A&A 524, A80 (2010)
DOI: 10.1051/0004-6361/201014934

Astronomy & Astrophysics

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

II. Theoretical performance assessment

D. Pelat[1], D. Rouan[2], and D. Pickel[1,2]

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

Received 4 May 2010 / Accepted 11 August 2010

ABSTRACT

Context. Nulling interferometry in the mid-IR using two telescopes (commonly referred to a Bracewell interferometer) is one possible way of directly detecting exoplanets in the habitable zone and their characterisation in terms of possible life signatures. A large wavelength domain is needed to simultaneously detect the infrared spectral features of a set of a bio-tracers. An achromatic phase shift of π is then required, and we previously presented a new concept for such a function that allows a simple design with only one device per beam. It is based on two cellular mirrors, called the chessboards, where each cell has a thickness that introduces, for any given central wavelength, a phase shift of $(2k + 1)\pi$ or of $2k\pi$ on the fraction of the wave it reflects.
Aims. We explore a more rigorous way to establish the optimum cell pattern design to attain the best theoretical performances for planet detection over a broad wavelength range. Two possible types of interferometres are now considered: on-axis and multi-axis.
Methods. We derived a rather simple iterative scheme for both designs, determining the thickness and XY position of the cells. The method confers to the chessboards a high degree of internal symmetry. Each design can be described as an *iterative Bracewell interferometer* characterised by an integer order. We demonstrate that their efficiencies increases with the power of that order.
Results. The device acts both spatially and versus wavelengths as an optical differential operator on the 3D light distribution. Its power is best understood in the on-axis case since its effect is to *push away* the stellar light from the centre over a very broad range of wavelengths, leaving space for an *out of phase* object to appear in the *cleaned* central region. We explore the theoretical performances for on-axis and multi-axis designs in the parameter space, and we especially compute the rejection factor for starlight and the attenuation factor for planet light and introduce the *relative nulling efficiency* metric. We show that, even with some realistic piston error added, the performances could meet the *Darwin* space project specifications for both designs, i.e., cancellation of the starlight by a factor of 10^5 over a wavelength range of 6–17 μm.

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

1. Introduction

A new concept of a quasi-achromatic nulling interferometer has been proposed in a preceding publication (Rouan & Pelat 2008, hereafter Paper I). The whole idea is based on cancelling the polynomial part, according to both the wavelength and spatial coordinates, of the starlight distribution at the centre of the focal plane of the interferometer. This objective was attained thanks to a particular setting of phase shifters on the interferometer pupil planes. The advantage of the proposed design depends on a rather simple arrangement, where the key device is a simple square grid of phase shifters (the so-called "chessboard") on the pupil plane of the telescopes.

In Paper I, several possible chessboards were proposed and a co-axial (Michelson) setting of the telescopes favoured. Here we concentrate more deeply on two high-performance configurations, one designed for the co-axial case and the other for the multi-axial (Fizeau) case. In Sect. 2, we outline the principles that led to the phase distribution and spatial repartition of the chessboards' cells. A first approximation of the system performances is also given, with particular emphasis on the principles that govern its effectiveness. In Sect. 3, detailed analytical calculations are carried out to determine the chessboards' nulling performances with more precision and to explore their robustness to a systematic or random piston on one arm of the interferometer.

2. The chessboard design

We recall that our device consists of two square grids of phase shifters each formed of n_0^2 adjacent cells. Finding a high-performance chessboard setting implies closely inspecting the complex amplitude of the electrical field of the interferometer output equipped with this device. If A stands for this amplitude, the point spread function is given by PSF $= |A|^2$. According to Fourier optics one gets

$$A(u,v) = \mathrm{Sum}(\exp \mathrm{UV} \cdot \exp \mathrm{PQ}) \mathrm{sinc}\left(\frac{\pi d}{n_0 \lambda} u\right) \mathrm{sinc}\left(\frac{\pi d}{n_0 \lambda} v\right). \quad (1)$$

In this expression, u and v are the Cartesian angular coordinates of a light ray diffracted by the pupil plane, d is the size of the pupil, n_0 the number of cells on one side of the chessboards,

and by definition $\mathrm{sinc}(x) = \sin(x)/x$. The symbolic expressions expUV and expPQ stand for two square matrices both with the "chessboard" format: $n_0 \times n_0$. The "dot" operation $\cdot$ between the two matrices is the element-to-element product *not* the matrices product. The term Sum( ) realises the sum of all elements of the matrix under the scope of that operator.

The matrix expUV takes the *positioning* of the cells into account, namely the place where the phase shifters are put on the grid of the chessboards. It is given by

$$\mathrm{expUV} = \mathrm{e}^{j\pi u\delta/n_0\lambda} \otimes \mathrm{e}^{j\pi v\delta/n_0\lambda}. \tag{2}$$

In this expression, δ is the displacement vector. It is a set of integers varying from $-(n_0 - 1)$ to $+(n_0 - 1)$ by step of 2, and the symbol $\otimes$ is the tensor product between the two vectors. For example, if $n_0 = 4$, we have $\delta = (-3, -1, 1, 3)$ and we get

$$\mathrm{expUV} = \begin{bmatrix} \mathrm{e}^{j\frac{3\pi d}{n_0\lambda}v} \\ \mathrm{e}^{j\frac{\pi d}{n_0\lambda}v} \\ \mathrm{e}^{j\frac{-\pi d}{n_0\lambda}v} \\ \mathrm{e}^{j\frac{-3\pi d}{n_0\lambda}v} \end{bmatrix} \begin{bmatrix} \mathrm{e}^{j\frac{-3\pi d}{n_0\lambda}u} & \mathrm{e}^{j\frac{-\pi d}{n_0\lambda}u} & \mathrm{e}^{j\frac{\pi d}{n_0\lambda}u} & \mathrm{e}^{j\frac{3\pi d}{n_0\lambda}u} \end{bmatrix}, \tag{3}$$

$$= \begin{bmatrix} \mathrm{e}^{j\frac{\pi d}{n_0\lambda}(-3u+3v)} & \mathrm{e}^{j\frac{\pi d}{n_0\lambda}(-u+3v)} & \mathrm{e}^{j\frac{\pi d}{n_0\lambda}(u+3v)} & \mathrm{e}^{j\frac{\pi d}{n_0\lambda}(3u+3v)} \\ \mathrm{e}^{j\frac{\pi d}{n_0\lambda}(-3u+v)} & \mathrm{e}^{j\frac{\pi d}{n_0\lambda}(-u+v)} & \mathrm{e}^{j\frac{\pi d}{n_0\lambda}(u+v)} & \mathrm{e}^{j\frac{\pi d}{n_0\lambda}(3u+v)} \\ \mathrm{e}^{j\frac{\pi d}{n_0\lambda}(-3u-v)} & \mathrm{e}^{j\frac{\pi d}{n_0\lambda}(-u-v)} & \mathrm{e}^{j\frac{\pi d}{n_0\lambda}(u-v)} & \mathrm{e}^{j\frac{\pi d}{n_0\lambda}(3u-v)} \\ \mathrm{e}^{j\frac{\pi d}{n_0\lambda}(-3u-3v)} & \mathrm{e}^{j\frac{\pi d}{n_0\lambda}(-u-3v)} & \mathrm{e}^{j\frac{\pi d}{n_0\lambda}(u-3v)} & \mathrm{e}^{j\frac{\pi d}{n_0\lambda}(3u-3v)} \end{bmatrix}. \tag{4}$$

The second matrix expPQ, takes the *values* of the phase shifters into account. Adopting the usual convention that the exponential of a matrix is the matrix of the exponentials, we have

$$\mathrm{expPQ} = \mathrm{e}^{j\pi P\Lambda}\mathrm{e}^{-j\pi uD/\lambda} + \mathrm{e}^{j\pi Q\Lambda}\mathrm{e}^{+j\pi uD/\lambda}. \tag{5}$$

The matrices P and Q are the chessboard matrices, of size $n_0 \times n_0$. All their elements are integers that give, in units of $\lambda_0/2$, the optical path difference (OPD) induced by the corresponding cell of the chessboard. For example, if $P = [0]$ and $Q = [1]$, there is no OPD on one telescope but one of $\lambda_0/2$ on the other one, inducing a phase shift of π at λ_0. This particular setting corresponds to a *Bracewell* nulling interferomer (Bracewell 1978). Finally, D is the distance between the two pupils' centres and Λ is a chromatic variable

$$\Lambda = \frac{\lambda_0}{\lambda} = \frac{1}{1 + \frac{\Delta\lambda}{\lambda_0}}, \quad \Delta\lambda = \lambda - \lambda_0. \tag{6}$$

(The Λ variable was named Δ in Paper I.)

