Available online at www.sciencedirect.com
ScienceDirect
C. R. Physique 8 (2007) 277–287
COMPTES RENDUS PHYSIQUE
http://france.elsevier.com/direct/COMREN/

Optical techniques for direct imaging of exoplanets/Techniques optiques pour l'imagerie directe des exoplanètes

An introduction to stellar coronagraphy

André Ferrari [a], Rémi Soummer [b,1], Claude Aime [a,*]

[a] *LUAN, université de Nice Sophia-Antipolis, parc Valrose, 06108 Nice cedex 2, France*
[b] *American Museum of Natural History, 79th St. at Central Park West, New York, NY 10024, USA*

Available online 19 June 2007

Abstract

This article gives a simple and original presentation of various coronagraphs inherited from the Lyot coronagraph. We first present the Lyot and Roddier phase mask coronagraphs and study their properties as a function of the focal mask size. We show that the Roddier phase mask can be used to produce an apodization for the star. Optimal coronagraphy can be obtained from two main approaches, using prolate spheroidal pupil apodization and a finite-size focal mask, or using a clear aperture and an infinite mask of variable transmission. ***To cite this article: A. Ferrari et al., C. R. Physique 8 (2007).***

Résumé

Introduction à la coronographie stellaire. Cet article propose une présentation simplifiée et originale des divers coronographes issus du coronographe de Lyot. Le coronographe de Lyot et celui de Roddier à masque de phase sont d'abord présentés. Leurs comportement est étudié en fonction de la taille du masque focal. Il est ainsi mis en évidence que le coronographe de Roddier peut agir comme un apodiseur pour l'étoile. On montre ensuite qu'une coronographie ayant des propriétées optimale peut être obtenue avec une taille de masque finie en apodisant la pupille par une fonction prolate ou avec un masque de transmission variable mais de taille infinie. ***Pour citer cet article : A. Ferrari et al., C. R. Physique 8 (2007).***

Keywords: Coronagraphy; Exoplanets

Mots-clés : Coronagraphie ; Exoplanètes

1. Introduction

Direct imaging of faint companions or planets around a bright star is a very difficult task, where the contrast ratio and the angular separation are the observable parameters. The problem consists of detecting a faint source over a bright and noisy background, mainly due to the diffracted stellar light. High contrast ratios and small angular separations correspond to the most difficult case. Typically, for giant exoplanets, contrast ratios of about 10^{-7} are expected in the near infrared (J;H;K bands), based on models for relatively young objects of about 100 Myr [1–3]. According to these

* Corresponding author.
E-mail addresses: andre.ferrari@unice.fr (A. Ferrari), claude.aime@unice.fr (C. Aime).
[1] R. Soummer is supported by a AMNH Kalbfleisch Research Fellowship.

1631-0705/$ – see front matter
doi:10.1016/j.crhy.2007.05.008

models, older objects would be an order of magnitude fainter. Terrestrial planets are much fainter than giant planets, about 3 to 4 orders magnitudes fainter depending on the wavelength range.

The aim of this article is to give a short review of basic concepts and techniques used in focal plane mask coronagraphy for exoplanets detection. The paper will focus on derivation of a general formalism which allows one to gain deeper insight in the behaviour of the principal coronagraphic techniques. It will not consider important problems related, for example, to the effects of adaptive optics residual speckles, slow-varying speckles caused by mechanical or thermal deformations, telescope central obstruction, chromaticity.

2. Coronagraph general formalism

A common formalism can be used for Lyot and Roddier and Roddier coronagraphs [4,5] in their classical or apodized version [6–8], four quadrants coronagraph (4QC) [9], band-limited mask coronagraph [10] and shaped pupil coronagraph [11].

We will follow the notations of [6]. The successive planes of the coronagraph are denoted by A, B, C and D. A is the entrance aperture, B denotes the focal plane with the image mask, C is the image of the aperture with the Lyot stop and D is the image in the focal plane after the coronagraph. The general setup is illustrated in Fig. 1. The position vector in each plane of the coronagraph will be noted in bold.

- We denote by $p(\mathbf{x})$ the telescope pupil and $a(\mathbf{x})$ its possible transmission if a pupil apodization is used. We assume that $\mathbf{x}$ in A is in units of wavelength.
- The mask transmission is $t(\mathbf{x})$. We will consider without loss of generality that $t(\mathbf{x}) = 1 - m(\mathbf{x})$, where $m(\mathbf{x})$ is the image mask. The units for the coordinates in B and D are angles on the sky.
- The Lyot stop, denoted as $p_s(\mathbf{x})$ will clearly acts as a low pass filter and consequently its size will be chosen at most equal to the size of the entrance pupil so that $p_s(\mathbf{x})p(\mathbf{x}) = p_s(\mathbf{x})$.

We denote as $\mathcal{P}$, $\mathcal{S}$ and $\mathcal{M}$ the areas defined by the pupil, the stop and, if pertinent, the mask. We will make the usual approximations of paraxial optics. Moreover we neglect the quadratic phase terms associated with the propagation of the waves or assume that the optical layout is properly designed to cancel it [12]. The constant propagation terms between the successive planes will be omitted. In this case the coronagraph can be easily described using classical Fourier optics: a Fourier transform exists between each of the four planes.

The wavefront complex amplitude just before the pupil plane is $\Psi(\mathbf{x})$. The expression of the complex amplitude in the successive planes A, B, C and D are:

$$\Psi_A(\mathbf{x}) = \Psi(\mathbf{x})p(\mathbf{x})a(\mathbf{x}) \tag{1}$$

$$\Psi_B(\mathbf{x}) = t(\mathbf{x})\hat{\Psi}_A(\mathbf{x}) = \hat{\Psi}_A(\mathbf{x}) - m(\mathbf{x})\hat{\Psi}_A(\mathbf{x}) \tag{2}$$

$$\Psi_C(\mathbf{x}) = \hat{\Psi}_B(\mathbf{x})p_s(\mathbf{x}) = \Psi_A(\mathbf{x}) - \big(\Psi_A(\mathbf{x}) * \hat{m}(\mathbf{x})\big)p_s(\mathbf{x}) \tag{3}$$

$$\Psi_D(\mathbf{x}) = \hat{\Psi}_C(\mathbf{x}) = \hat{\Psi}_A(\mathbf{x}) - \big(\hat{\Psi}_A(\mathbf{x})m(x)\big) * \hat{p}_s(\mathbf{x}) \tag{4}$$

where $\hat{g}$ is the Fourier transform of g and $*$ denotes convolution.

