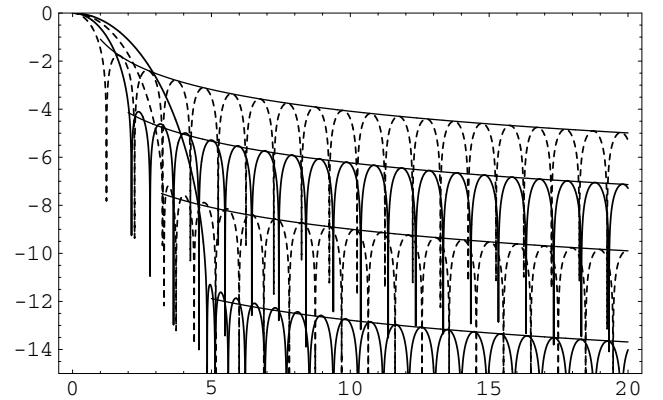

Fig. 14. Representation on a logarithmic scale of the diffraction patterns in the focal plane. From top to bottom, the Airy pattern and the GPSF apodized apertures for the parameter c equal to 6, 10, and 15 (same as in Fig. 13). All curves are normalized to 1 at the origin. Envelopes have the form ρ^{-3}. See Table 1 for attenuation values.

More details on the interest of GPSFs as apodizers for stellar coronagraphy can be found in Aime (2005a) and papers cited in this reference. The numerical computation of GPSFs is non trivial, and we used a special program written by Falloon (Falloon et al. 2003) in *Mathematica* (Wolfram 1999).

Corresponding PSFs are given on a linear scale in Fig. 13 and on a logarithmic scale in Fig. 14. The reader will note that GPSFs have a particularity of great interest for the present study. They behave differently from the other apodizations. The rate of attenuation of the wings remains r^{-3} as for the Airy pattern, while starting at a much lower level. We denote the attenuation factor as $\kappa(c)$ (see Table 1).
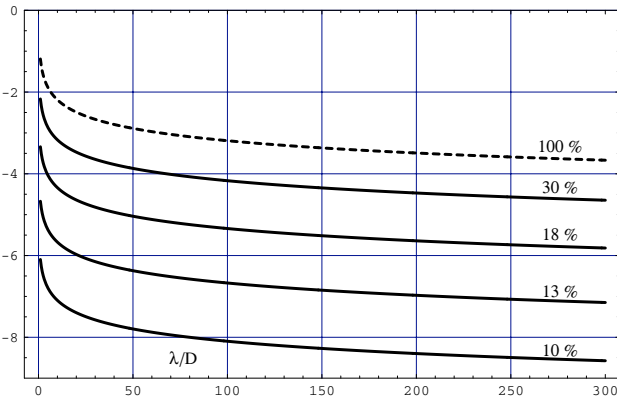
This property permits us to directly use the results derived for the raw aperture for prolate apodized apertures. Indeed, as soon as α becomes larger than the core of the PSF, we can just multiply the Airy pattern with $\kappa(c)$ in Eq. (10) to obtain the result for the GPSFs. Since the integration is linear, Eq. (23) can then be rewritten as

$$I_\odot(\alpha, c) \sim \kappa(c) \frac{2}{\pi^3 \rho}. \qquad (24)$$

We have represented a set of curves $I_\odot(\alpha, c)$ in Fig. 15 for the values c equal to 4, 6, 8, and 10, indicating the throughput in intensity $\Phi_I(c)$ of the corresponding aperture. For comparison, the result for the raw un-apodized telescope is also represented. It clearly appears that, to be efficient, the apodization must be

Table 1. Attenuation $\kappa(c)$, throughput $\Phi_I(c)$ and resolution $\sqrt{\varpi(c)}$ function of the prolate parameter c.