To obtain the best achromatism, we propose a progressive approach below where we first look for the best phase-shift distribution and second for their spatial distribution. We start our investigation with the mathematically simpler co-axial design.

2.1. The co-axial chessboards

In the co-axial configuration, the two pupils are superimposed thanks to some optical arrangement. A practical example is given in Fig. 1, where it includes an imaging lens in the last stage since any efficient detection requires concentrating the light beam on a small surface. To consider this design, we set $D = 0$ in (5). The expPQ matrix therefore reduces to $\mathrm{e}^{j\pi P\Lambda} + \mathrm{e}^{j\pi Q\Lambda}$. As a first step we do not consider the phase shifters' location on the pupils; i.e., we set the displacement vector δ to zero making all elements of expUV equal to one. Accordingly, the Sum( ) term in (1) takes the simpler form of $\mathrm{Sum}(\mathrm{e}^{j\pi P\Lambda} + \mathrm{e}^{j\pi Q\Lambda})$, which we call *the chromatic term*.

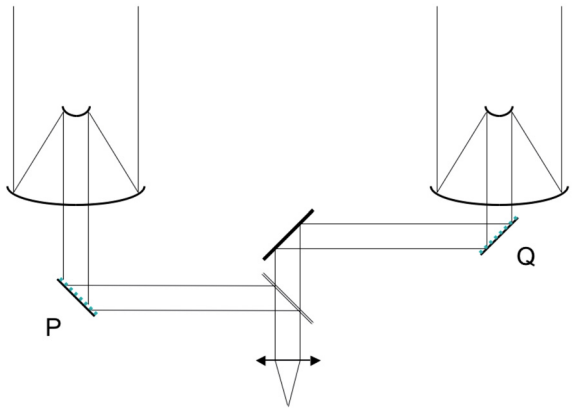

Fig. 1. A possible optical scheme of the nulling interferometer in co-axial configuration.

2.1.1. The multigrad solution

The idea, explained in Paper I, is to cancel the chromatic term up to the Nth order of its Taylor series. For this purpose, we need the Taylor series expansion of the elements of $\mathrm{e}^{j\pi P\Lambda}$ and $\mathrm{e}^{j\pi Q\Lambda}$ matrices. For example, if n_s stands for the numerical values contained in P, the Taylor expansion, relative to $\varepsilon = \Delta\lambda/\lambda_0$, is given by

$$\mathrm{e}^{j\pi n_s\Lambda} = \sum_{k=0}^{\infty} \tfrac{1}{k!}\varepsilon^k \mathrm{e}^{j\pi n_s} p_k(j\pi n_s), \tag{7}$$

where p_k stands for a polynomial of degree k. (Note that p_k is not a monomial because of the chain rule needed to compute the derivatives.) Of course, a similar formula exists for $\mathrm{e}^{j\pi m_s\Lambda}$ involving the values m_s contained in Q. Cancelling the chromatic term up to order N implies solving the equation

$$\sum_{s=1}^{n_0^2} \sum_{k=0}^{N} \tfrac{1}{k!}\varepsilon^k \left(\mathrm{e}^{j\pi n_s} p_k(j\pi n_s) + \mathrm{e}^{j\pi m_s} p_k(j\pi m_s)\right) = 0, \text{ or} \tag{8}$$

$$\sum_{k=0}^{N} \tfrac{1}{k!}\varepsilon^k \sum_{s=1}^{n_0^2} \left(\mathrm{e}^{j\pi n_s} p_k(j\pi n_s) + \mathrm{e}^{j\pi m_s} p_k(j\pi m_s)\right) = 0. \tag{9}$$

This polynomial is zero for any ε if, and only if, all its coefficients are zero. Therefore we have to solve the system

$$\sum_{s=1}^{n_0^2} \left(\mathrm{e}^{j\pi n_s} p_k(j\pi n_s) + \mathrm{e}^{j\pi m_s} p_k(j\pi m_s)\right) = 0, \tag{10}$$

for $k = 0, \ldots, N$. If $c_{\ell k}$ stands for the coefficients of the polynomials p_k, this reduces to

$$\sum_{\ell=0}^{k} c_{\ell k} (j\pi)^\ell \sum_{s=1}^{n_0^2} (\mathrm{e}^{j\pi n_s} n_s^\ell + \mathrm{e}^{j\pi m_s} m_s^\ell) = 0. \tag{11}$$

This system can be solved by induction, because if a solution is found up to order k, the solution up to order $k + 1$ is found if

$$\sum_{s=1}^{n_0^2} (\mathrm{e}^{j\pi n_s} n_s^{k+1} + \mathrm{e}^{j\pi m_s} m_s^{k+1}) = 0. \tag{12}$$

Therefore, the chromatic term is nullified up to order N if one solves the following system of equations

$$\sum_{s=1}^{n_0^2} (\mathrm{e}^{j\pi n_s} n_s^k + \mathrm{e}^{j\pi m_s} m_s^k) = 0\,; \quad k = 0, \ldots, N, \tag{13}$$

for the unknown integers n_s and m_s and the parameter n_0.

As in a Bracewell interferometer, we impose the constraint that $\mathrm{e}^{j\pi n_s} = +1$ and $\mathrm{e}^{j\pi m_s} = -1$, therefore the n_s and m_s are respectively even and odd integers. As introduced in Paper I, Appendix C, one ends up with a special set of diophantine equation called a *multigrad* system, namely

$$\sum_{s=1}^{n_0^2} n_s^k = \sum_{s=1}^{n_0^2} m_s^k, \; k = 1, \ldots, N. \tag{14}$$

If two sets of integers, say $\{n_s\}$ and $\{m_s\}$, solve a multigrad system up to order N, they are called multigrad sets. This property is shortly written as $\{m_s\} \overset{N}{=} \{n_s\}$. Hereafter, the two sets are supposed to be of equal size, for short $\{m_s\} \overset{0}{=} \{n_s\}$ (adopting the convention that $0^0 = 1$).

The key theorem is that if one has $\{n_s\} \overset{N}{=} \{m_s\}$, then it follows that $\{n_s, m_s + c\} \overset{N+1}{=} \{m_s, n_s + c\}$. This is true for any c, but is nontrivial only if $c \neq 0$ and $\{m_s\} \neq \{n_s\}$. In this way, two multigrad sets up to order N generate two new multigrad sets up to order $N + 1$. During this operation, the size of the sets is multiplied by two. But in order that the number of cells on the chessboards remains a square, the size of the sets must be multiplied by four. This is done by applying two times the key theorem. Moreover, if the sets $\{n_s\}$ and $\{m_s\}$ possess the desired parity, we keep it if we choose an odd constant c, say $c = 1$. We get the final result:

if $\{n_s\} \overset{N}{=} \{m_s\}$, then

$\{n_s, m_s + 1, m_s + 1, n_s + 2\} \overset{N+2}{=} \{m_s, n_s + 1, n_s + 1, m_s + 2\}$.