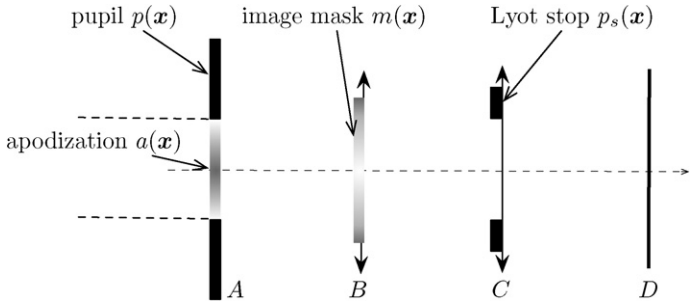

Fig. 1. Schematic illustration of Lyot, Roddier and Roddier, 4QC and band-limited mask coronagraph.

The effect of the coronagraph clearly appears in Eq. (3). The first term is the direct wave corresponding to the entrance pupil. The second term corresponds to the wave diffracted by the mask for which the light diffracted outside the aperture in C has been removed. A coronagraph correctly designed for exoplanets imaging can operate one of these two techniques:

(i) cancel the on-axis star without altering an off-axis source: for the star the two terms in (3) must interfere destructively whereas for the planet the second term of (3) must cancel. In this case the response of the coronagraph is no longer invariant by translation;
(ii) concentrate the on-axis star light reducing the off-axis diffracted light. This approach corresponds to the apodization techniques [12].

These two techniques can be achieved by a proper choice of the apodization $a(\mathbf{x})$ and the transmission $t(\mathbf{x})$. This article will focus on the first solution. For a general overview of pure apodization techniques see [12] and included references.

Section 2 will present the main solutions to this problem that have been derived in the literature. First the 'historical' Lyot coronagraph and the Roddier coronagraph are presented. Whereas the first attempts to cancel the star light, it appears that the second one can also operate by an apodization of the star light. Whereas these coronagraphs can provide acceptable results when the dynamic is not too high, their performance appears to be insufficient for the detection of Earth-like planets. Section 3 addresses the problem of optimal coronagraphy which try to maximize a criterion quantifying the star light rejection. In this context a complete star light rejection can be obtained using an infinite size mask (4QC and band-limited mask) or a finite size π phased mask (Roddier coronagraph) with a properly apodized entrance pupil. In the case of a Lyot coronagraph, an apodized aperture allows us to maximize the star light rejection for a given mask size. Through all this presentation we will try to be, when possible, as general as possible with respect to the geometry of the system without specializing, for example, on a specific pupil shape.

3. Lyot and Roddier coronagraphs

The first solution was proposed by Lyot [4] and consists in cancelling the major contribution of the star energy located in the telescope point spread function (PSF) inside a disk of radius r_m. This is simply achieved setting for $t(\mathbf{x})$ an opaque mask at the center of the image plane B.

The results will be presented for a circular aperture of radius R and without central obstruction. The use of polar coordinates will be preferred. Under this assumption $p(\mathbf{x}) = \Pi(r/R)$, $m(\mathbf{x}) = \Pi(r/r_m)$ and $p_s(\mathbf{x}) = \Pi(r/r_s)$ where $\Pi(r) = 1$ for $r \in [0, 1)$ and 0 if $r \geqslant 1$ and for example: $\hat{p}(\mathbf{x}) = Rr^{-1}J_1(2\pi Rr)$. Note that as long as $\Psi_A(\mathbf{x})$, $a(\mathbf{x})$ and $m(\mathbf{x})$ are radial functions, the complex amplitudes at the different stages of the coronagraph will exhibit the same symmetry.

Fig. 2 illustrates the response of the Lyot coronagraph to an on-axis point source ($\Psi(\mathbf{x}) = 1$) for different values of the mask radius r_m. These values of r_m are located on $\hat{\Psi}_A(r)$ in Fig. 3.

We seek to obtain the best subtraction of the two wavefronts $\Psi_A(r)$ and $\Psi_A(r) * \hat{m}(r)$ inside the Lyot stop. Fig. 2 clearly shows that this result is achieved increasing the mask size. In fact, as r_m increases, $\hat{m}(\mathbf{x})$ will be more 'concentrated' around the origin. As long as $\hat{m}(\mathbf{x})$ always verifies $\int \hat{m}(\mathbf{x})\,\mathrm{d}\mathbf{x} = m(\mathbf{0}) = 1$ we have $\hat{m}(\mathbf{x}) \to \delta(\mathbf{x})$ as $r_m \to +\infty$, which is the neutral element for the convolution product. On the other hand, as r_m decreases $\hat{m}(r)$ widens around $\hat{m}(0) = \pi r_m^2$. At the limit $r_m \to 0$, $\hat{m}(\mathbf{x}) \to \pi r_m^2$ and $\Psi_A(\mathbf{x}) * \hat{m}(\mathbf{x})$ tends to the constant $\pi^2 R^2 r_m^2$, which does not subtract to $\Psi_A(\mathbf{x}) = 1$. It is interesting to note that this problem is formally equivalent to the problem of digital low-pass filter design with finite impulse response using the windowing method where the infinite impulse response of the ideal filter is truncated using a window, see for example [13].

The results of Fig. 2 suggest the following comments:

– The performance of the coronagraph is not strictly linear with the mask size. The choice of mask b is as an evidence more appropriate than mask c. This is confirmed by the general shape of the second plot in Fig. 3 which gives the integrated intensity in C for a Lyot stop of radius R.
– A large amount of the residual star energy is located at the edges of the pupil. Consequently a moderate reduction of the diameter of the Lyot stop radius r_s will provide a significant gain in starlight rejection with a reasonable