c	$\kappa(c)$	$\Phi_I(c)$	$\sqrt{\varpi(c)}$
0	1	1	1
1	0.9	0.88	0.99
2	0.6	0.64	0.99
3	0.29	0.43	0.95
4	0.105	0.30	0.89
5	0.03	0.229	0.83
6	0.007	0.185	0.77
7	0.0016	0.155	0.72
8	0.00033	0.13	0.68
9	6.4×10^{-5}	0.118	0.64
10	1.2×10^{-5}	0.105	0.61
11	2.3×10^{-6}	0.095	0.58
12	3.8×10^{-7}	0.087	0.56
13	7.0×10^{-8}	0.080	0.54
14	1.2×10^{-8}	0.074	0.52
15	2.0×10^{-9}	0.069	0.50

Fig. 16. The x-axis corresponds to minus the logarithm of $\kappa(c)$, i.e. the logarithm of the gain in dynamic. The continuous line curve corresponds to $\sqrt{\varpi(c)}$, or the loss in resolution. The dashed line curve is the throughput $\Phi_I(c)$ of the apodized aperture. See Table 1 for numerical values.
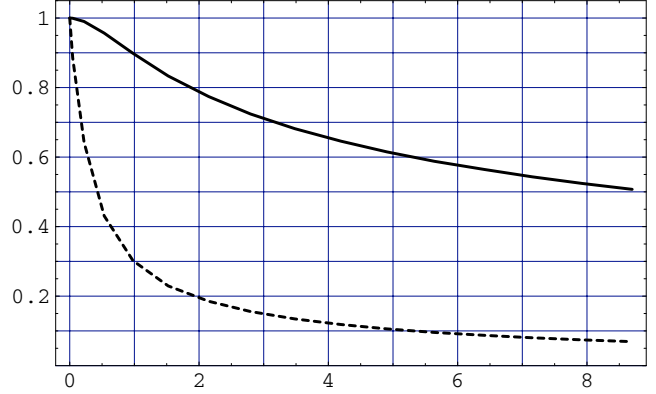

Fig. 15. Representation of $I_\odot(\alpha, c)$ of Eq. (24) for different values of the parameter c; the throughput of the aperture is indicated in percent. The dashed line curve corresponds to the clear aperture with transmission 100%.

strong. This is a result that was already obtained for the detection of exoplanets (Aime 2005a). Indeed, increasing the strength of apodization, i.e. increasing c, there is first a strong decrease in throughput without a very significative reduction of the wings level. Indeed, up to $c \sim 4$, the transmission goes down to 30% for a gain of a factor 10.

The apodization corresponding to $c = 6$, with a throughput of 18.5% (see Table 1) starts to give a rejection factor that is sufficient for seeing the solar corona, but not very close to the solar limb (below 10^{-6} at $460 \lambda/D$). The curve corresponding to $c = 10$, with a throughput of about 10%, should permit the observation of the corona down to the solar limb. The loss in classical resolution of GPSF apodized apertures is not very much. For the critter $\varpi(c)$ based on the equivalent solid angle (angular surface) proposed in Aime (2005a), the resolution for that apodization is down to 0.36. In terms of the usual linear resolution $\sqrt{\varpi(c)}$, it is only a loss of 1.6 times. The main drawback of apodization may be the reduction in throughput, and not the loss in resolution.

6. Signal-to-noise ratio considerations

As seen above, increasing the parameter c (i.e. the apodization) strongly reduces the level of diffracted light. Is there an optimal value for c? One may think that once the background

due to the halo of diffraction has become much lower than the intensity of the corona, there is no need to reduce it further, since by doing so we also lose in throughput. A way to quantify that limit is to perform a signal-to-noise ratio (*SNR*) analysis. We follow a procedure similar to that of Aime (2005a); the quantities called "intensity" thereafter must be expressed in number of photons. We are interested in comparing *SNR*s in a relative sense for apodized vs. un-apodized apertures.

Let us denote as $C(\rho)$ the intensity of the corona observed with a clear un-apodized aperture. All things being equal and neglecting the difference in resolution, the intensity for the apodized aperture is $C_c(\rho) = C(\rho) \times \Phi_I(c)$, where we recall that $\Phi_I(c)$ represents the reduction in throughput due to apodization.