This remarkable property is at the heart of the scheme proposed below.

2.1.2. The iterative Bracewell scheme

This scheme runs as follows. One starts with a Bracewell nulling interferometer $\{0\} \overset{0}{=} \{-1\}$ and constructs the next one by two applications of the key theorem. In this way, one first gets $\{0, 0, 0, 2\} \overset{2}{=} \{-1, 1, 1, 1\}$. Indeed, we have (always with $0^0 = 1$)

$$\begin{aligned} 0^0 + 0^0 + 0^0 + 2^0 &= -1^0 + 1^0 + 1^0 + 1^0, \\ 0^1 + 0^1 + 0^1 + 2^1 &= -1^1 + 1^1 + 1^1 + 1^1, \\ 0^2 + 0^2 + 0^2 + 2^2 &= -1^2 + 1^2 + 1^2 + 1^2. \end{aligned} \tag{15}$$

As expected, this new interferometer cancels the chromatic term up to order $N = 2$. To construct the next one, we add $c = -1$ instead of $c = +1$ in order to gain symmetry and in an effort to keep the z-heights of the cells as low as possible. Finally, to generate the successive phase chessboards, we propose the following iterative scheme

$$\begin{aligned} P_0 &= \{0\}, \quad Q_0 = \{-1\}, \\ P_{n+1} &= \{P_n, Q_n \pm 1, Q_n \pm 1, P_n \pm 2\}, \\ Q_{n+1} &= \{Q_n, P_n \pm 1, P_n \pm 1, Q_n \pm 2\}, \end{aligned} \tag{16}$$

where $\pm$ means $(-1)^n$ (i.e. $+1$ if n is even and -1 if it is odd). With this process, the chessboards P_n and Q_n retain the shape of two squares possessing $n_0 = 2^n$ cells on each of their sides. Their combined effect is to cancel the chromatic term up to order $N = 2n$. The higher n, the more efficent the corresponding nulling interferometer.

The iterative scheme presented above leads to constructing the interferometer (P_n, Q_n) possessing, for example, the phase-shift distribution shown on Table 1. We observe that the progression of the number of cells that produce the same phase shift follows that of the binomial coefficients in the polynomial expansion of $(1 + x)^{2n+1}$. This is easily proved by two applications of the recursive formula $\binom{2n}{q} = \binom{2n-1}{q-1} + \binom{2n-1}{q}$, which leads to $\binom{2n+1}{q} = \binom{2n-1}{q-2} + 2\binom{2n-1}{q-1} + \binom{2n-1}{q}$, which is our iterative scheme. (In these formulæ we adopt the convention that $\binom{n}{q} = 0$ if $q \leq 0$.) Table 2 gives the phase-shift, together with the corresponding number of cells (the formulæ are valid for two consecutives orders).

Table 1. Number of cells in the interferometer (P_n, Q_n) according to the phase shift $\exp(jk\pi\lambda_0/\lambda)$.

n	k							
	-3	-2	-1	**0**	1	**2**	3	**4**
0			1	**1**				
1			1	**3**	3	**1**		
2	1	**5**	10	**10**	5	**1**		
3	1	**7**	21	**35**	35	**21**	7	**1**

Notes. The cells of the "even" P_n pupil are shown in boldface. The number of cells, alternately odd and even, follows the binomial coefficients.

Table 2. Number of cells in (P_n, Q_n) producing a given phase shift.

Interferometer order n	$e^{jk\pi\lambda_0/\lambda}$	$e^{-j(2p+1)\pi\lambda_0/\lambda} e^{jq\pi\lambda_0/\lambda}$
0, 1	$\binom{2n+1}{k+1}$	$\binom{2n+1}{q}$
2, 3	$\binom{2n+1}{k+3}$	$\binom{2n+1}{q}$
$2p, 2p+1$	$\binom{2n+1}{k+2p+1}$	$\binom{2n+1}{q}$

Notes. Column 2 shows the number of cells producing a $e^{jk\pi\lambda_0/\lambda}$ phase shift. To obtain more tractable binomial coefficients, Col. 3 is obtained with $q = k + 2p + 1$.

2.1.3. First estimation of the iterative Bracewell scheme performance

In this section we study the nulling efficiency of an interferometer built on this principle. To compare the interferometers (P_n, Q_n) in the same conditions, differing only by their order n, the light collecting surface of their mirrors must be kept constant by reducing the size of the cells to a square of surface $d^2/2^{2n}$. Because at this point we have not distributed the cells on the pupil plane, we only consider A_n, the sum of the amplitudes. This leads to the formula

$$A_n = \frac{d^2}{2^{2n}} e^{-jK\pi\lambda_0/\lambda} \sum_{q=0}^{2n+1} \binom{2n+1}{q} e^{jq\pi\lambda_0/\lambda}, \tag{17}$$

$$= 2d^2 e^{-jK\pi\lambda_0/\lambda} \left(\frac{1+z}{2}\right)^{2n+1}, \tag{18}$$

where we have set $z = e^{j\pi\lambda_0/\lambda}$ and $K = 2\left\lfloor \frac{n}{2} \right\rfloor + 1$. The common phase factor $e^{-jK\pi\lambda_0/\lambda}$ sets the OPD reference, and the function $\lfloor x \rfloor$ is the "floor" function (i.e. it returns the greatest integer less than or equal to x.)

If $\lambda = \lambda_0$, then $z = -1$ and $A_n = 0$, therefore the nulling is perfect for any n. The device is strictly equivalent to a Bracewell nulling interferometer. Now, if $\lambda \neq \lambda_0$, since the module of the complex $(1 + z)/2$ is less than one, $A_n \to 0$ as n, the order of the interferometers, increases. This is true for all λ except for the values such that $z = 1$, i.e. $\lambda = \infty, \lambda_0/2, \lambda_0/4, \ldots, \lambda_0/2k, \ldots$ Aside from the actual distribution of cells on the pupils, we have obtained a quasi-achromatic nulling interferometer.

However, one must judge the perfomances of the device relative to some benchmark. We have demonstrated above that, for an on-axis object, our device has powerful nulling efficiency on a given set of wavelengths. This is of course different for any off-axis object such as a planet which is best detected when its position induces an OPD= $D\,\theta_{\text{planet}} = \lambda_0/2$. In that situation, the recombination, at λ_0, is fully constructive for the planet. Therefore, the interferometer with a phase shift of π at λ_0 added to one of

its pupils will serve as a reference. In so doing, we transform the nulling interferometer into a non-destructive one for the wavelengths set that was just mentioned above. The phase addition on all the, say, odd phase shifters changes the sign of the odd, $q = 2p + 1$, binomial coefficients in the expression of A_n. If A'_n stands for the function obtained in this way, one gets

$$A'_n = 2d^2 \mathrm{e}^{-jK\pi\lambda_0/\lambda} \left(\frac{1-z}{2}\right)^{2n+1} \cdot \tag{19}$$

The only change between the expressions of A_n and A'_n is the sign of z. If $\lambda = \lambda_0/2$, then $z = 1$ and $A'_n = 0$, we also have a nulling interferometer. Again, if $\lambda \neq \lambda_0/2$, since the modulus of the complex $(1-z)/2$ is also less than one, $A'_n \to 0$ as n increases. This is true for all λ except for the values that make $z = -1$; that is, $\lambda = \lambda_0, \lambda_0/3, \lambda_0/5, \ldots, \lambda_0/(2k+1), \ldots$ The result is an associated "out of phase" nulling interferometer but, fortunately, less efficient than the original one. The off-axis planet will indeed suffer some attenuation but much less than the on-axis star, as the ratio of A_n to A'_n demonstrates.

