The achromatic chessboard, a new concept of a phase shifter for nulling interferometry

I. Theory

D. Rouan and D. Pelat

1 LESIA, Observatoire de Paris, CNRS, UPMC, Université Paris Diderot, 5 place Jules Janssen, 92190 Meudon, France
e-mail: daniel.rouan@obspm.fr
2 LUTH, Observatoire de Paris, CNRS, Université Paris Diderot, 92190 Meudon, France
e-mail: didier.pelat@obspm.fr

Received 20 September 2007 / Accepted 21 February 2008

ABSTRACT

Context. Direct detection of a planet around a star and its characterisation for identification of bio-tracers in the mid-IR requires a nulling interferometer. Such an instrument must be efficient in a large wavelength domain in order to have the capability of simultaneously detecting the infrared spectral features of several bio-tracers: CO$_2$, O$_3$, and H$_2$O.

Aims. A broad wavelength range can be effective provided that an achromatic phase shift of $\pi$ can be implemented, with good enough accuracy to achieve a deep nulling at all considered wavelengths. A new design concept for such an achromatic phase shifter is presented here. The major interest of this solution is that it allows a simple design with only one device per beam.

Methods. The heart of the system consists in two cellular mirrors where each cell has a thickness that introduces, for a given central wavelength, a phase shift of $(2k + 1)\pi$ or of $2k\pi$ on the fraction of the wave it reflects. Each mirror is put in one of the collimated beams of the interferometer. Because of the odd/even distribution, a destructive interference is produced on axis for the central wavelength when recombining the two beams. If the number of cells of a given thickness follows a rather simple law based on the Pascal’s triangle, we then show that the nulling is also efficient for a wavelength that is not too far from the central wavelength.

Results. The effect of achromatization is more efficient the more cells there are. For instance, with two mirrors of $64 \times 64$ cells, where the cells’ phase shift ranges between $-6\pi$ and $+6\pi$, one reaches a nulling of $10^{-6}$ on a wavelength range [0.6, 1.25, 16], i.e. on more than one complete octave. This is why we claim that this device produces a quasi-achromatic phase shift; especially, it could satisfy the specifications of a space mission as DARWIN. In a second step, we study the optimum way to distribute the cells in the plane of the pupil. The most important criterion is the isolation of the planet image from the residual image of the star. Several algorithms and their nulling performances are presented.

Key words. instrumentation: interferometers – techniques: high angular resolution – techniques: interferometric – space vehicles: instruments – stars: planetary systems

1. Introduction

There is a huge brightness contrast ($10^{6}$–$10^{7}$) between a star and a planet that orbits it, so that directly detecting the planet requires cancelling as much of the stellar light as possible. Today, it is generally admitted that there are two main instrumental paths to this goal: one is coronagraphy on a single telescope (e.g. Guyon 2007) generally in the visible to near-IR domain, the other nulling interferometry – proposed essentially in the thermal infrared – using at least two telescopes recombined in a clever way (e.g. Léger et al. 1996; Woolf & Angel 1997). This paper focuses on the second category. Bracewell (1978) was the first to propose the concept of a nulling interferometer, where the light collected by two telescopes is coherently recombined, after applying a $\pi$ phase shift to one of the two arms of the interferometer. The system of fringes projected onto the sky then shows a central dark fringe. If the star image is put on this central dark fringe, it is strongly attenuated (actually, the stellar photons are all sent to the second output of the interferometer). If, at the same time, the planet is on a bright fringe, i.e. a region of transmission unity, then it can in principle be detected in a much more efficient way, especially because of the reduced photon noise.

Obtaining an achromatic $\pi$ phase shift is one of the main keys to success because a) the wavelength domain is broad where the set of reputedly unambiguous spectroscopic signatures of life, CO$_2$ (15 $\mu$m), O$_3$ (9.5 $\mu$m) and H$_2$O (5–8 $\mu$m), is to be found (Léger et al. 1996; Ollivier 2007), b) in photon-starving experiments such as those considered here. Any design based on a monochromatic detection is out of question.

Several solutions have been presented as approaches to an achromatic $\pi$ phase shift in a large range of wavelengths; for a review, see Rabbia (2004). The three main solutions are stacking dielectric plates of different indexes and dispersions (Mieremet et al. 2000), crossing a focus (Rabbia et al. 2005), and using a periscope optical train to reverse the electric field (Serabyn & Colavita 2001). The first solution only provides a quasi-achromatization since it relies on an approximate linearization of the optical path vs. wavelength. The two last methods should provide an intrinsic achromatic phase shift, but one that is strictly limited to $\pi$. In addition, instrumental effects (differential polarization for instance) may also affect the resulting null. The two
first methods introduce a specific subsystem in one arm only of the interferometer, resulting in an asymmetric design with the drawback of a difficult balancing of the two arms. All those solutions require the use of several optical components, generally requiring that the beams be folded several times. Is there a way of reaching a \( \pi \) phase shift with a simple optical component, while keeping the interferometer design symmetric? We present a new concept that has this in mind.

Rouan (2003, 2004, 2006, 2007) presented a method of defining interferometric configurations of telescopes that provide a \( \theta^\pi \) nulling function for any even value of \( n \). Achieving a high value of \( n \) is important in order to produce the trough shape of the nulling function and thus a good cancellation of the leaks due to the finite angular size of the star. The principle is based on a remarkable property discovered by Prouhet (1851) about a peculiar partition in two sets of the \( N = 2^L \) first integers, according to the Prouhet-Thué-Morse sequence. This property is such that, by selecting half of the \( N \) telescopes according to this sequence and applying a \( \pi \) phase shift to the output of each one, facilitates cancellation of the \( (L-1) \) first terms of the development in \( \theta \) of the recombined amplitude, thus leading to the desired \( \theta^{2L} \) behaviour of the intensity.