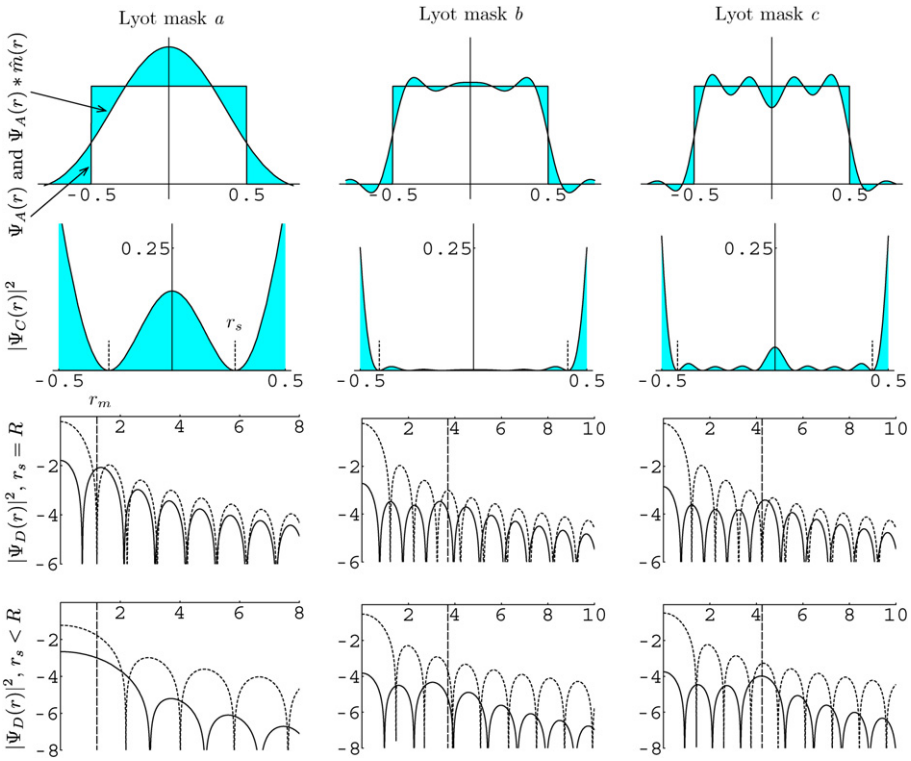

Fig. 2. Illustration of Lyot coronagraphy for circular aperture (radial cuts) and an on-axis point source. The first row represents the two complex amplitudes that should interfere destructively in C. The second row gives the intensity in C and the third and fourth raw the intensity in D. In the two last rows, the continuous line represents the intensity in D for a coronagraphed point source and the dashed line the PSF associated to a pupil of radius r_s. The vertical dashed line in the last row indicates the radius of the mask.

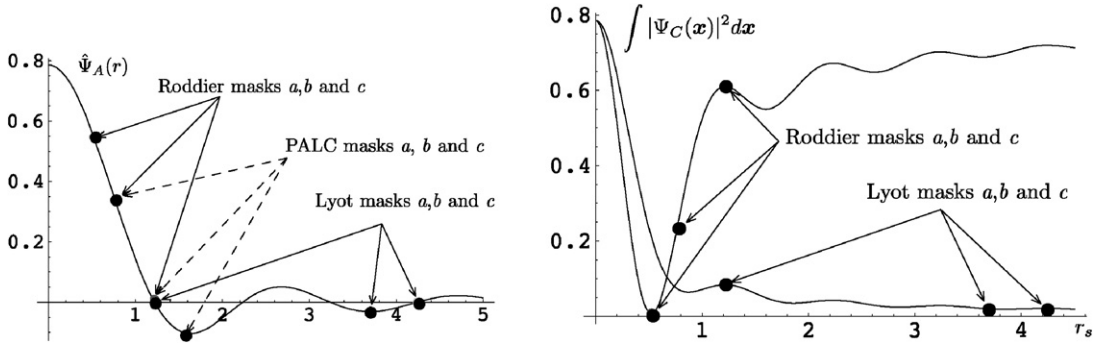

Fig. 3. The left plot locates the different masks used in the simulation on the response of an on-axis point source. The right plot gives the integrated intensity in C for $r_s = R$ as a function of the mask size.

loss of transmission. The effect of a Lyot stop reduction is visible in the last two lines of the plot. The value of r_s for the last line is indicated in the plots of the intensity in C by a dashed line.

– The off-axis transmission of the coronagraph is of course a crucial point for planet detection. As long as the angular distance from the optical axis is sufficient to guaranty that the response of the planet in B will not interfere with the mask, we can consider that the response of the planet will be the shifted PSF of a pupil with radius r_s. In order to quantify the loss of transmission for an off-axis source, this PSF has been added in Fig. 2. Finally, note that a solution to the analytical computation of $\Psi_C(r,\theta)$, $\Psi_D(r,\theta)$ of a point source which can be close to the optical axis has been addressed in [14].

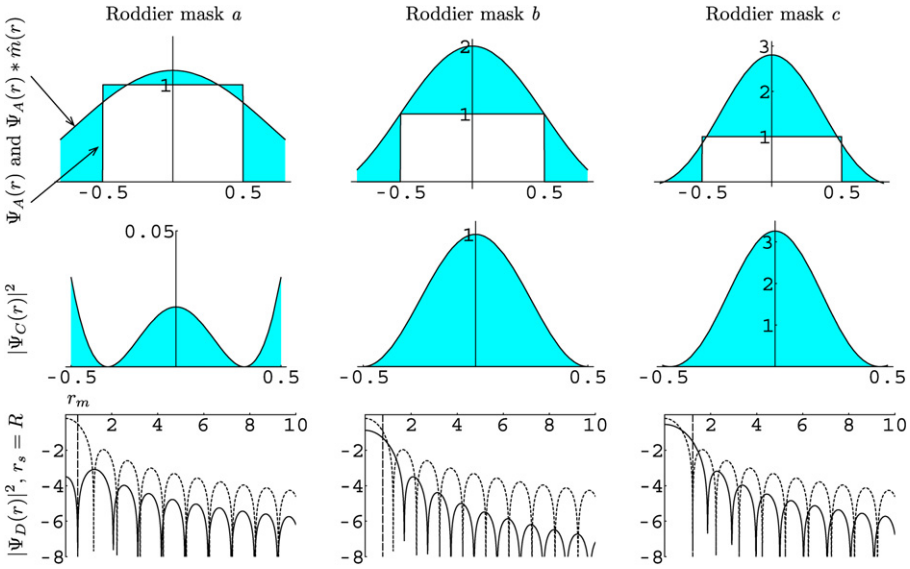

Fig. 4. Illustration of Roddier coronagraphy for circular aperture (radial cuts) and an on-axis point source. The first row represents the two complex amplitudes that should interfere destructively in C. The second row gives the intensity in C and the third raw the intensity in D. In the last row, the continuous line represents the intensity in D for a coronagraphed point source and the dashed line the PSF associated to a pupil of radius r_s.