We define ρ_n as A_n/A'_n, and name it the *relative nulling efficiency*. One gets

$$\rho_n(z) = \left(\frac{1+z}{1-z}\right)^{2n+1} = \rho_0^{2n+1}(z). \tag{20}$$

If $|\rho_0|$ is less than one and n large enough, the on-axis nulling can be made more efficient than the off-axis one by any factor. Note that $\rho_0(z)$ is the relative nulling efficiency of the original Bracewell interferometer.

A deeper understanding of how the iterative Bracewell interferometers works is obtained through careful analysis of the transformation induced by ρ_0 on the complex plane $\mathbb{C}$ of the Fresnel representation. In fact, under ρ_0, the unit circle is transformed into the imaginary axis[1]. One gets

$$\rho_0(z) = j\cot\Big(\frac{1}{2}\arg(z)\Big), \quad \arg(z) = \pi\frac{\lambda_0}{\lambda} \cdot \tag{21}$$

This means that $\rho_0(z)$, and the Bracewell relative nulling efficency that follows is obtained by a stereographic projection of z toward the real 1 (see Fig. 2). The image of the half-circles $\frac{\pi}{2} < \theta \pmod{2\pi} < \frac{3\pi}{2}$ is the segment $(-j, j)$ shown in green in the figure. All points in this segment have a modulus less than one, therefore $|\rho_n(z)|$ can be made as small as desired, provided one chooses an interferometer of sufficiently high-order n. This defines many wavelengths orders, the first one being: $\frac{2}{3}\lambda_0 < \lambda < 2\lambda_0$. Within these bandwiths, an "in phase" object, say a star, is strongly attenuated while an "out of phase" one, say a planet, is much less affected. This property opens interesting perspectives for broadband spectroscopy, or detection, of exoplanets, all the more so for intensity, since the nulling behaves as $|\rho_0|^{4n+2}$.

From the dispersion of the optical paths difference ΔOPD produced by the chessboards, it is possible to obtain a very good approximation of the full width at half maximum (*FWHM*) of the planet attenuation spectrum within the full bandwidths, e.g. $(\frac{2}{3}\lambda_0, 2\lambda_0)$. Because the OPDs distribution is binomial, one gets $\sigma(\mathrm{OPD}) = \sqrt{2n+1}\lambda_0/4$ or approximately $\Delta\mathrm{OPD} \propto \sqrt{n}\lambda_0$. Now, $\Delta\mathrm{OPD} = c\Delta t$ and $\Delta E = h\Delta\nu = hc\Delta\lambda/\lambda^2 \approx hc\Delta\lambda/\lambda_0^2$. By applying the Heisenberg uncertainty principle, we have $\Delta t\Delta E \geq h/4\pi$, from which we get

$$\frac{\Delta\lambda}{\lambda_0} \propto \frac{1}{\sqrt{n}} \cdot \tag{22}$$

The constant of proportionality should not be very different from one, see Fig. 7 for an example where it is approximately equal to $1/\sqrt{2}$. While the planet-to-star ratio greatly increases with the order n of the chessboards, the effective bandwidth reduction in $\sqrt{n}$ implies longer exposure times to get the full range of wavelengths.

Finally, the proposed scheme generates, from one order to the next, a fourfold increase in the number of cells. In Appendix A, we investigate to what extent so many cells are necessary.

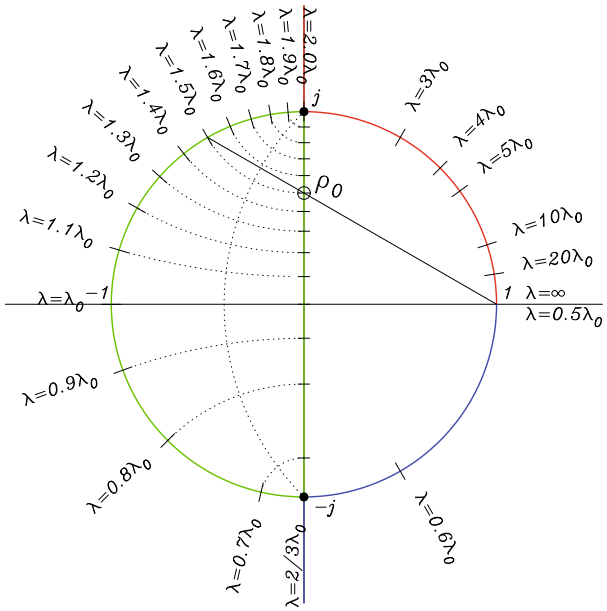

Fig. 2. Graphic determination of the relative nulling efficiency ρ_0 of a Bracewell nulling interferometer through a stereographic projection of the Fresnel vector $z = \mathrm{e}^{\pi\lambda_0/\lambda}$ toward point 1 on the real axis. The relative nulling efficiency of the *iterative* Bracewell scheme (P_n, Q_n) is ρ_0^{2n+1}. This quantity decreases to 0, as n increases where $|\rho_0| < 1$. That is when $\rho_0(z)$ is in the $(-j, j)$ segment. Therefore, a planet dephased by π at λ_0 relative to its star may be detected in many bandwidths whose first order is $\frac{2}{3}\lambda_0 < \lambda < 2\lambda_0$.

2.1.4. Placing the cells on the chessboards

To actually build the interferometer, it is necessary to distribute the chessboards cells on the pupil plane, an operation that leads to a specific PSF on the image plane. The previous performance analysis is valid for a perfectly parallel light beam, therefore, for any distribution of cells, it is only relevant for the central pixel of the image. For the other pixels, it is necessary to ensure that the cells distribution generates a PSF that does not ruin the mighty nulling power of the chessboards as derived above. This can be done by extending the same idea to the *spatial* dimension of the amplitude A, as given by Eq. (1).

As mentioned in Paper I, we gain insight into this problem by performing a Taylor expansion of the expUV matrix to get

$$\mathrm{expUV} = \sum_{k=0}^{\infty} \frac{1}{k!}\Big(j\frac{\pi d}{n_0\lambda}\Big)^k [\mathrm{UV}]^k. \tag{23}$$

[1] More precisely, ρ_0 performs a Möbius transformation of $\mathbb{C}$. In that case, it is a reflection (inversion) relative to the circle of center 1 and radius $\sqrt{2}$, followed by a reflection about the imaginary axis (see e.g. Ratcliffe, J. G. 1994, Chap. 4). This transformation leaves invariant the two points j and $-j$.

In this expression, the symbol [UV] stands for a *positioning* matrix of size $n_0 \times n_0$. For example, in the case $n_0 = 4$, we have

$$[\mathrm{UV}] = \begin{bmatrix} -3u+3v & -u+3v & u+3v & 3u+3v \\ -3u+v & -u+v & u+v & 3u+v \\ -3u-v & -u-v & u-v & 3u-v \\ -3u-3v & -u-3v & u-3v & 3u-3v \end{bmatrix}. \quad (24)$$

By convention the elements of $[\mathrm{UV}]^k$ are the elements of [UV] raised to the power k. Following the above example, we get

$$[\mathrm{UV}]^k = \begin{bmatrix} (-3u+3v)^k & (-u+3v)^k & (u+3v)^k & (3u+3v)^k \\ (-3u+v)^k & (-u+v)^k & (u+v)^k & (3u+v)^k \\ (-3u-v)^k & (-u-v)^k & (u-v)^k & (3u-v)^k \\ (-3u-3v)^k & (-u-3v)^k & (u-3v)^k & (3u-3v)^k \end{bmatrix}. \quad (25)$$

Again, the key observation is to recognise that the elements of $[\mathrm{UV}]^k$ are the sampled values of the 2-D polynomial $(x+y)^k$. Therefore, if we were able to arrange the cells in such a way that the expPQ matrix would turn out to be the coefficients of a *finite difference 2-D differential operator* of order $k+1$, the expression $\mathrm{expPQ} \cdot [\mathrm{UV}]^k$ would be zero. Indeed all terms of degree r less than or equal to k in (23) would also be cancelled, since $\mathrm{expPQ} \cdot [\mathrm{UV}]^r$ would be the finite difference derivative of order $k+1$ of a 2-D polynomials of degree less than or equal to k.