This signal is seen superimposed to the background due essentially to the diffraction halo, with the form $I_\odot(0)I_{\odot,c}(\alpha) \times \Phi_I(c)$, where $I_\odot(0)$ is the Sun brightness at the disk center, plus other unwanted noisy contributions. Their modeling is difficult, so to simplify our notations, we write them in the form $\sim \varepsilon I_\odot(0)\Phi_I(c)$, a model that can easily be modified if necessary.

These quantities represent a bias term that must be subtracted as much as possible in a pre-processing of the image. One might imagine that an un-apodized aperture might permit observation of the solar corona if that subtraction could be done perfectly. For practical reasons (dynamic of detectors, flat-field problem, etc.) that will be impossible. But let us assume that it can be done for the *SNR* analysis. The bias might be suppressed, but the fluctuations of the background remain uncorrected and constitute the noisy term, to which we must add the fluctuations of the number of photons for the signal of the corona itself to be complete. 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. Making use of Eq. (24), we have

$$\sigma^2 = \Phi_I(c)(I_\odot(0)\kappa(c)\frac{2}{\pi^3\rho} + \varepsilon I_\odot(0) + C(\rho)). \tag{25}$$

Writing the *SNR* as the ratio of the signal divided by σ, we have

$$\frac{C_c(\rho)}{\sigma} = \frac{C(\rho)\sqrt{\Phi_I(c)}}{\sqrt{I_\odot(0)\kappa(c)\frac{2}{\pi^3\rho} + \varepsilon I_\odot(0) + C(\rho)}}. \tag{26}$$

The exact expression of the *SNR* will depend on the number of photons collected by the experiment and the value of all terms.
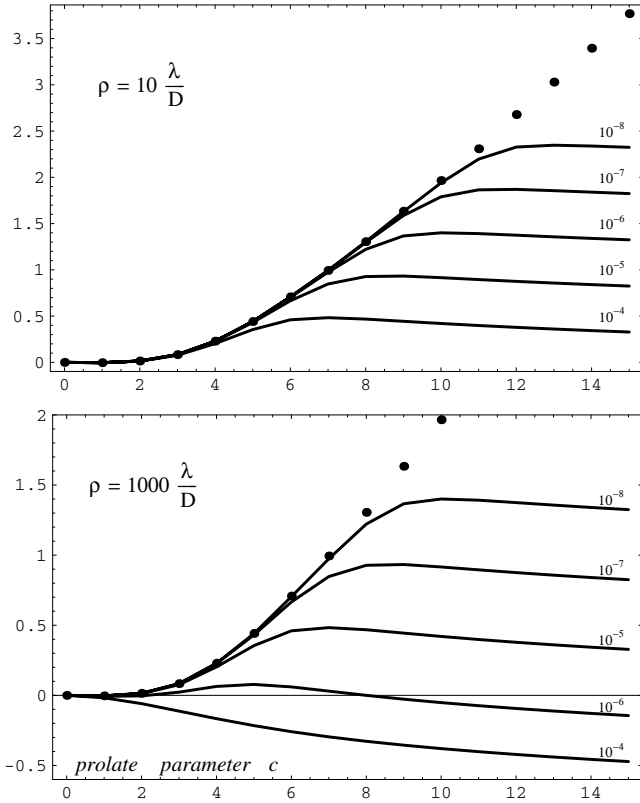
Fig. 17. Representation as a function of c of the $SNR(c,\rho,\varepsilon)$ logarithm of Eq. (27), ρ values of 10 (*top*) and 1000 (*bottom*) λ/D, and different values of ε indicated in the figure. Points correspond to the limit $(\Phi_I(c)/\kappa(c))^{0.5}$ for $\varepsilon = 0$.

A simplified expression can be obtained by comparing the SNR of the apodized aperture to that of the clear one in a relative way. If we assume that $C(\rho)$ is small compared to the other sources of noise, we can suppress it in the denominator, and we obtain the expression of the relative SNR that characterizes the effect of apodization. We have

$$SNR(c,\rho,\varepsilon) \sim \sqrt{\Phi_I(c)}\sqrt{\frac{2 + \varepsilon\pi^3\rho}{2\kappa(c) + \varepsilon\pi^3\rho}}. \tag{27}$$

When $\varepsilon = 0$, the quantity equals the root square of the ratio of $\Phi_I(c)$ over $\kappa(c)$, which increases with the apodization strength (neglecting the term $C(\rho)$ as already indicated). In practice, the SNR will be strongly limited by the term ε, which means that a strong effort to limit the terms of diffusion is mandatory, which is not surprising.