The Lyot technique was improved by Roddier [5] replacing the opaque mask by a π phase mask: $m(\mathbf{x}) = 2\Pi(r/r_m)$. Fig. 4 show that for very small values of r_m, the Roddier coronagraph tends to behave like the Lyot coronagraph, trying to cancel out the light of the star. This behavior totally differs when r_m increases: for mask size a or b the effect of this coronagraph is to produce in the plane C an apodized version of the wavefront originated by an on-axis source. Note that a coronagraphic technique relying on a similar principle was proposed in [15] using a $\pi/2$ phase mask and a defocus. The effect of this apodization is visible in D: the main lobe of the response is broadened whereas the side lobes decrease. Moreover, it is worthy to note that this apodization has a maximum value for $\mathbf{x} = 0$ that can be higher than 1.

These results suggest the following comments:

- The choice of r_m remains a crucial point. However, according to the preceding remark the integrated energy in C plotted in Fig. 3 is no longer a valid criterion for selection of an optimal value of r_m.
- As the second row of Fig. 4 shows, a reduction of the Lyot stop is not crucial as long as the Roddier coronagraph operates in its 'apodization mode'.
- The effect of a Roddier coronagraph on a planet relies on the same principle as the Lyot coronagraph. Consequently the apodization effect mentioned above will not operate for an off-axis source. Moreover, the fact that for this latter coronagraph the mask will be smaller and a reduced stop is not necessary will increase its performances in terms of resolution and planet flux.

According to the previous remarks, a Roddier coronagraph with an extended mask can be used as an angular selective apodizer for the star's wavefront complex amplitude. The star rejection of such a system being insufficient for planet detection such a device should be considered as the first stage of a 'classical' coronagraph.

4. Optimal coronagraphy

According to Eq. (3) the ideal coronagraph equation is:

$$\Psi_A(\mathbf{x}) = \big(\Psi_A(\mathbf{x}) * \hat{m}(\mathbf{x})\big)\, p_s(\mathbf{x}) \tag{5}$$

where $\Psi_A(\mathbf{x}) = a(\mathbf{x})\,p(\mathbf{x})$. Following Lyot and Roddier coronagraphs, a first solution is to set $m(r) = \varepsilon\Pi(r/r_m)$ where $\varepsilon = 1$ for Lyot and $\varepsilon = 2$ for Roddier and to find, if it exists, the optimal apodization $a(\mathbf{x})$ solution of the integral

equation (5). This solution was first proposed by [16] and [17]. The other solution is to suppress the constraint on the mask shape and to find the function $m(\mathbf{x})$ solution of Eq. (5) when $a(\mathbf{x}) = 1$. This approach includes the band-limited masks proposed by [10] and the four quadrants coronagraph [9].

4.1. Finite size masks: apodized Lyot and Roddier coronagraphs

As long as $m(\mathbf{x})$ has a bounded symmetric support $\mathcal{M}$, we can write using the notations of Appendix A: $\hat{m}(\mathbf{x}) = \varepsilon K_M(\mathbf{x})$. Consequently, Eq. (A.3) implies that if $a(\mathbf{x})$ is proportional to a spheroidal prolate function associated to $\mathcal{M}$ and $\mathcal{P}$, the residual in C when $p_s(\mathbf{x}) = p(\mathbf{x})$ is obtained replacing in (3) the second term by $\varepsilon\Lambda\Psi_A(\mathbf{x})$:

$$\Psi_C(\mathbf{x}) = (1 - \varepsilon\Lambda)\Psi_A(\mathbf{x}) \tag{6}$$

which states that the residual wavefront in C originated by an on-axis source is proportional to the apodized entrance pupil. If we assume that the angular distance of the planet is sufficient, its response in C is obtained setting $\varepsilon = 0$ in (6). The tradeoff between the reduction of the planet flux by the apodization and the reduction of the star flux by the coronagraph can be evaluated by the quantity η defined as the ratio between the total planet intensity in D and the total star intensity outside the mask in D.

From Eq. (6) the star intensity in D is simply $(1 - \varepsilon\Lambda)^2|\hat{\Psi}_A(\mathbf{x})|^2$. The total energy outside the mask is:

$$(1 - \varepsilon\Lambda)^2\left(\int_{\mathbb{R}^2} \left|\hat{\Psi}_A(\mathbf{x})\right|^2 d\mathbf{x} - \int_{\mathcal{M}} \left|\hat{\Psi}_A(\mathbf{x})\right|^2 d\mathbf{x}\right) = (1 - \varepsilon\Lambda)^2\left(\int_{\mathbb{R}^2} \left|\hat{\Psi}_A(\mathbf{x})\right|^2 d\mathbf{x} - \Lambda \int_{\mathbb{R}^2} \left|\hat{\Psi}_A(\mathbf{x})\right|^2 d\mathbf{x}\right) \tag{7}$$

where the last equality comes from the fact that $\varepsilon\Lambda$ is the ratio between the energy encircled in the bounded frequency domain $\mathcal{M}$ and the total energy of $\Psi_A(\mathbf{x})$. As a consequence we have:

$$\eta = (1 - \varepsilon\Lambda)^2(1 - \Lambda) \tag{8}$$

The objective will be to minimize η for a given coronagraph type ($\varepsilon = 1$ or 2). This is achieved by a modification of the size of $\mathcal{M}$ which will affect Λ and the pupil apodization shape $a(\mathbf{x})$.

Any eigenvalue of the integral equation minimizing (8) leads to a valid solution. However, we will only consider in the sequel the largest eigenvalue, i.e. corresponding to the maximum encircled energy behind the focal plane mask. This choice will in fact 'generally' lead to a positive prolate function which will be normalized by its maximum value on the pupil in order to achieve an apodization with maximum transmission. Finally, it is worthy to note that this development does not require a specific shape of the pupil and the mask. However, further analytical derivations based on Eqs. (A.4) and (A.5) require that the mask is a scaled version of the pupil (in this case if the pupil has a central obstruction $m(\mathbf{x})$ has a hole). This point will not be considered here.

– We consider first the case of an apodized Roddier coronagraph, $\eta = (1 - 2\Lambda)^2(1 - \Lambda)$. The results of Appendix A prove that $1/2$ is a valid eigenvalue and consequently a mask size can be chosen to achieve $\eta = 0$, i.e. a total extinction of the star.
  For example, in the case of a circular pupil without obstruction, Fig. 8 in Appendix A shows that the coefficient $c = r_m R \approx 0.25$ corresponds to $\Lambda = 1/2$. This value defines the apodization shape and fixes the mask size to $r_m \approx 0.25/R$. Fig. 5 gives the corresponding apodization. This figure also illustrates the loss in transmission for an off axis planet comparing the prolate apodized pupil PSF with the PSF of the unapodized pupil.