It seems, however, impossible to attain this objective exactly, but it is possible to approach it fairly closely. The first-order approximation, according to $\varepsilon = \Delta\lambda/\lambda_0$, of expPQ is

$$\mathrm{expPQ} \approx \mathrm{e}^{j\pi P}(1 - \varepsilon j\pi P) + \mathrm{e}^{j\pi Q}(1 - \varepsilon j\pi Q). \quad (26)$$

In the case where P is even and Q odd, this formula reduces to $\mathrm{expPQ} \approx -\varepsilon j\pi(P-Q)$. It is not too difficult to find an arrangement of cells on the pupil plane that makes $P-Q$ equal to a matrix whose elements are those of a finite difference differential operator (*fiddop*). We found that it is achievable by the following iterative scheme

$$P_0 = [0], \quad Q_0 = [-1],$$
$$P_{n+1} = \begin{bmatrix} Q_n \pm 1 & P_n \pm 2 \\ P_n & Q_n \pm 1 \end{bmatrix}, \quad Q_{n+1} = \begin{bmatrix} P_n \pm 1 & Q_n \pm 2 \\ Q_n & P_n \pm 1 \end{bmatrix}. \quad (27)$$

In this expression $\pm$ again means $(-1)^n$. (This scheme is the same as the iterative Bracewell one discussed above. Also $P_n \pm 1 = -Q_n^{\mathrm{T}}$, $Q_n \pm 1 = -P_n^{\mathrm{T}}$, where e.g. P_n^{T} stands for the transposition of P_n.) It follows, by induction, that $P_n - Q_n$ is, as desired, a *fiddop* of order n. To start the induction, we have

$$P_1 - Q_1 = \begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix} - \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} -1 & +1 \\ +1 & -1 \end{bmatrix}. \quad (28)$$

It is indeed a gradient operator that, as such, will cancel the first-order (constant) term in the Taylor expansion of expUV. More generally we have

$$P_{n+1} - Q_{n+1} = \begin{bmatrix} Q_n \pm 1 & P_n \pm 2 \\ P_n & Q_n \pm 1 \end{bmatrix} - \begin{bmatrix} P_n \pm 1 & Q_n \pm 2 \\ Q_n & P_n \pm 1 \end{bmatrix} \quad (29)$$
$$= \begin{bmatrix} Q_n - P_n & P_n - Q_n \\ P_n - Q_n & Q_n - P_n \end{bmatrix}. \quad (30)$$

Therefore, $P_{n+1} - Q_{n+1}$ is the gradient of $P_n - Q_n$ and, since the later is a *fiddop* of order n, then $P_{n+1} - Q_{n+1}$ is a *fiddop* of order $n+1$. It is remarkable that all lines and columns of $P_n - Q_n$ are a Prouhet-Thue-Morse sequence of -1 and $+1$ of length 2^n, as is clear from the way the $P_n - Q_n$ were constructed.

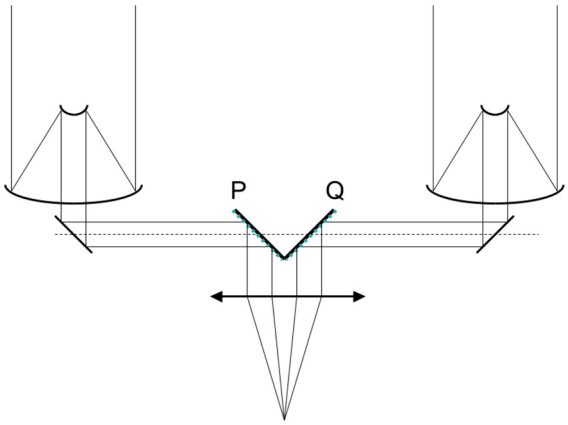

Fig. 3. A possible optical scheme of the nulling interferometer in multi-axial configuration.

We recall that the Prouhet-Thue-Morse sequence P_n is a string of length 2^n whose kth element is given by $P_n(k) = (-1)^{s(k)}$ for $0 \le k < 2^n$, where $s(k)$ is the sum of the binary digits of k; e.g. $P_3 = +1, -1, -1, +1, -1, +1, +1, -1$. This sequence is also constructed iteratively by $P_0 = +1$ and $P_n = P_{n-1} \| -P_{n-1}$, where $\|$ is the concatenation operator (Prouhet 1851; Allouche & Mendès-France 2008).

2.2. The multi-axial chessboards

For this design, we set $D \neq 0$ in (5) (see Fig. 3 for a practical realisation). It is convenient to write the expPQ matrix as

$$\mathrm{expPQ} = \cos(\pi uD)(\mathrm{e}^{j\pi P\Lambda} + \mathrm{e}^{j\pi Q\Lambda}) + j\sin(\pi uD)(-\mathrm{e}^{j\pi P\Lambda} + \mathrm{e}^{j\pi Q\Lambda}). \quad (31)$$

From this we get

$$\mathrm{Sum}(\mathrm{expPQ} \cdot \mathrm{expUV}) = \cos(\pi uD)\ \mathrm{Sum}\left((\mathrm{e}^{j\pi P\Lambda} + \mathrm{e}^{j\pi Q\Lambda}) \cdot \mathrm{expUV}\right) - j\sin(\pi uD)\ \mathrm{Sum}\left((\mathrm{e}^{j\pi P\Lambda} - \mathrm{e}^{j\pi Q\Lambda}) \cdot \mathrm{expUV}\right). \quad (32)$$

In the first term of the right-hand side member, we recognise the co-axial chromatic term defined above. We know how to cancel it, up to a given polynomial order. One way to cancel the second term is to generate an electrical field amplitude A, which is antisymmetric with respect to u (direction of the interferometer base), and to perform a sum of the electric field *prior* to its quadratic detection. Such a sum can be achieved in an elegant and simple way thanks to a *single-mode* fiber optics. The efficiency of this cancelling device has actually been demonstrated (e.g., Wallner et al. 2004; Buisset et al. 2006).

To obtain $A(-u,v) = A(u,v)$, and because the $\sin(\pi uD)$ function is odd, it is sufficient to make sure that $\mathrm{Sum}\left((\mathrm{e}^{j\pi P\Lambda} - \mathrm{e}^{j\pi Q\Lambda}) \cdot \mathrm{expUV}(u,v)\right)$ is even. It is clear from Eq. (3) that $\mathrm{expUV}(-u,v)$ is equal to $\mathrm{expUV}(u,v)$ after a symmetry around the pupil's y-axis. Therefore, if P and Q were left invariant following the same symmetry, our goal would be attained. Moreover, since the Michelson chromatic term must also be made as low as possible, one must use the co-axial chessboards, one way or the other. One solution is to put the co-axial chessboards, shrunken by a factor of two in the upper right corner of the multi-axial chessboards, for example, and to fill the empty cells by means of successive x and y symmetries. This setting

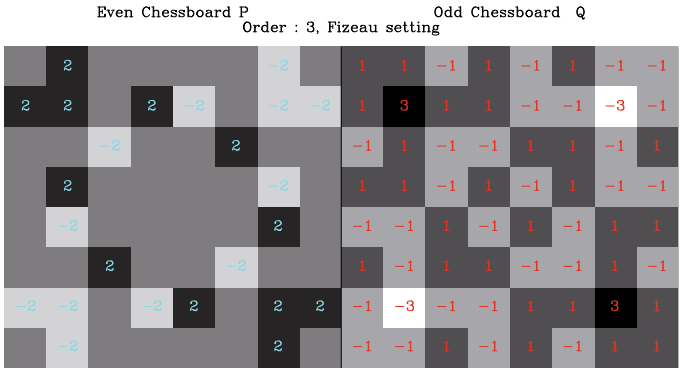

Fig. 4. The (P_3^*, Q_3^*) chessboards. The P_2, Q_2 chessboards from which they were built are in the top right corners. The $2k\pi$ or $(2k+1)\pi$ phase shifts at λ_0 are indicated in the cells' centres by their $2k$ (blue) or $2k+1$ (red) values. The value zero is not shown.

already gives good results, but one can gain more nulling power if one requires that $A(-u,-v) = A(u,v)$.