We continue in the same spirit while tackling the achromatisation of the \( \pi \) phase shift. The idea is to apply a similar principle of cancelling the first terms of Taylor’s development of the amplitude, but now with respect to the relative wavelength difference \( \Delta \lambda / \lambda_0 \). The central component of the new concept is a chessboard mirror, i.e. a mirror made with a pattern of cells of different thicknesses.

In this first paper, we only consider the case of an interferometer with two square pupil telescopes and mainly establish the theoretical grounds on which this new concept lies. We show that there is indeed a peculiar distribution of the cells’ thicknesses on the chessboard mirror that fulfils our goals of a quasi-achromatization of the \( \pi \) phase shift. We also show that an adequate distribution of the cells on the mirror’s surface can add more nulling power. We depict several configurations that have good properties for detecting a planet, especially one. We then study the performances that a perfect device would require and show that it would satisfy the specification of the currently proposed missions, such as Darwin (Leger & Herbst 2007). Several appendices explain the formalism and the mathematical background for the readers who will be seeking a deeper investigation. A second paper will address the questions of the performances in a real world where a device cannot be perfect, and will present the first results of an experimental demonstrator.

2. The concept of the chessboard phase shifter

The cartoon of Fig. 1 depicts an interferometer made with two identical afocal telescopes whose output beams are recombined in the so-called co-axial mode thanks to a beam-splitter. To simplify the design and forthcoming computations, we assume that the shape of each telescope pupil is limited by a square. In a collimated beam issued from the first telescope, we introduce a square cellular mirror. Alternatively, transparent plate could be used. This cellular mirror has a chessboard design because it features \( n \times n \) square cells arranged in columns and rows as on a chessboard. The cells have different thicknesses, and all are chosen to produce an optical path difference (OPD hereafter) that is an odd multiple of \( \lambda_0 / 2 \), where \( \lambda_0 \) is the central wavelength. Figure 2 gives an example of what a \( 4 \times 4 \) cellular mirror could look like. Symmetrically, a cellular mirror, where cells produce OPD that are even multiples of \( \lambda_0 / 2 \), is put into the output beam of the second telescope. By design, each cell then introduces a phase shift of \( (2k + 1) \pi \) or of \( 2k \pi \) on the fraction of the wave at wavelength \( \lambda_0 \) it reflects. Hereafter, we call odd and even the two cellular mirrors. Of course, the number of cells is the same on the two mirrors and their geometry is unique (square).

Note that a Fizeau (or multi-axial) recombination, i.e. done at the image plane level and not at the pupil level, could also be considered, with the advantage of keeping the design fully symmetric. It has been shown (e.g., Bausset et al. 2006; Wallner et al. 2004) that nulling is as efficient with this design as in a more classical co-axial recombination (Michelson type), provided that the star image (i.e. the PSF) is formed on single-mode fiber optics that will make the overlap integral of the product between the incident electric fields and the fundamental mode distribution of the fiber. In that case, the nulling comes from the amplitude being anti-symmetric with respect to the \( y \) axis. However, in the following, we stick to the co-axial configuration and especially will show images of the resulting PSF for that type of recombination. The effect of nulling on the putative planet detection is seen with a much better contrast since most of the stellar light intensity has been cancelled.

The amplitude integrated over the whole pupil area, at the output of one chessboard mirror is

\[
a = \sum_{k=1}^{N} \exp(i\phi_k)
\]

where \( \phi_k \) is the phase difference between one cell and a reference cell, \( \phi_k = 2\pi d_k / \lambda \), and \( d_k \) is the OPD produced by cell \( k \). Assuming that each OPD is a multiple of \( \lambda_0 / 2 \), where \( \lambda_0 \) is the central wavelength, the phase then reads \( \phi_k = n_k \pi \lambda_0 / \lambda \). If \( \lambda \) is within a rather narrow range with respect to \( \lambda_0 \), then one can write \( \phi_k = n_k(1 - \Delta \lambda / \lambda_0) \).
If on one cellular mirror all the OPD are even multiples of \(\lambda_0/2\) and, on the other mirror, they are odd multiples of \(\lambda_0/2\), then the complex amplitudes take the form

\[
a_+ = \sum_{k=1}^{N} \exp(j\pi m_k \Delta \lambda/\lambda_0),
\]

where \(m_k\) are even integers, and

\[
a_- = \sum_{k=1}^{N} - \exp(j\pi n_k \Delta \lambda/\lambda_0),
\]

where \(n_k\) are odd integers.

When, after reflection on the cellular mirrors, the two beams are recombined, then the integrated amplitude \(a\) of the wave at the output of the recombiner (or, equivalently, at the focus of the final lens) is

\[
a = a_+ + a_- = \sum_{k=0}^{N} \exp(j\pi m_k \Delta \lambda/\lambda_0) - \sum_{k=0}^{N} \exp(j\pi n_k \Delta \lambda/\lambda_0),
\]

\[\phi = \pi \Delta \lambda/\lambda_0\) and developing each exponential in series of \((n_k \phi)^o\), so that the two terms of the above expression become

\[
a_+ = \sum_k 1 + j\phi \times \sum_k m_k - \phi^2 \times \sum_k m_k^2 - j\phi^3 \times \sum_k m_k^3 + ...
\]

and

\[
a_- = - \sum_k 1 - j\phi \times \sum_k n_k + \phi^2 \times \sum_k n_k^2 + j\phi^3 \times \sum_k n_k^3 + ...
\]

Clearly, since the number \(N\) of cells is the same for the two mirrors, the first term of each expression mutually cancels out. In order to cancel out the other terms, at least up to a given order \(O\), we need to find one set \([m_k]\) of even integers and one set \([n_k]\) of odd integers, such that \(\sum_k m_k = \sum_k n_k\) for all \(o \leq O\). Note that several of the \(m_k\) (respectively of the \(n_k\)) can be identical. This problem is known as the Prouhet-Tarry-Escott problem (Borwein & Ingalls 1994), with the additional peculiar condition on the odd/even reparation of the integers.