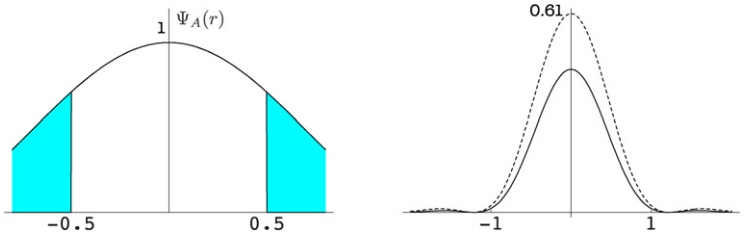

Fig. 5. Illustration of the Roddier apodized coronagraphy for circular aperture (radial cuts). The left plot shows the optimally prolate apodized pupil. The right plot shows the PSF corresponding to the prolate apodized pupil and the PSF of the unapodized pupil.

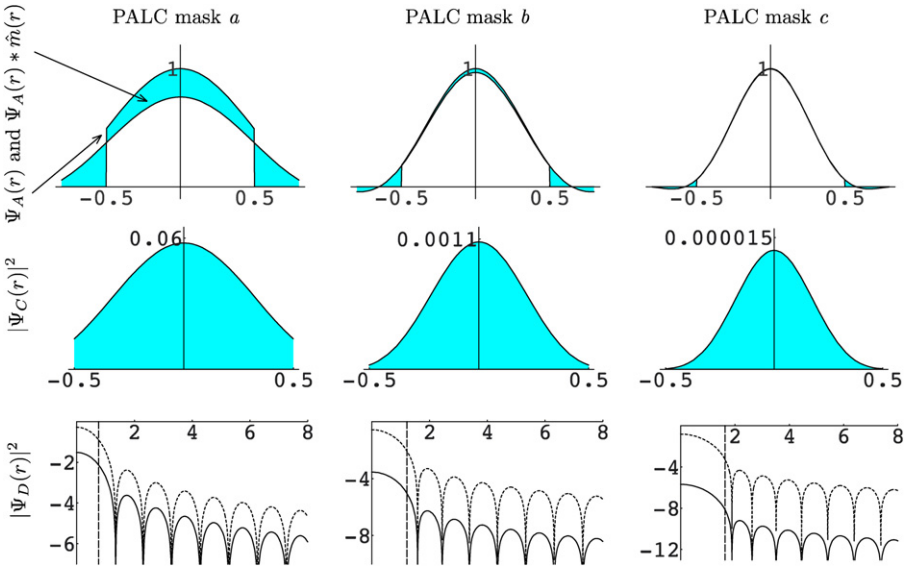

Fig. 6. Illustration of the prolate apodized coronagraphy for circular aperture (radial cuts) and an on-axis point source. The first row represents the two complex amplitudes that should interfere destructively in C. The second row gives the intensity in C. In the last row, the continuous line represents the intensity in D for a coronagraphed point source and the dashed line the PSF associated to the apodized telescope pupil.

– We consider now a prolate apodized Lyot coronagraph (PALC), $\eta = (1 - \Lambda)^3$. The eigenvalues being upper-bounded by 1, a prolate apodized Lyot coronagraph (PALC) cannot achieve total extinction of an on-axis point source. The trivial solution $\Lambda = 1$ would correspond to an infinite size opaque mask. However, approximate solutions can be obtained for eigenvalues Λ close to 1 and finite mask size. Taking advantage of the rapid saturation of the eigenvalue curve we can choose a mask size corresponding to an eigenvalue close to 1.
Fig. 6 illustrates the response of the PALC to an on-axis point source for different values of the mask radius r_m given in Fig. 3. To each value of r_m corresponds a coefficient c and an optimal apodization function given in the first row. The second row gives the residue in C and the last row the response of the star and the PSF associated to the apodized telescope pupil. Note that contrarily to the unapodized Roddier coronagraph operating in its 'apodized regime' (see columns 2 and 3 of Fig. 4) the residue in C is here both apodized *and* attenuated. An analog comment can be done for the apodization alone coronagraphs.
The Lyot stop wave amplitude is proportional to the entrance apodized pupil amplitude and creates the possibility for multiple stage coronagraphs described in [18]. A multiple stage PALC only requires a single apodizer in the entrance pupil. The Lyot stop plane is naturally apodized and can be used as the entrance pupil of a second coronagraphic stage without further loss of throughput due to apodization. In this case the residual amplitude in the second Lyot stop plane is $(1 - \Lambda)^2 \Psi_A(\mathbf{x})$ and the parameter η becomes $(1 - \Lambda)^5$.

4.2. Infinite size masks band limited and four quadrants coronagraphs

A simple solution to define a mask $m(\mathbf{x})$ that verifies (5) is to chose $m(\mathbf{x})$ such that $\hat{m}(\mathbf{x})$ is nonzero only on a small bounded region $\mathcal{B}$ included inside the pupil. In this case if we define $\mathcal{S}$ such as $\forall \mathbf{x}_0 \in \mathcal{S}$ the support of $\hat{m}(\mathbf{x}_0 - \mathbf{x})$ is strictly included in the pupil $\mathcal{P}$:

$$\forall \mathbf{x}_0 \in \mathcal{S}, \quad \Psi_A(\mathbf{x_0}) * \hat{m}(\mathbf{x}_0) = \int_{\mathcal{P}} \hat{m}(\mathbf{x}_0 - \mathbf{x})\, \mathrm{d}\mathbf{x} = m(\mathbf{0}) \tag{9}$$

Consequently, if we impose $m(\mathbf{0}) = 1$ we will clearly verify Eq. (5) on $\mathcal{S} \subset \mathcal{P}$. Fig. 7 illustrates the definition of the various sets $\mathcal{S}$, $\mathcal{P}$ and $\mathcal{B}$. A simple solution to block the residue in $\mathcal{P} \backslash \mathcal{S}$ is to use as a Lyot stop the indicator function of $\mathcal{S}$, $p_s(\mathbf{x}) = \mathbb{1}_{\mathcal{S}}(\mathbf{x})$ ($p_s(\mathbf{x}) = 1$ if $\mathbf{x} \in \mathcal{S}$ and 0 elsewhere). Consequently a properly defined band limited,