To achieve this objective, one starts with the co-axial chessboard shrunk by a factor of two, and one again put it in the upper right-hand corner of the multi-axial chessboard under construction. We fill the lower right-hand corner by a x-axis symmetry *with a change of sign* and finally we fill the remaining left-hand side half plane by a y-axis symmetry, again with a change in sign. We symbolically note 'unfold' this procedure. If P_{n-1} and Q_{n-1} stands for two uni-axial chessboards, we construct the corresponding multi-axial chessboards P_n^* and Q_n^* by way of the operations

$$P_n^* = \text{unfold}(P_{n-1}), \quad Q_n^* = \text{unfold}(Q_{n-1}). \tag{33}$$

(P_n^* and Q_n^* have the same number of cells as P_n and Q_n). We have, for example,

$$P_1^* = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad Q_1^* = \begin{bmatrix} +1 & -1 \\ -1 & +1 \end{bmatrix}, \tag{34}$$

$$P_2^* = \begin{bmatrix} -2 & 0 & 0 & +2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ +2 & 0 & 0 & -2 \end{bmatrix}, \quad Q_2^* = \begin{bmatrix} -1 & -1 & +1 & +1 \\ -1 & +1 & -1 & +1 \\ +1 & -1 & +1 & -1 \\ +1 & +1 & -1 & -1 \end{bmatrix}. \tag{35}$$

The "unfold" procedure offers many advantages.

1. Since the combination (P_{n-1}, Q_{n-1}) cancels the chromatic term up to order $N = 2(n-1)$, the (P_n^*, Q_n^*) is expected to cancel it, at least up to the same order $N = 2n-2$. In fact, the change in sign during the unfolding process implements the following property: if $\{P_n\} \overset{2n}{=} \{Q_n\}$, then $\{P_n, -P_n\} \overset{2n+1}{=} \{Q_n, -Q_n\}$. Therefore the pair (P_n^*, Q_n^*) cancels the chromatic term up to order $N = 2n-1$.
2. We expect $P_n^* - Q_n^*$ to be, as $P_{n-1} - Q_{n-1}$, a *fiddop* of, at least, order $n-1$. This is true if n is even, but thanks to the symmetries of the Prouhet-Thue-Morse sequence, this order is increased to n if n is odd.
3. Finally, the symmetry of the multi-axial chessboards is such that it cancels the $[UV]^k$ matrices when k is odd (see Eq. (25)).

Because of these supplementary properties, the chessboards (P_n^*, Q_n^*) designed for the multi-axial case are expected to be just slightly less successful than their co-axial (P_n, Q_n) counterparts.

Table 3 is a summary of the nulling properties of the chessboards considered in this paper.

Table 3. The nulling power of the chessboards considered in this paper.

Type	Size $n_0 \times n_0$	*Chromatic* nulling order N	*Spatial* nulling order (*fiddop*)
co-axial (P_n, Q_n)	$2^n \times 2^n$	$2n$	n
multi-axial (P_n^*, Q_n^*)	$2^n \times 2^n$	$2n-1$	$n-1$, if n even n, if n odd

3. Performance of the design

A first evaluation of the performances of the chessboards have been established in the preceding section; however, it is possible to obtain far more detailed information by working out the analytical expression of the electrical field amplitude. This is made possible because relatively simple operations allow passing, almost mechanically, from one chessboard of order n to the next one. These are, in the following order,

- start with a Bracewell setting;
- shrink the chessboards by a factor of two;
- shift it to its proper place;
- add a phase; and
- perform symmetries with a change in sign.

All these operations, when performed on the pupil plane, have simple Fourier equivalents on the image plane. Therefore, a computation of the amplitude function can be considered. A straightforward, though lengthy, analytical work leads to the formulæ

$$A_n = \left[P_n \mathrm{e}^{-j2\theta_u \frac{D}{d}} + Q_n \mathrm{e}^{j2\theta_u \frac{D}{d}}\right] \mathrm{sinc}\left(\frac{1}{2^{n-1}}\theta_u\right) \mathrm{sinc}\left(\frac{1}{2^{n-1}}\theta_v\right), \tag{36}$$

$$P_n = \frac{1}{2}(S_n + D_n), \quad Q_n = \frac{1}{2}(S_n - D_n), \tag{37}$$

where P_n and Q_n are, in this context, two functions on the image plane. They are *not* the chessboard matrices. The expressions for S_n and D_n differ according to the interferometer setting. In the co-axial case, we set $D = 0$ and obtain

$$S_n = S'_n, \quad D_n = D'_n, \tag{38}$$

while in the multi-axial one, $D \neq 0$ and we find

$$\begin{aligned} S_n &= S''_n(\theta_u, \theta_v, -\phi^*) + S''_n(-\theta_u, -\theta_v, -\phi^*) \\ &\quad + S''_n(-\theta_u, \theta_v, \phi^*) + S''_n(\theta_u, -\theta_v, \phi^*), \end{aligned} \tag{39}$$

$$\begin{aligned} D_n &= D''_n(\theta_u, \theta_v, -\phi^*) + D''_n(-\theta_u, -\theta_v, -\phi^*) \\ &\quad + D''_n(-\theta_u, \theta_v, \phi^*) + D''_n(\theta_u, -\theta_v, \phi^*), \end{aligned} \tag{40}$$

$$S''_n(\theta_u, \theta_v, \phi^*) = \frac{1}{4}\mathrm{e}^{j(\theta_u+\theta_v)} S'_{n-1}\left(\frac{1}{2}\theta_u, \frac{1}{2}\theta_v, \phi^*\right), \tag{41}$$

$$D''_n(\theta_u, \theta_v, \phi^*) = \frac{1}{4}\mathrm{e}^{j(\theta_u+\theta_v)} D'_{n-1}\left(\frac{1}{2}\theta_u, \frac{1}{2}\theta_v, \phi^*\right). \tag{42}$$

The S'_n and D'_n functions reflect the self similarity of the iterative Bracewell scheme and may be called the "nulling functions":

$$\begin{aligned} S'_n(\theta_u, \theta_v, \phi^*) &= \left(1 + \mathrm{e}^{j2\phi^*}\right) \prod_{\ell=0}^{n-1} \cos\left(\phi^* + \tfrac{1}{(-2)^\ell}\theta_u\right) \\ &\quad \times \cos\left(\phi^* + \tfrac{1}{(-2)^\ell}\theta_v\right), \end{aligned} \tag{43}$$

$$\begin{aligned} D'_n(\theta_u, \theta_v, \phi^*) &= \left(1 - \mathrm{e}^{j2\phi^*}\right) \prod_{\ell=0}^{n-1} \sin\left(\phi^* + \tfrac{1}{(-2)^\ell}\theta_u\right) \\ &\quad \times \sin\left(\phi^* + \tfrac{1}{(-2)^\ell}\theta_v\right). \end{aligned} \tag{44}$$