An illustration of the results is given in Figs. 17 and 18 that give different representations of the normalized SNR logarithm of Eq. (27). Interpreting these SNR curves is not straightforward. The value $\log(SNR) = 0$ means that the SNRs are the same for the apodized and clear aperture. Even there, there is a fundamental interest in using an apodized aperture since the computation on the SNR is based on a variance analysis that assumes a suppression of biases for the clear aperture, which we know to be very difficult to realize in practice.

In the SNR of Eq. (27), ε and ρ play the same role. This is clearly illustrated in Figs. 17 and 18. A GPSF apodized aperture will be very efficient for seeing the solar corona very close to the solar limb. Considering the various results of this section and the former one, a GPSF apodized aperture of parameter c on the order of 10 appears to be a good compromise.
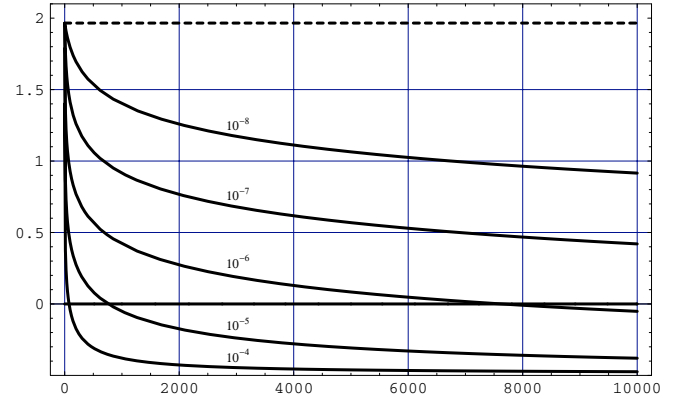

Fig. 18. Representation of the logarithm of $SNR(c,\rho,\varepsilon)$ in Eq. (27) as a function of ρ for $c = 10$ and various values of ε. The horizontal dashed line corresponds to the limit $(\Phi_I(c)/\kappa(c))^{0.5}$ for $\varepsilon = 0$.

7. Conclusion

In this paper we have studied the theoretical form of the diffraction halo produced by the solar disk at the level of the corona using a perfect diffraction-limited telescope. The computation is first made for a 1D model of the problem that permits us to obtain complete analytical expressions for the halo and check the effect of the center-to-limb solar variation. We find that the halo diffracted by the Sun outside the limb $I_\odot(\alpha)$ can be approximated well by a slow decrease in α^{-1}, where α denotes the distance from the limb. Using apodized apertures makes it possible to obtain steeper decreases of the halo. We treat the example of a triangle apodization giving a decrease in α^{-3}.

Formal equations are written for the 2D case, and it is shown that, for a Sun of uniform brightness, the expression may take the form of a 1D integral. Nevertheless, the problem is difficult to solve. An analytic expression can be worked out using the LSF, which is shown to give a valid approximation to the problem. The expression is found in excellent agreement with a numerical computation that uses the exact integral.

As for the 1D case, the diffraction halo for a 2D circular telescope can be approximated well by a simple law in power minus one of the distance (approximately $0.0645\,\rho^{-1}$ in units of λ/D), which makes any direct observation of the solar corona impossible.

As an alternative to the classical Lyot coronagraphy, the use of an apodized entrance aperture may be the solution for imaging the corona. We have proposed to use GPSFs for apodizing functions. These functions are the optimal apodizers in terms of light concentration. They are similar to the Airy pattern, making the analytical approach simpler. For a perfect experiment, the diffraction halo is lowered by a factor $\kappa(c)$ that can give the required reducing factor, as shown in Table 1. To be efficient, the apodization must be strong, and we may expect a factor 10^5 for an apodized aperture of throughput 10%.