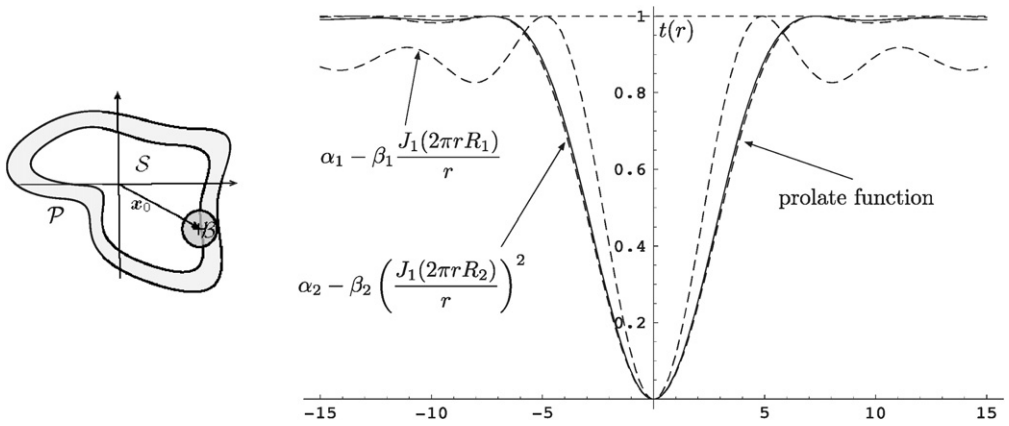

Fig. 7. Left: relation between the pupil ($\mathcal{P}$), support of the mask Fourier transform ($\mathcal{B}$) and the Lyot stop $\mathcal{S}$ in a band-limited coronagraph. Right: illustration of three different masks. Dashed lines: second and fourth order Bessel cardinal masks. Continuous line: prolate spheroidal function.

and necessary infinite size mask $m(\mathbf{x})$ must verify: $m(\mathbf{0}) = 1$, $0 \leqslant m(\mathbf{x}) \leqslant 1$ and $m(\mathbf{x})$ is band-limited on $\mathcal{B} \subset \mathcal{P}$. This approach was proposed by [19].

A fundamental result for understanding of band limited coronagraphs states that the complex amplitude in D associated to an off-axis point source at an angle $\mathbf{x}_1$ is the response of the Lyot stop attenuated by the transmission at $\mathbf{x}_1$:

$$\Psi_D(\mathbf{x}) = t(\mathbf{x_1})\,\hat{p}_s(\mathbf{x} - \mathbf{x}_1) \tag{10}$$

This result is demonstrated in [19]. This result suggests the following comments:

(i) The loss of throughput of a band-limited coronagraph is principally due to the undersizing of the Lyot stop. Consequently $\mathcal{B}$ should be as small as possible. However the general consequences of a reduction of $\mathcal{B}$ are a widening of the central lobe of $m(\mathbf{x})$ and an increase of the ripples in the tails of $m(\mathbf{x})$. From Eq. (10) the first effect implies an increase of a blind zone around the star and the second a reduction of detectability for given locations of a planet.

(ii) The behavior of the coronagraph when the star is slightly off-axis is perfectly described by Eq. (10): in order to reduce the effect of a misalignmemt $m(\mathbf{x})$ must be as 'flat' as possible in $\mathbf{0}$. This effect is quantified in [19] by the degree of the first term in the series expansion of $t(\mathbf{x})$. Note that the flatness of $m(\mathbf{x})$ around $\mathbf{x} = \mathbf{0}$ is, of course, closely related to the constraints previously mentioned. Finally, a technique to construct a mask of given order adding and multiplying simple band-limited masks is proposed in [19].

To conclude, Fig. 7 compares the behavior of three different radial transmissions in the case of a circular aperture. To facilitate this comparison the three masks have been normalized to the same $\mathcal{B}$ in order to guarantee that the Lyot stop will be the same: for these simulations the flux reduction equals $4/9$. The first transmission is a Bessel cardinal function and the second one the square of a Bessel cardinal function. These functions have been properly scaled and shifted in amplitude in order that $t(r)$ occupies all the interval $[0, 1]$. The properties detailed in the item (i) of the discussion above can also be expressed in term of concentration of the energy of $m(\mathbf{x})$ around the origin. In this case, the additional band-limited constraints leads naturally to a prolate function. Such a solution is illustrated in Fig. 7. In this case the coefficient c has been chosen empirically equal to 6.8 in order to achieve a good compromise between the ripples and the width of the main lobe. This result shows that the behavior of such a prolate mask and a squared Bessel cardinal mask are similar. Finally, it is important to emphasize that the masks illustrated in Fig. 7 correctly behave with respect to the requirement (i). On the contrary they do not fulfil (ii) (the order of the squared Bessel cardinal and the prolate is only 4) and will have a poor behavior in the case of a resolved star or a misalignment.

To conclude, it is important to mention in this section devoted to optimal coronagraphy with infinite size masks the 4QC. This coronagraph will be developed in a separate article by D. Rouan et al. [20] in this volume. It relies on the fact that for a circular aperture and a transmission $t(\mathbf{x}) = \operatorname{sign}(x)\operatorname{sign}(y)$, the complex amplitude in C is identically zero inside the pupil for a point source on the axis. This beautiful result relies on a nontrivial property of the Fourier

transform (see the paper of D. Rouan and included references for a proof). A coronagraphic technique based on a similar principle is developed in [21].

Acknowledgements

The authors would like to thank Peter Falloon for his Mathematica program.

Appendix A. Prolate spheroidals functions

This section presents some facts about prolate spheroidals functions. For a detailed presentation refer to the seminal papers [22] and references therein. Application of prolate spheroidals functions to optics can be found in [23].