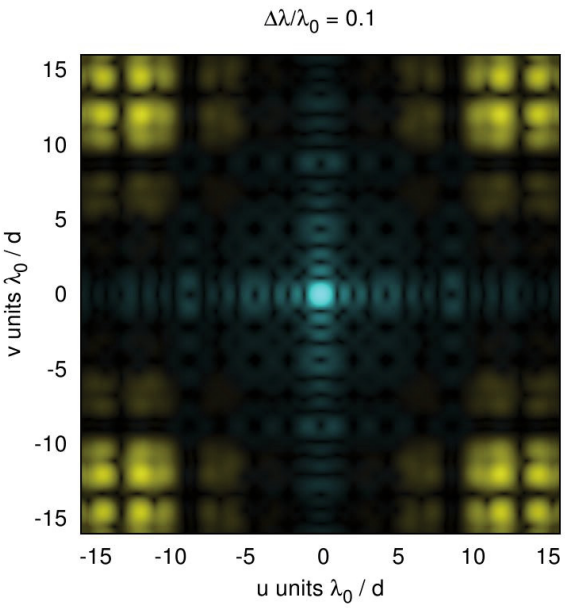

Fig. 5. Image of a star system seen through a Michelson interferometer $d/D = 1$ equipped with the *multi-axial* (P_n^*, Q_n^*) design of order $n = 7$. The system consists of a star and a 10^6 fainter planet. The light from the star is shown in yellow and the one from the planet in blue. The image intensity I has been enhanced by a $I^{0.25}$ filter to see the very low levels (here: the star). The effect of the device is to push away the light from the star, while that of the planet fits in the space left empty. To illustrate the achromatic effect, the observation is simulated for a wavelength detuned by 10% compared to the central wavelength λ_0.

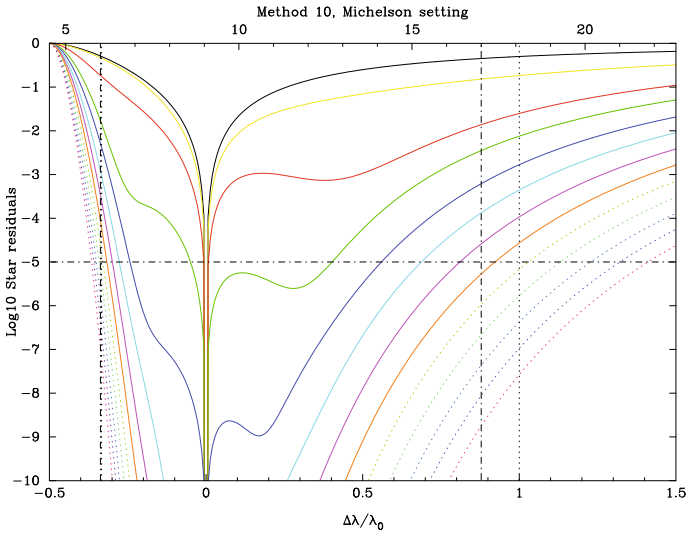

Fig. 6. Extinction spectrum of a star by *iterative Bracewell* nulling uniaxial (Michelson) interferometers of order $n = 0, 1, \ldots, 12$. The original Bracewell $n = 0$ is in black, order 1 in yellow, 2 in red, etc. The x-axis is the wavelength, either $\Delta\lambda/\lambda_0$ (*bottom*) or λ in μm for $\lambda_0 = 9.05\,\mu$m (*top*). The R_n^* integration has been performed over an Airy square of size $2\lambda_0/d$ where d is the size of the mirrors. The maximum bandwidth is indicated between the two dashed vertical lines, and it meets the DARWIN project nominal bandwidth indicated by the two broken vertical lines. The y-axis is the log of the integrated extinction factor, 0 means no extinction, and the horizontal line at -5 sets the nominal DARWIN extinction factor of 10^5.

For a given λ, the θ_u, θ_v are the *spatial* variables and ϕ^* is the *chromatic* variable (because it depends on λ_0). We then obtain

$$\theta_u = \frac{\pi}{2}\frac{d}{\lambda}u\,, \quad \theta_v = \frac{\pi}{2}\frac{d}{\lambda}v, \tag{45}$$

$$\phi^* = \begin{cases} -\phi & \text{if } n \text{ is even,} \\ +\phi & \text{if } n \text{ is odd.} \end{cases} \qquad \phi = \frac{\pi}{2}\frac{\lambda_0}{\lambda}\,. \tag{46}$$

One recalls that: D is the separation between the two telescopes, d the mirror size, λ the wavelength considered, λ_0 the wavelength for which the chessboards are tuned and u, v are the Cartesian angular coordinates of the light diffracted by the pupils. Figure 5 shows the theoretical image of a planetary system similar to ours seen with the (P_7^*, Q_7^*) device. The image was calculated using these formulæ.

To account for a possible path difference between the two arms of the interferometer, we introduce a systematic piston on one arm. If ϵ is the piston value, instead of (36), we have

$$A_n = \left[P_n \mathrm{e}^{-j2\theta_u \frac{D}{d}} + Q_n \mathrm{e}^{j(2\theta_u \frac{D}{d} + 2\phi_\epsilon)}\right] \mathrm{sinc}\left(\frac{1}{2^{n-1}}\theta_u\right) \times \mathrm{sinc}\left(\frac{1}{2^{n-1}}\theta_v\right), \tag{47}$$

$$\text{with } \phi_\epsilon = \pi\frac{\epsilon}{\lambda}\,. \tag{48}$$

3.1. Extinction of the star, attenuation of the planet, and planet-to-starlight ratio

Let us define R_n^* as the residual stellar light intensity. We have $R_n^* = |A_n|^2$. We consider this function successively for the co-axial and multi-axial cases below.

3.1.1. Co-axial case

Residual light from the star. One finds, in the co-axial case

$$R_n^*(u, v, \lambda; \lambda_0, \epsilon) = \left(\cos\phi_\epsilon \cos\phi^* \prod_{\ell=0}^{n-1} \cos\left(\phi^* + \frac{1}{(-2)^\ell}\theta_u\right) \cos\left(\phi^* + \frac{1}{(-2)^\ell}\theta_v\right) - \sin\phi_\epsilon \sin\phi^* \prod_{\ell=0}^{n-1} \sin\left(\phi^* + \frac{1}{(-2)^\ell}\theta_u\right) \sin\left(\phi^* + \frac{1}{(-2)^\ell}\theta_v\right)\right)^2 \times \mathrm{sinc}^2\left(\frac{1}{2^{n-1}}\theta_u\right)\mathrm{sinc}^2\left(\frac{1}{2^{n-1}}\theta_v\right). \tag{49}$$

The starlight extinction spectrum shown in Fig. 6 is defined as the integration $\int_{\mathrm{Airy}} R_n^*(u, v, \lambda; \lambda_0, 0)\,\mathrm{d}u\mathrm{d}v$. Of course, as expected, the nulling effect may reach considerable values over a large bandwidth. For example, an extinction factor of 10^5 can be easily achieved over the bandwidth 6.2–17.5 μm by an interferometer of order $n = 7$ (orange curve) tuned at $\lambda_0 = 9.05\,\mu$m.