The good news is that solutions do exist for any value of \(O\)! Even more, there are several solutions for a given value of \(O\) and the number of solutions increases very rapidly with \(O\). For instance, let us consider a nulling up to \(O = 2\). There, we are looking for at least one set of even integers \(m_k\) and one set of odd integers \(n_k\), such that

\[
\sum_k m_k = \sum_k n_k \quad \text{and} \quad \sum_k m_k^2 = \sum_k n_k^2.
\]

One solution (that can be found for instance by trying all possible combinations of first odd and even integers) is

\[
[m_k] = [2, 2, 4], \quad [n_k] = [1, 3, 3, 3].
\]

One checks that \(2 + 2 + 2 + 4 = 1 + 3 + 3 + 3 = 10\) and that \(2^2 + 2^2 + 2^2 + 4^2 = 1^2 + 3^2 + 3^2 + 3^2 = 28\).

At this point, an infinity of solutions can be built from a given one. It suffices to add any integer to all elements of the two sets. For instance the sets \([m_k] = [4, 4, 4, 6]\), \([n_k] = [3, 5, 5, 5]\), obtained by adding 2 to each element of the solution given as first example, constitute a valid solution. This is so, because \(\sum (m_k + L)^2 = \sum m_k^2 + oL \lambda_0^{(o-1)} + \sum (L^2 m_k^{o-2} + \ldots + L^o)\), where any elementary even sum \(\sum_k (L m_k^{o-1})\) is equal to the equivalent odd sum \(\sum_k (L n_k^{o-1})\).

From the optical point of view, it is clear that all those solutions are totally equivalent, which is simply adding the same phase shift (or piston) everywhere. In the following, we use this property of invariance by translation. Hereafter we consider only solutions where at least one cell has a null thickness (0), simply meaning that it corresponds to the reference level of the mirror surface.

In the following section, we examine how proper solutions can be built in a coherent and systematic way for any value of \(O\).

### 2.1. The various solutions

As said, there is a rapidly increasing number of solutions when \(O\) increases. For instance, for \(O = 3\), a systematic search gives 16 solutions when the number of cells per mirror is set to 32. However, all those solutions are not equivalent from the instrumental point of view. We are looking for the solutions best adapted to our optical problem. One natural condition is that the dispersion of the thickness on the surface of the mirror is not too large. Obviously if the cellular mirror comprises a few cells with thickness much greater than the other, then there will be both a problem of manufacturing of the mirror (steep edges) and of mutual shadowing by the cells. For instance, the solution \([m_k] = [4, 6, 6, 0, 0, 0, 0], [n_k] = [1, 1, 1, 1, 1, 1, 9]\) is valid for \(O = 2\), but the odd mirror exhibits a tower of high thickness emerging from a flat pond, and this is clearly not desirable.

To avoid a large dispersion in thickness, one can put as a condition that the ordered set of thickness, increases by one unit from one cell to the next. In other terms, we want to use all thicknesses in a set of incrementing integers. This is a guarantee of a rather low variance of the thickness on the surface of the mirror.

### 2.2. The optimized solution

We have found an elegant and singular solution that satisfies this maximum flatness condition and is based on Pascal’s triangle (the rule to derive the binomial coefficients recursively). More precisely, for a nulling up to \(O\), we consider mirrors with \(2^O\) cells, the number of cells of thickness \(k\) being the binomial coefficient \(\binom{O}{k}\). A clear way to present the solution is simply to write the first integers from 0 to \(O\) on one line and the binomial coefficients of order \(O\) on the next line:

| OPD introduced by cell | Number of cells |
|-----------------------|-----------------|
| 0                     | 1               |
| 1                     | \(\binom{O}{1}\) |
| 2                     | \(\binom{O}{2}\) |
| ...                  | ...            |
| \(O-1\)               | \(O-1\)        |
| \(O\)                 | 1               |

The first line gives the OPD introduced by the cell (in units of the elementary step \(\lambda_0/2\)), and the second line gives the number of cells of this kind. For instance, if \(O = 6\), then

| Thickness of cell | Number of cells |
|-------------------|-----------------|
| 0                 | 1               |
| 1                 | \(6\)           |
| 2                 | \(15\)          |
| ...               | ...            |
| 4                 | \(4\)           |
| 5                 | \(5\)           |
| 6                 | \(1\)           |

This means that the even mirror is built as follows: a) it is made of \(O/2 + 1\) (if \(O\) is even) or \((O + 1)/2\) (if \(O\) is odd) subsets of cells; b) all cells in the subset \(k\) have an OPD equal to \(k\); c) the number of cells in the subset \(k\) is given by the coefficient of the binomial, \(\binom{O}{k}\). All even values of \(k\), such that \(0 \leq k \leq O\), are thus found at the surface of the even mirror. For instance, 1 cell of null OPD, \(\binom{O}{0} = O(0-1)/2\) cells of OPD 2, \(\binom{O}{2}\) cells of OPD 4, ..., 1 cell of OPD \(O\) if \(O\) is even, or \(O\) cells of OPD \(O - 1\) if \(O\) is
odd. Symmetrically, the odd mirror is built with the same rule; but with \( k \) being odd, it thus exhibits \( O \) cell of OPD 1, \((\frac{1}{2})\) cells of OPD 3, \((\frac{1}{2})\) cells of OPD 5, ..., 1 cell of OPD \( O \) if \( O \) is odd or \( O \) cells of OPD \( O - 1 \) if \( O \) is even. Note that there are \( O/2 \) (if \( O \) is even) or \((O + 1)/2\) (if \( O \) is odd) subsets of odd cells.