In order to gain deeper insight in their principal properties using simple mathematics the Cartesian coordinates have been preferred. Derivations using specific coordinates can be found in the references. We consider a real valued square-integrable function f of d variables having a bounds support $\mathcal{P}$ (in our case $f(\mathbf{x}) = a(\mathbf{x})p(\mathbf{x})$ and $\mathcal{P}$ is the telescope pupil):

$$\hat{f}(\mathbf{x}) = \int_{\mathcal{P}} f(\mathbf{y}) \mathrm{e}^{-\imath 2\pi \mathbf{x}\mathbf{y}^t} \,\mathrm{d}\mathbf{y} \tag{A.1}$$

The ratio between the energy encircled in the bounded frequency domain $\mathcal{M}$ (here the Lyot mask) and the total energy of $\hat{f}$ is obtained integrating the squared modulus of (A.1) on $\mathcal{M}$:

$$\int_{\mathcal{P}} K_{\mathcal{M}}(\mathbf{x}-\mathbf{y}) f(\mathbf{y}) f(\mathbf{x}) \,\mathrm{d}\mathbf{y}\,\mathrm{d}\mathbf{x} \Big/ \int_{\mathcal{P}} f(\mathbf{y})^2 \,\mathrm{d}\mathbf{y} \tag{A.2}$$

where $K_{\mathcal{M}}(\mathbf{x})$ is the inverse Fourier transform of the indicator function of $\mathcal{M}$, $\mathbb{1}_{\mathcal{M}}(\mathbf{x})$. Standard results on functional analysis, see for example [24], prove that the maximum of this ratio is the largest eigenvalue Λ of the integral equation:

$$\int_{\mathcal{P}} K_{\mathcal{M}}(\mathbf{x}-\mathbf{y})\psi(\mathbf{y}) \,\mathrm{d}\mathbf{y} = \Lambda\psi(\mathbf{x}), \quad \mathbf{x} \in \mathcal{P} \tag{A.3}$$

or equivalently: $(\psi(\mathbf{x})\mathbb{1}_{\mathcal{P}}(\mathbf{x})) * K_{\mathcal{M}}(\mathbf{x}) = \Lambda\psi(\mathbf{x})$, $\mathbf{x} \in \mathcal{P}$. Note that, according to (A.3), in order to obtain $\psi(\mathbf{x})$ real, $K_{\mathcal{M}}(\mathbf{x})$ must be real and consequently $\mathcal{M}$ must have a central symmetry.

The prolate spheroidal functions are defined as the solution of this integral equation. The extrema of (A.2) is reached when $f(\mathbf{x})$ is the corresponding eigenfunction. The kernel being positive defined the eigenvalues of (A.3) are positive and from (A.2) upper bounded by 1. The eigenfunctions of (A.3) are orthogonal and complete on $L_2(\mathcal{P})$ and orthogonal on $\mathbb{R}^d$ when extended on outside $\mathcal{P}$ using (A.3).

Considerable simplifications occur when $\mathcal{P}$ and $\mathcal{M}$ are both scaled versions of a 'normalized' domain $\mathcal{U}$: $\exists \mathbf{R} \in \mathbb{R}^d$, $\exists \mathbf{r}_m \in \mathbb{R}^d$ such that $\mathbf{x} \in \mathcal{P} \Leftrightarrow \mathbf{x}./\mathbf{R} \in \mathcal{U}$, $\mathbf{x} \in \mathcal{M} \Leftrightarrow \mathbf{x}./\mathbf{r}_m \in \mathcal{U}$, where $\mathbf{a}./\mathbf{b}$ (resp. $\mathbf{a}.\mathbf{b}$) defines the vector having $a(k)/b(k)$ (resp. $a(k)b(k)$) as components. A simple change of variable in the definition of $K_{\mathcal{M}}(\mathbf{x})$ shows that $K_{\mathcal{M}}(\mathbf{x}) = (\prod_{k=1}^{d} r_m(k)) K_{\mathcal{U}}(r_m.\mathbf{x})$. Substitution of this result in (A.3) shows that $\psi(\mathbf{x}) = \phi(R.\mathbf{x})$ where $\phi(\mathbf{x})$ is the solution of the normalized problem:

$$\left(\prod_{k=1}^{d} r_m(k) R(k)\right) \int_{\mathcal{U}} K_{\mathcal{U}}\big(\mathbf{r}_m.\mathbf{R}.(\mathbf{x}-\mathbf{y})\big)\phi(\mathbf{y}) \,\mathrm{d}\mathbf{y} = \Lambda\phi(\mathbf{x}), \quad \mathbf{x} \in \mathcal{U} \tag{A.4}$$

- Consider for example the case where $d = 2$ and $\mathcal{U}$ is a disk. Consequently $\mathcal{P}$ and $\mathcal{M}$ are ellipses. Eq. (A.4) shows that the prolate functions associated to an ellipse are obtained by a scaling of the prolate function associate to a disk.
- If $\forall k$, $R(k) = R$ and $r_m(k) = r_m$, Eq. (A.4) shows that the prolate spheroidal functions for a given $\mathcal{U}$ form essentially a one-parameter family of functions depending of the parameter $c = r_m R$.

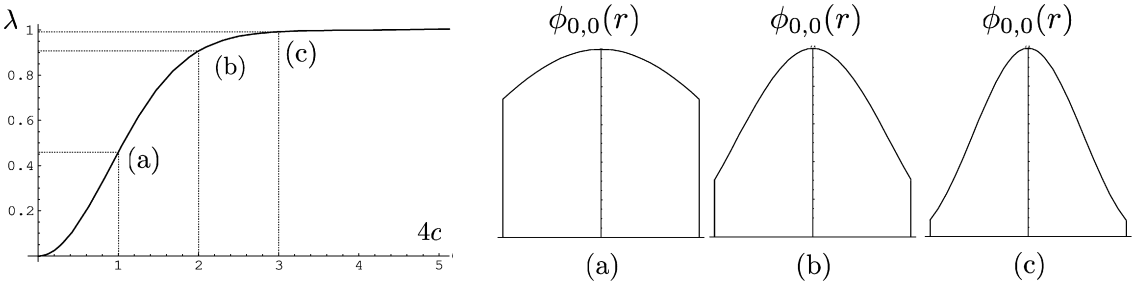

Fig. 8. Circular prolate functions. Left plot gives the largest eigenvalue Λ as a function of $4c$. The right plot gives the prolate functions for three values of c.