The performances given in this paper are for a perfect apodized aperture. In our signal-to-noise analysis, we estimated the effect of a constant background, due for example to a diffusion by the surfaces of the optics. Numerical simulations could be made to check the effects of departure for the ideal condition of all the optical system (apodization and optics). The best way to test the final performances of the instrument would be to realize a demonstrator, a project that we are now considering in our laboratory.

A difficult point will certainly be to realize the apodization itself. One possibility is to use an entrance window reflecting the

light so that the telescope aperture matches the required prolate transmission. A mandatory point is that the apodizing window be the first one the light encounters. Re-imaging the aperture for apodization is unsatisfactory. Indeed, as we have seen, an unapodized aperture produces a huge diffraction halo. An imaging system will necessarily truncate it, and cause other sources of unwanted diffraction. Moreover that will reduce the advantage of apodization alone compared to Lyot coronagraphy, i.e. that apodization does not require optics with a large field of view. This is certainly a very important point for coronagraphs with large entrance apertures.

The drawbacks of apodization are the poor aperture transmission, in the 10% range, and a loss of resolution, roughly a factor 1.6 compared to the theoretical diffraction limit of the telescope. These points may be compensated for by using a larger aperture, if necessary. Nevertheless, we have shown that GPSF apertures give better *SNR*s than do clear apertures for imaging the corona.

Now it would be very interesting to compare apodization techniques to the classical Lyot coronagraph. A formal comparison between the two techniques remains to be done, and we intend to study that in a future work. A precise determination of the effect of Lyot coronagraphy is much more difficult to obtain than the results reported in the present paper. For each point source of the solar disk we will have to express the remaining intensity through the Lyot coronagraph and sum all contributions coming from the entire disk, as we have done here. The problem is to derive an analytical expression for Lyot coronagraphy, in the conditions of a huge Lyot mask not centered on the point source, a problem never treated in the exoplanet domain. One can predict that the residual contributions will come primarily from the edge close to the Sun.

The question of using a Prolate Apodized Lyot Coronagraph (PALC), i.e. a Lyot coronagraph with an entrance apodized aperture, may also be considered. We might think in the direction of that technique, but with a negative a priori. Indeed, in the PALC technique, the aperture apodization parameter c and the Lyot mask must match. Relation (5) of Aime (2005b) shows that the parameter c is $\pi/2$ times the angular Lyot mask diameter divided by λ/D, which would correspond to unreasonable c values and related apodizations.

References

Aarts, R. M., & Weisstein, E. W. 2004, From MathWorld–A Wolfram Web Resource, http://mathworld.wolfram.com/StruveFunction.html
Aime, C. 2005a, A&A, 434, 785
Aime, C. 2005b, PASP, 117, 1012
Aime, C., & Soummer, R. 2004, ApJ, 612, L85
Aime, C., Soummer, R., & Ferrari, A. 2002, A&A, 389, 334
Falloon, P. E., Abbott, P. C., & Wang, J. B. 2003, J. Phys. A, 36, 5477
Frieden, B. R. 1971, Progress in Optics, ed. E. Wolf, 9
Jacquinot, P., & Roizen-Dossier, B. 1964, in Progr. Opt., 3, 31
Kasdin, N. J., Vanderbei, R. J., Spergel, D. N., & Littman, M. G. 2003, ApJ, 582, 1147
Lyot, B. 1932, Z. Astrophys., 5, 73
Nisenson, P., & Papaliolios, C. 2001, ApJ, 548, L201
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., & Falloon, P. E. 2003, A&A, 397, 1161
Vanderbei, R. J., Spergel, D. N., & Kasdin, N. J. 2003, ApJ, 590, 593
Weisstein, E. W. 2003, From MathWorld–A Wolfram Web Resource, http://mathworld.wolfram.com/GeneralizedHypergeometric-Function.html
Wolfram, S. 1999, The Mathematica Book, Fourth edition (Cambridge University Press), http://integrals.wolfram.com/index.jsp