One can demonstrate that this solution does satisfy the condition of nulling: \( \sum_{i=1}^{O} m_{o}^{i} = \sum_{i=1}^{O} n_{o}^{i} \) for all \( o \leq O \) (see Appendix).

3. \( X-Y \) distribution of cells in the pupil: which criterion?

If the optimum solution provides the distribution in \( z \) of the cells, it does not say anything about the distribution in \( x \) and \( y \), i.e. on the surface of the mirrors. Once again, some common sense can be exercised to select one over the other. For instance, if all the large thicknesses are on one side of a mirror, or if there is a continuous gradient of thickness across the surface, then this is equivalent to introducing some tilt on the wavefront, a bias to be avoided; see Fig. 3-right for an example of configurations with the same \( z \)-distribution but different \( x-y \) distributions, one being better than the other. A first condition that can be stated is that the global tilt be zero. This can be expressed by requiring that the first moments of the phase of the reflected wavefront be zero:

\[
\sum_{i,j} i \times m_{ij} = 0 \quad \text{and} \quad \sum_{i,j} j \times m_{ij} = 0,
\]

where \( i, j \) are the coordinates of a cell in a chessboard pattern, the origin being at the centre of the mirror. This does not seem to be a hard constraint to fulfill. To go further, one can ask that higher moments of the phase be equal to zero, so that higher order aberrations are cancelled:

\[
\sum_{i,j} i^{p} \times m_{ij} = 0 \quad \text{and} \quad \sum_{i,j} j^{p} \times m_{ij} = 0.
\]

In a final step, all moments, including the cross terms could also be considered:

\[
\sum_{i,j} i^{p} \times j^{q} \times m_{ij} = 0.
\]

Another criterion is to consider the image plane, i.e. the plane conjugated to the pupil plane through the Fourier transform (in an on-axis recombination as mentioned above). We can look for patterns of cells, such that the nulling for the starlight extends as far as possible from the centre, for instance along the two main axes, with of course less of attenuation for the planet. Assuming that the cells on the even and odd mirrors are distributed respectively according to a matrix \( P \) and a matrix \( Q \), where each element \( e_{i,j} \) is the integer determining the OPD of the cell \((i, j)\).

At a wavelength \( \lambda \neq \lambda_{o} \), the Fourier transform (hereafter noted \( \hat{1} \)) of the combined pupil reads

\[
1[a_{+} + a_{-}] = \hat{1}(\exp jP\phi \times \exp jQ\phi) \times \text{cell},
\]

where \( \text{cell} \) means a function whose value is 1 inside a square of the size of one cell and 0 outside. Note that, with this condensed notation, \( \exp jP\phi \) means in reality the 2D Dirac’s comb with a period of one cell width and an amplitude \( \exp jP_{k}\phi \). Using the properties of the convolution \( \ast \), this expression becomes

\[
1[a_{+} + a_{-}] = \hat{1}(\exp jP\phi \times \exp jQ\phi) \times \text{PSFC},
\]

where \( \text{PSFC} \) is the point spread function given by a unique cell. Since the term \( \text{PSFC} \) only corresponds to the envelope which modulates the resulting PSF, we can restrain the study to \(1[\exp jP\phi - \exp jQ\phi] \), the term that carries the information on the distribution of the cells. Because we are dealing with Dirac’s combs, the Fourier transforms reduces to a discrete sum

\[
\hat{1}[\exp jP\phi - \exp jQ\phi] = \sum_{k,l} (\exp jP_{k,l} - \exp jQ_{k,l}) \exp j\alpha(kx + ly),
\]

where \( P_{k,l} \) and \( Q_{k,l} \) are the elements of matrix \( P \) and \( Q \), respectively, and \( \alpha \) is a coefficient resulting from the Fourier transform. After a Taylor development, this expression can be written

\[
\sum_{k,l} \{ j\phi(p_{k,l} - q_{k,l}) - \phi^{2}(p_{k,l}^{2} - q_{k,l}^{2}) \times j\phi^{3}(p_{k,l}^{3} - q_{k,l}^{3}) \times j\phi^{3}(p_{k,l}^{3} - q_{k,l}^{3}) + \ldots \exp j\alpha(kx + ly).
\]

Considering this expression, we see that a sufficient condition to get a null contribution at any point of the \( y \) axis \((x = 0)\), is to have simultaneously:

\[
\sum_{k,l} (p_{k,l} - q_{k,l}) = 0, \quad \sum_{k,l} (p_{k,l}^{2} - q_{k,l}^{2}) = 0, \quad \sum_{k,l} (p_{k,l}^{3} - q_{k,l}^{3}) = 0,
\]

etc. for any value of \( l \).

We note that this is the same condition, applied here to each column, that we found for the whole pupil, on the power of sets of odd and even numbers, i.e. \( \sum_{o=1}^{O} m_{o}^{i} = \sum_{o=1}^{O} n_{o}^{i} \) for all \( o \leq O \).

This means that if we can distribute the cells along the columns of the odd and even mirror, so that for each column the condition \( \sum_{o=1}^{O} m_{o}^{i} = \sum_{o=1}^{O} n_{o}^{i} \) is satisfied up to an index \( O \), then the amplitude is not only very low at the centre, but also along the \( y \) axis. Of course, the same result applies along the \( x \) axis if the condition can be satisfied for all lines. This does not guarantee that the nulling will be as deep outside those two main axes, but by continuity of the PSF, there is a good chance that it remains fairly good. We see below that such distributions of cells indeed satisfy this condition, along both the lines and the columns (see also Appendix D for a different but complementary approach).