Additional simplifications in this last case are obtained when the region $\mathcal{U}$ is symmetric: $\mathbf{x} \in \mathcal{U} \Leftrightarrow -\mathbf{x} \in \mathcal{U}$. In this case the solutions of (A.4) are real, either even with a real eigenvalue or odd with a pure imaginary eigenvalue. Moreover, finding the solutions of (A.4) is equivalent to finding the solutions of:

$$\alpha\phi(\mathbf{x}) = \int_{\mathcal{U}} \mathrm{e}^{-\imath 2\pi c \mathbf{x}\mathbf{y}^t} \phi(\mathbf{y})\,\mathrm{d}\mathbf{y}, \quad \mathbf{x} \in \mathcal{U}, \text{ where } \Lambda = |\alpha|^2 c^d \tag{A.5}$$

This equation states that the Fourier transform of $\phi(\mathbf{x})\mathbb{1}_{\mathcal{U}}(\mathbf{x})$ is proportional to $\phi(\mathbf{x}/c)$. This result can be easily verified taking the complex conjugate of (A.5), substituting in the integral $\phi(\mathbf{y})$ by the right side of (A.5) and identifying with (A.4).

Circular prolate functions correspond to the case $d = 2$ where $\mathcal{U}$ is the unit disk. They decompose, up to a normalisation factor as $\phi_{N,n}(r,\theta) = R_{N,n}(r)[\cos, \sin](N\theta)$. Consequently (A.5) shows that they can be defined by their invariance to a finite Hankel transform of order N, e.g.:

$$2\pi \int_0^1 R_{0,0}(r) J_0(2\pi r\rho) r\,\mathrm{d}r = \frac{\sqrt{\Lambda}}{r_m} R_{0,0}\left(\frac{\rho}{r_m}\right) \tag{A.6}$$

Fig. 8 represents the radial circular prolate function $\phi_{0,0}(r)$ for three different values of c.

A crucial point is the numerical computation of prolate functions. For simple geometries, the functions can be computed using rapidly converging series as the one derived for a circular aperture in [22]. For general geometries, prolate spheroidal functions can be computed directly solving Eq. (A.3) using one of the numerous iterative algorithm that can be found in the literature for the computation of the eigenelements of linear integral operators, e.g. the algorithm used in [17,25].

References

[1] G. Chabrier, I. Baraffe, Theory of low-mass stars and substellar objects, Annual Review of Astronomy and Astrophysics 38 (2000) 337.
[2] I. Baraffe, G. Chabrier, T. Barman, F. Allard, P.H. Hauschildt, Evolutionary models for cool brown dwarfs and extrasolar giant planets. The case of hd 20945, Astronomy and Astrophysics 402 (2003) 701.
[3] A. Burrows, D. Sudarsky, I. Hubeny, Spectra and diagnostics for the direct detection of wide-separation extrasolar giant planets, The Astrophysical Journal 609 (2004) 407.
[4] B. Lyot, The study of the solar corona and prominences without eclipses (George Darwin lecture, 1939), Monthly Notices of the Royal Astronomical Society 99 (1939) 580.
[5] F. Roddier, C. Roddier, Stellar coronagraph with phase mask, Astronomical Society of the Pacific 109 (1997) 815–820.
[6] C. Aime, R. Soummer, A. Ferrari, Total coronagraphic extinction of rectangular apertures using linear prolate apodizations, Astronomy and Astrophysics 389 (2002) 334–344.
[7] R. Soummer, C. Aime, P. Falloon, Stellar coronagraphy with prolate apodized circular apertures, Astronomy and Astrophysics 397 (2003) 1161–1172.
[8] R. Soummer, Apodized pupil Lyot coronagraphs for arbitrary telescope apertures, The Astrophysical Journal 618 (2004) 161–164.
[9] D. Rouan, P. Riaud, A. Boccaletti, Y. Clénet, A. Labeyrie, The four-quadrant phase-mask coronagraph I. Principle, Publication of the Astronomical Society of the Pacific 112 (2000) 1479–1486.
[10] M. Kuchner, W. Traub, A coronagraph with a band-limited mask for finding terrestrial planets, The Astrophysical Journal 570 (2002) 900–908.
[11] N.J. Kasdin, R.J. Vanderbei, D.N. Spergel, M.G. Littman, Extrasolar planet finding via optimal apodized pupil and shaped-pupil coronagraphs, The Astrophysical Journal 582 (January 2003) 1147–1161.

[12] C. Aime, Radon approach to shaped and apodized apertures for imaging exoplanets, Astronomy and Astrophysics 434 (May 2005) 785–794.
[13] A. Oppenheim, R. Schafer, Discrete-Time Signal Processing, Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 1989.
[14] A. Ferrari, Analytical analysis of Lyot coronagraphs response, The Astrophysical Journal 657 (March 2007) 1201.
[15] F. Martinache, PIZZA: a phase-induced zonal Zernike apodization designed for stellar coronagraphy, Journal of Optics A: Pure and Applied Optics 6 (August 2004) 809–814.
[16] P. Baudoz, 1999, Imagerie interférométrique grand champ et applications, Thèse de Doctorat, Université Nice Sophia Antipolis.
[17] O. Guyon, F. Roddier, Direct exoplanet imaging possibilities of the nulling stellar coronagraph, in: Interferometry in Optical Astronomy, vol. 4006, SPIE, Bellingham, WA, 2000, pp. 377–387.
[18] C. Aime, R. Soummer, Multiple-stage apodized pupil Lyot coronagraph for high contrast imaging, in: Astronomical Telescopes and Instrumentation 2004, SPIE, Bellingham, WA, October 2004, pp. 456–461.
[19] M. Kuchner, J. Crepp, J. Ge, Eighth-order image masks for terrestrial planet finding, The Astrophysical Journal 628 (2005) 466–473.
[20] D. Rouan, J. Baudrand, A. Boccaletti, P. Baudoz, D. Mawet, P. Riaud, The Four Quadrant Phase Mask coronagraph and its atavars, C. R. Physique 8 (2007), 10.1016/j.crhy.2007.05.004 (this issue).
[21] G. Foo, D.M. Palacios, G.A. Swartzlander, Optical vortex coronagraph, Optics Letters 30 (December 2005) 3308–3310.
[22] D. Slepian, Prolate spheroidal wave functions, Fourier analysis and uncertainty, IV: Extensions to many dimensions; generalized prolate spheroidal functions, Bell System Technical Journal 43 (1964) 3009–3058.
[23] B. Frieden, Progress in Optics, North-Holland, Amsterdam, 1971 (Ch. VIII, pp. 311–407).
[24] F. Riesz, B. Sz.-Nagy, Functional Analysis, Dover Publications, New York, 1990.
[25] T.J. Schulz, Iterative transform algorithm for the computation of optimal beams, J. Opt. Soc. Am. 21 (2004) 1970–1974.