3.1. The optimized solutions

We found several ways to geometrically distribute the cells on the mirror’s surface, following one or the other of the prescriptions established in the previous section. We describe hereafter three of them, each with its advantages and limits. The first one corresponds to the first criterion specifying that the first moments in the pupil must be zero, while the two other patterns satisfy the second criterion on the nulling along the \( x \) and \( y \) axis in the image plane. Other solutions, which are essentially variants of those three, have been also explored and the section on performances includes them.
3.1.1. First method

This first method tries to fulfill the condition on minimizing the moments on the wavefront after the reflection on the chessboard mirror, and it is based on a recurrence scheme to build the pair of mirrors of order $O$ from the pair of order $O-1$. Let us assume that three square matrix $B$, $P$, $Q$ are defined with dimension $2^{O-1} \times 2^{O-1}$. The even matrix is $P$ and the odd one is $Q$, while $B$ is an auxiliary matrix. We then build the matrix of dimension $2^O \times 2^O$ as follows

\[
\begin{align*}
P_0 &= P_{0-1} - Q_{0-1} - B_{0-1} \\
Q_0 &= Q_{0-1} + P_{0-1} - B_{0-1} \\
B_0 &= B_{0-1} - Q_{0-1} \\
\end{align*}
\]

As a starting point we choose $P = 0$, $Q = 1$, $B = 1$, such that at a first step

\[
\begin{align*}
P_1 &= 0 \quad 0 \\
Q_1 &= 1 \quad -1 \\
B_1 &= 1 \quad -1 \\
\end{align*}
\]

Fig. 4 shows an example of the pattern of cells obtained with that method, and Fig. 7a displays the resulting PSF with a planet $10^{-6}$ fainter than the star. One can verify that the first moments of the wavefront, up to the order $O$, do satisfy the condition of being null.

3.1.2. Second method

The aim of this distribution is to produce an efficient nulling along both the $x$- and $y$-axes in the image plane. We start from the remark that, if one writes the first $2^O$ integers in their binary form and associate the number of bits set to 1 to each integer, the frequency of those numbers are distributed according to the binomial coefficients. For instance if $O = 3$, then

\[
\begin{array}{cccccccc}
\text{Integer} & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 \\
\text{Binary} & 000 & 001 & 010 & 011 & 100 & 101 & 110 & 111 \\
\# of 1 & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 \\
\end{array}
\]

The occurrence of bits set to 1 is indeed \( \binom{3}{1} = \binom{3}{2} = 3 \) for the one or two bit set, and \( \binom{3}{0} = \binom{3}{3} = 1 \) for the zero or three bit set.

If we reproduce the elements of this line through all permutations, we will keep the property of Pascal’s triangle. We can then easily build a square matrix with $2^O$ lines of this kind. However, we wish to do it so that we also produce columns that share the same property. One practical way to obtain this result is to write $i \oplus j$ (exclusive OR) at the intersection of line $i$ and column $j$. The number of bits set to 1 on the resulting number then gives the OPD of the cell in steps of $\lambda_0/4$. An example is given below:

\[
\begin{array}{cccccccc}
000 & 0 & 1 & 0 & 1 & 1 & 0 & 1 \\
001 & 1 & 0 & 2 & 1 & 2 & 1 & 3 \\
010 & 1 & 2 & 0 & 1 & 2 & 3 & 1 \\
011 & 2 & 1 & 1 & 0 & 3 & 2 & 2 \\
100 & 1 & 2 & 2 & 3 & 0 & 1 & 2 \\
101 & 2 & 1 & 1 & 3 & 1 & 2 & 1 \\
110 & 2 & 2 & 4 & 0 & 2 & 2 & 3 \\
111 & 2 & 2 & 4 & 0 & 2 & 2 & 3 \\
\end{array}
\]

One can check that: a) each line or each column contains the right number of cells following the Pascal’s triangle rule; b) the lines and columns are distributed exactly the same way (symmetric matrix with respect to the second diagonal). After this operation, the odd and even numbers are interleaved. The odd and even mirrors are then built by simply adding 1 to either the even cells or the odd cells, as illustrated in the following table:

\[
\begin{array}{cccccccc}
0 & 2 & 2 & 2 & 2 & 2 & 4 & 2 \\
0 & 2 & 2 & 2 & 2 & 4 & 2 & 2 \\
0 & 2 & 2 & 2 & 4 & 2 & 2 & 2 \\
0 & 2 & 2 & 4 & 0 & 2 & 2 & 2 \\
\end{array}
\]

Fig. 5 shows an example of the pattern of cells obtained with that method and Fig. 7b displays the resulting PSF with a planet $10^{-6}$ fainter than the star.

3.1.3. Third method

The third method also aims at producing an efficient nulling along both the $x$- and $y$-axes, but is now based on a recurrence scheme that is applied to a single matrix $R_O$ (with $R_0 = 0$):

\[
R_O = R_{O-1} - (-1)^O \quad R_{O-1} = R_{O-1} + (-1)^O
\]

For instance:

\[
\begin{array}{cccccccc}
R_1 & 0 & -1 & 1 & 0 \\
R_2 & 0 & -1 & 1 & 0 & 1 & 0 & 2 & 0 \\
\end{array}
\]
To extract the odd and even matrices, one simply adds 1 to the even elements or to the odd elements in matrix $R$, as in method 2. One can check that, again, the desired condition on the sum of powers along each column or each line is satisfied. Figure 6 shows an example of the pattern of cells obtained with that method, and Fig. 7c displays the resulting PSF, assuming a planet $10^{-6}$ fainter than the star.

In fact, two versions of this pattern exist, one that is antisymmetric and produces an asymmetric PSF and one that is symmetric and produces symmetric PSF. It seems for practical reasons (e.g. alignment of the optical setup) that the second pattern would be preferable.

4. Simulation and predicted performances

The most important question is now to evaluate whether the achromatic phase shifter we propose is indeed efficient, both in terms of nulling and of wavelength range. For that purpose, we have developed a numerical simulation to assess the respective performances of the various configurations we found. Essentially, it consists in a) building a pair of phase masks for a given type and a choice of dimensions; b) assuming a source exactly on axis, to perform a FFT on the sum of the complex amplitudes on the two masks, for a given shift of wavelength with respect to the central one (this produces the resulting nullled image of the star, of course after taking the square of the modulus); c) to add a phase shift, generally $\pi$, on one mask, in order to mimic the presence of a planet that would be on a constructive interference fringe, and again to perform the FFT and take the square of the modulus. In Fig. 7 we compare the different images of the planet – assumed to be at a level of $10^{-6}$ of the stellar flux – and residuals of the star, obtained for the typical case $(\lambda - \lambda_o)/\lambda_o = 0.2$.

The assessment is then done: a) on the quality of the nulling versus the wavelength, for instance in computing the residual flux in a circular diaphragm centred on the star position; b) on the magnitude of the unavoidable partial extinction of the planet; c) on the wavelength range where a good nulling is obtained.

The simple metric we used to compare the different configurations is to compute the ratio of the residuals of the planet flux to the star flux, each measured within one resolution element of a single telescope, i.e. an area of $\lambda/D$ around the main axis in the image plane. Another possible metric would be to estimate a signal-to-noise ratio for the planet detection, assuming either photon noise (planet + stellar residuals) or a noise proportional to the stellar residuals since it now seems clear that a lack of stability in the nulling will probably be the main source of noise in nulling interferometers (Lay 2004).

We present in Fig. 8 the results of the model in terms of nulling efficiency vs. $(\lambda - \lambda_o)/\lambda_o$ for the different configurations we have tested and for two numbers of cells in the phase mask.

The following remarks can be made. First, it is especially rewarding to see that the specifications on nulling efficiency that a mission like DARWIN would require (a contrast of $10^{-6}$ and in fact a specification of $10^{-5}$ on the nulling efficiency) could easily be met with the solution we propose; e.g., theoretical contrasts as high as $10^{15}$ are predicted on a significant wavelength domain.

Second, the performance in terms of achromaticity is fair, since the level of $10^{-6}$ is reached on the domain $[0.6\lambda_o, 1.25\lambda_o]$ with the best configuration. This corresponds to more than one complete octave in wavelength. It means that it is possible to cover the whole wavelength range of DARWIN with only two devices. We also note that there is an asymmetry in the useful wavelength domain, because the performances are within the specifications in a broader range on the short wavelength side.

Third, the number of cells is important, as shown by comparing curves at the top ($8 \times 8$) and the bottom ($64 \times 64$) in Fig. 8. However, increasing this number beyond $64 \times 64$ does not bring a very significant gain.

Fourth, among the different chessboard patterns, clearly the third method presents much better results – at least when we consider the raw numbers – since, for instance, it gives a $10^6$ times higher contrast than the next best configuration at $(\lambda - \lambda_o)/\lambda_o = -0.2$. This does not mean that it would always be the best, a more thorough analysis is needed, taking deviation from a perfect system into account once the device is manufactured.

5. Conclusion

We have presented a new concept of a quasi-achromatic phase shifter, a function that is at the heart of any nulling interferometer. We think that it can simplify the design and the setup of the interferometer since its relies on a simple component of bulk optics symmetrically introduced in the two arms of the interferometer. In that sense, it could overcome certain limitations in other solutions that use multicomponent systems and/or a complex asymmetric design. The heart of the concept is a chessboard...
Diophantine patterns of cells, the numbers or letter referring to the various distributions that satisfy this condition, a very unique development in nulling interferometers. It is the peculiar distribution of the dihedral mirror with a large number of cells (typically 64 × 64) put in the pupil plane. It is the peculiar distribution of the different cell thicknesses that is at the origin of the quasi-achromatic capability. It allows cancelling out the first terms of the Taylor’s development in Δl/λ of the amplitude, up to any given order. Among the various distributions that satisfy this condition, a very unique one is especially optimum. We also show that there is a peculiar x − y distribution of the cells that gives very good performances in terms of planet-to-star contrast. Reaching a very deep nulling over a broad wavelength range is a mandatory condition for detecting and ultimately characterising terrestrial planets with nulling interferometers. This solution can be of interest only if it reaches the proper performance of nulling over a broad enough range. Indeed, we show that the specification of 10−6 in contrast that a space mission such as DARWIN is aiming at can be reached on a complete octave in wavelength in the best case of a 64 × 64 chessboard mirror.

Finally, we want to introduce a new terminology. The set of equations on which this new concept is based is diophantine (i.e. polynomial equations where solutions are integers), as was a similar set of equations used for very deep nulling interferometers that one of us has been exploring (Rouan 2007) for few years; as a consequence, we propose to denote diophantine optics as this new type of optics where peculiar relations between powers of integers are used to obtain very specific effects.

Acknowledgements. The authors thanks Marie Ygouf for her help. This work also received the support of PHASE, the high angular resolution partnership between ONERA, Observatoire de Paris, CNRS and University Denis Diderot Paris 7.

Appendix A: The z-distribution based on Pascal’s triangle

We want to demonstrate that the condition \( \sum_k m^o_k = \sum_k m^p_k \) for all \( o \leq O \) is fulfilled with the solution given by Pascal’s triangle. We use recurrence reasoning. The nulling relation we want to demonstrate read

\[
\sum_{k=0}^{O-1} (-1)^k k^O \binom{O}{k} = 0
\]

for all \( o \leq O \).

The expression on the left can be rewritten

\[
\sum_{k=0}^{O-1} (-1)^k k^{o-1} \binom{O}{k}
\]

We now use the relationship \( k\binom{O}{k} = O\binom{O-1}{k-1} \) to transform this expression into

\[
\sum_{k=0}^{O-1} (-1)^k k^{o-1} O\binom{O-1}{k-1}
\]

We note that when \( o \neq 0 \) the summation can start at \( k = 1 \), so that the equation to verify is

\[
\sum_{k=1}^{O-1} (-1)^k k^{o-1} O\binom{O-1}{k-1} = 0.
\] (A.1)

Now let us assume that the nulling relation holds for \( p-1 \) and \( O-1 \)

\[
\sum_{k=0}^{O-1} (-1)^k (k+1)^{o-1} \binom{O-1}{k} = 0.
\]

This can be rewritten as

\[
\sum_{k=1}^{O-1} (-1)^k k^{o-1} \binom{O-1}{k-1} = 0.
\]

This is the same expression as in Eq. (A.1), where it is just multiplied by \( O \). This demonstrates that, if the relation holds for \( p-1 \), then it holds for \( p \). It is then sufficient to find a value of \( p \) for which the equation is true, to prove that it is true for any \( p \). For \( p = 0 \), the condition becomes

\[
\sum_{k=0}^{O} (-1)^k \binom{O}{k} = 0,
\]

where the left part is actually the development of \((1 - 1)^O\) and is obviously 0.
Appendix B: Image in the focal plane

To assess the device’s performance after insertion of the phase grids into the beam, it is convenient to calculate the image at the telescope focus. This image is obtained, at wavelength \( \lambda \), by squaring the Fourier transform module of the pupils. One formally obtains image \( \mathcal{I}(u, v) = \text{module}^2 \mathcal{I}(u, v) \). We have

\[
\text{module}(u, v) = \sin \left( \frac{\pi d}{n_0 A} \right) \sin \left( \frac{\pi d}{n_0 A} \right) \left| \text{Sum}(\expUV \cdot \expPQ) \right|.
\]  

(B.1)

In this expression, \( u \) and \( v \) are the coordinates on the focal plane, \( d \) the size of the mirror, \( n_0 \) the size of the phase grids, and the definition \( \sin(x) = \sin x / x \) was adopted. The remaining term involves the two square matrices \( \expUV \) and \( \expPQ \). The operation \( \cdot \) between the two matrices is the element by element product \( \text{not} \) the matrices product. The term \( \text{Sum}(\Delta) \) is just the sum of all the elements in a matrix \( A \) and, as usual, \( |z| = \sqrt{x^2 + y^2} \), with \( z \) the complex conjugate. The two matrices above in (B.1) have format \( n_0 \times n_0 \) and are given by

\[
\expUV = e^{j\pi d/n_0} \cdot e^{j\pi d/n_0},
\]

(B.2a)

\[
\expPQ = e^{j\pi \Delta x} \cdot e^{j\pi \Delta y} + e^{j\pi \Delta x} \cdot e^{j\pi \Delta y}.
\]

(B.2b)

We have adopted the convention that the exponential of a matrix is the matrix of the exponentials. Now, \( D \) is the distance between the two telescopes and \( \delta \) is a vector of integers varying from \(-n_0 - 1\) to \(+n_0 - 1\) in steps of 2. For example if \( n_0 = 2 \), we have \( \delta = [-3, -1, 1, 3] \). The symbol \( \otimes \) is the tensor product, accordingly \( \expUV \) is a square matrix constructed as a tensor product of two vectors (sometimes called the outer product or the cartesian product). Finally the chromatic term \( \Delta \) is

\[
\Delta = \frac{1}{1 + \lambda^2/\lambda_0^2}, \quad \Delta \lambda = \lambda - \lambda_0.
\]

(B.3)

The above expression is valid for the star of the system. For the planet, one must correct image \( \mathcal{I}(u, v) \) with the relative monochromatic luminosity at wavelength \( \lambda \) and, if one neglects the planet tilt on the telescope pupils, add the phase shift on one grid. Typically, we set the \( Q \) grid to \( Q + 1 \), to obtain maximum planet detectivity through a \( \pi \) phase shift.

Appendix C: First order: the Prouhet-Pascal solution

At a first-order approximation, we set \( D = 0 = \delta = 0 \), and the corresponding interferometer is an on-axis one, having all cells of each phase shift grids on top of the others. In this case, we get \( \expUV = 1 \) and \( \expPQ = e^{j\pi \Delta x} \cdot e^{j\pi \Delta y} \), leading to the chromatic term: \( \text{Sum} = \text{Sum}(e^{j\pi \Delta x} + e^{j\pi \Delta y}) \). As discussed earlier, its Taylor development is given by

\[
\text{Sum} = \sum_{k=1}^{N} \left( m_k^n - n_k^n \right) + \sum_{k=1}^{N} \left( m_k^n - n_k^n \right) \cdot (j\phi)^2 \sum_{k=1}^{N} \left( m_k^n - n_k^n \right) + \ldots \ (C.1)
\]

where \( \phi = \pi A \).

The device can be made achromatic up to order \( O \) if we simultaneously solve the above system of diophantine equations:

\[
\sum_{k=1}^{N} m_k = \sum_{k=1}^{N} n_k \quad \sum_{k=1}^{N} m_k^2 = \sum_{k=1}^{N} n_k^2 \quad \ldots \ldots \quad \sum_{k=1}^{N} m_k^n = \sum_{k=1}^{N} n_k^n.
\]

This system is called a multigrad system and is equivalent to the abbreviated notation: \( \{m_k, k = 1, N\} \equiv \{n_k, k = 1, N\} \), or simply \( \{m_k\} \equiv \{n_k\} \) if there is no ambiguity. Two sets of integers, solutions to a multigrad system, are called multigrade sets; for example, we have \( \{-11, -3, 3, 11\} \equiv \{-9, -7, 7, 9\} \). Solving a multigrad system corresponds to the classical Prouhet-Tarry-Escott problem. Our problem is a particular multigrad system, since the two sets \( \{m_k\} \) and \( \{n_k\} \) must have the same size and, more important, the \( n_k \) must be even and \( n_k \) odd.

The solution of our problem hinges on the property that if \( \{a_k\} \equiv \{b_k\} \), then \( \{a_k, b_k + c\} \equiv \{b_k, a_k + c\} \) for any \( c \). (These properties can be easily proved by induction using the fact that the highest power coefficient of a polynomial in \( x \) is invariant under the translation \( x + c \).) Starting from \( \{-1, 1\} \equiv \{0, 0\} \), we conserve parity by setting \( c = 1 \), and we get \( \{-1, 1, 1, 1\} \equiv \{0, 0, 0, 2\} \). For the next step, we regain symmetry by setting \( c = -1 \) and now get \( \{-1, -1, -1, 1, 1\} \equiv \{-2, 0, 0, 0, 0, 2\} \). What we use in fact is \( \{-1, 1\} \equiv \{0, 0\} \) and the property that if \( \{m_k\} \equiv \{n_k\} \) then \( \{n_k - 1, m_k, m_k + 1\} \equiv \{m_k - 1, n_k, n_k + 1\} \), the parity being preserved.

Appendix D: Second order: distributing the cells on the pupil

In the preceding appendix, we built a nulling interferometer with the even cells and the odd cells one on top of the others. The practical realisation of such an interferometer is conceivable, but is far too complicated to be considered here. The cells must therefore be distributed spatially on the even and odd pupils. Take for example the two \( 4 \times 4 \) grids \( n_0 = 4 \). The Prouhet-Pascal solution tells us to distribute the two sets \( \{-1, -1, -1, 1, 1, 1\} \) and \( \{-2, 0, 0, 0, 0, 0, 2\} \) on the pupils, but does not tell us where. To gain some insight into this difficult problem, it is necessary to scrutinise the \( \expUV \) matrix a step further and look at its expansion. We have

\[
\expUV = \sum_{k=0}^{\infty} \frac{1}{k!} \left( \frac{\pi d}{n_0 A} \right)^k \lbrack UV \rbrack^k,
\]

(D.1)

with (always for the example \( n_0 = 4 \) case):

\[
\lbrack UV \rbrack^k = \left[ \begin{array}{cccc}
(-3u + 3v)^k & (-u + 3v)^k & (u + 3v)^k & (3u + 3v)^k \\
(-3u - v)^k & (-u - v)^k & (u - v)^k & (3u - v)^k \\
(-3u - c)^k & (-u - c)^k & (u - c)^k & (3u - c)^k \\
(-3u - 3v)^k & (-u - 3v)^k & (u - 3v)^k & (3u - 3v)^k
\end{array} \right].
\]

(D.2)

We observe that the even-order matrices \( \lbrack UV \rbrack^k \) are axisymmetric, while the \( \lbrack UV \rbrack^{2p+1} \) odd ones are anti-axisymmetric. Therefore it easy to nullify all the contributions from the odd ones in (B.1), and it suffices to distribute the cells in an axisymmetric manner. This cancels out the first-order term in (D.1) and all the odd ones but, of course, the even ones still contribute. Next, we observe that the \( \lbrack UV \rbrack^k \) matrices are 2D polynomials in \( u, v \) coordinates, therefore they will be cancelled out by a 2D finite differential operator. Following these lines, we start with the \( 2 \times 2 \) matrices corresponding to the doubled \( \{-1, 1\} \equiv \{0, 0\} \) solution:

\[
P = \begin{bmatrix}
0 & 0 \\
0 & 0
\end{bmatrix}, \quad Q = \begin{bmatrix}
-1 & 1 + 1 \\
1 & -1
\end{bmatrix}.
\]
Indeed, \( \exp P Q \approx P - Q \) is a differential (gradient) operator. We obtain the next matrix by using the \([n_k - 1, m_k, m_k, n_k + 1]^{\alpha^2} [m_k - 1, n_k, n_k, m_k + 1] \) property. Since there were still some degrees of freedom left, we found the best arrangement by keeping the \( P - Q \) differential operator property in order to minimize the star/planet ratio of the residual light. We get

\[
P = \begin{bmatrix}
+2 & 0 & 0 & -2 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
-2 & 0 & 0 & +2
\end{bmatrix},
\]
\[
Q = \begin{bmatrix}
+1 & +1 & -1 & -1 \\
+1 & -1 & +1 & -1 \\
-1 & +1 & -1 & +1 \\
-1 & -1 & +1 & +1
\end{bmatrix}.
\]

The algorithm of the third method proceeds following these lines from an \( n_0 \times n_0 \) matrix to the next \( 2n_0 \times 2n_0 \) one. This algorithm is very efficient because it produces a very low star/planet ratio. Moreover it possesses the desirable property of producing axisymmetrical images which helps finding the location of the planet.

References

Borwein, P., & Ingalls, C. 1994, Enseign. Math., 40, 3
Bracewell, R. N. 1978, Nature, 274, 780

D. Rouan and D. Pelat: The achromatic chessboard. I. 589