To understand more deeply how the nulling works, one considers R_n^* at the centre of the image plane (i.e. $\theta_u = 0$ and $\theta_v = 0$), leading to

$$R_n^*(\lambda; \lambda_0, \epsilon) = \left(\cos\phi_\epsilon \cos^{2n+1}\phi^* - \sin\phi_\epsilon \sin^{2n+1}\phi^*\right)^2. \tag{50}$$

Finally, we suppose a perfect design: $\epsilon = 0$. One ends up with the chessboard's extinction capability characterised by

$$R_n^*(\lambda; \lambda_0) = \cos^{4n+2}\left(\frac{\pi}{2}\frac{\lambda_0}{\lambda}\right). \tag{51}$$

As n goes to infinity, the starlight residual goes to zero at all wavelengths, except those where the cosine is equal to ± 1; that is, $\lambda = \infty, \lambda_0/2, \lambda_0/4, \ldots, \lambda_0/2k, \ldots$

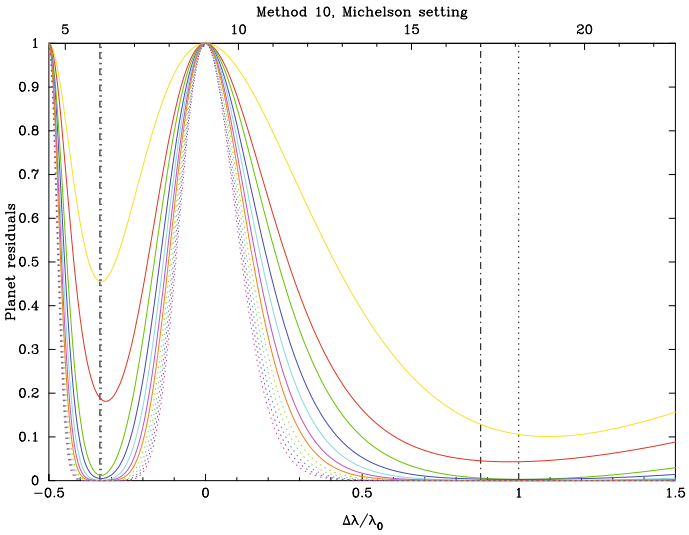

Fig. 7. Attenuation spectrum of a planet following the same setting, except for the linear y-axis, as in Fig. 6. Near the wavelength λ_0, where the device is tuned, the planet only suffers very small attenuation. The full width at half maximum (*FWHM*) is given, with a very good approximation, by $FWHM = 1/\sqrt{2n}$, where n is the interferometer order.

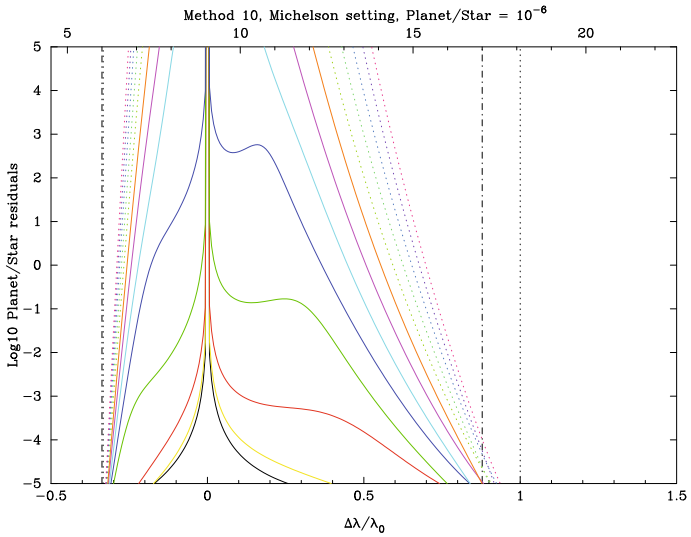

Fig. 8. Planet-to-starlight ratio spectrum of a planet 10^6 times fainter than the star around which it revolves, seen through a uni-axial (Michelson) interferometer. When the value on the y-axis is greater than 0, it is seen more light from the planet than from the star, but at the same time the effective bandwidth of the planet is reduced. A value slightly above 0 seems to be an acceptable objective.

Light from the planet. Since it is off-axis, the planet wavefront is tilted relative to the pupil plane. As in Paper I, we assume that the planet was observed when this tilt induced a phase shift of π at λ_0 between the two arms of the interferometer. If, on the pupil plane, we neglect the effect of the tilt, the residual planet intensity is given by $R_n^{\oplus} = r(\lambda)R_n^*(\phi'_\epsilon)$, where $\phi'_\epsilon = \phi_\epsilon + \pi$. In this expression $r(\lambda)$ stands for the planet/star intensity ratio at λ, typically $r \approx 10^{-6}$ in the thermal infrared. At the image centre, we have

$$R_n^{\oplus}(\lambda;\lambda_0,\epsilon) = r(\lambda)\left(\sin\phi_\epsilon\cos^{2n+1}\phi^* - \cos\phi_\epsilon\sin^{2n+1}\phi^*\right)^2, \quad (52)$$

and, in the absence of piston,

$$R_n^{\oplus}(\lambda;\lambda_0) = r(\lambda)\sin^{4n+2}\left(\frac{\pi}{2}\frac{\lambda_0}{\lambda}\right). \quad (53)$$

As n goes to infinity, the planet light also goes to zero at all wavelengths, except those where the sine is equal to ± 1, that is: $\lambda = \lambda_0, \lambda_0/3, \ldots, \lambda_0/(2k+1), \ldots$ Figure 7 shows the attenuation spectrum of the planet.

Planet-to-starlight ratio. To successfully detect and to do the spectroscopy of an exoplanet, one must reduce all sources of noise. For the lower levels n of our device, the starlight residuals play a predominant role over other sources of noise (such as the exozodiacal light for example). One must therefore consider the ratio $R_n^{\oplus}/R_n^*$, which gives the planet residual intensity relatives to the star residual intensity. This quantity may be called the *planet-to-starlight ratio* and denoted r_n. Typically, if $r_n \geq 1$, the source of noise from the star ceases to be dominant and the planet is likely to be detected. We obtain

$$r_n(\lambda;\lambda_0,\epsilon) = r(\lambda)\left(\frac{\tan\phi_\epsilon \pm \tan^{2n+1}\phi}{1 \pm \tan\phi_\epsilon\tan^{2n+1}\phi}\right)^2, \quad (54)$$

where again $\pm$ means $+$ if n is even and $-$ if it is odd. If $n \to \infty$, and to the extent that $|\tan\phi| > 1$, then $r_n \to r\cot^2\phi_\epsilon$. This expression reduces to $r_n \approx r(\lambda)(\lambda/\pi\epsilon)^2$ for low piston value. For example, if $r = 10^{-6}$, $\lambda = 10\,\mu$m, and $\epsilon = 3$ nm, we get $r_n \approx 1$. That means that a planet 10^6 times fainter than its star is likely to be detected at 10 μm even with a piston of 3nm. If the piston is increased by a factor of 3, then the planet-to-starlight ratio diminishes by approximately a factor of 10. Figure 8 shows the planet-to-starlight ratio for a zero piston.

This discussion is, for the central pixel and a zero piston, in accordance with the previous analysis in 2.1.3. In particular, we have $r/r_n = |\rho_n|^2$.

The planet-to-starlight ratio bandwidth. The condition $|\tan\phi| > 1$ defines the planet-to-starlight ratio bandwidth or, for short, the contrast bandwidth. If $\epsilon = 0$, it defines a *first order* where r_n goes to infinity like $\tan^{4n+2}\phi$ as n (the interferometer order) increases. The contrast bandwidth of the first order is $\pi/4 \leq \phi \leq 3\pi/4$; which is, $\frac{2}{3}\lambda_0 \leq \lambda \leq 2\lambda_0$. For example, if $\lambda_0 = 9\,\mu$m, one gets the bandwidth $6\,\mu\text{m} \leq \lambda \leq 18\,\mu$m, a range consistent with the 6–17 μm one demanded by the DARWIN spatial project.

Higher orders do indeed exists. The kth order is given by $\frac{2}{4k-1}\lambda_0 \leq \lambda \leq \frac{2}{4k-3}\lambda_0$, $k = 1, 2, \ldots$; elsewhere r_n goes to zero. If $\epsilon \neq 0$, the bandwidth is shrunk. One can show that the first order is now roughly given by $\phi_1 \leq \phi \leq \pi - \phi_1$, with $\phi_1 = \arctan(\tan^{-1/(2n-1)}\frac{\pi}{2}\epsilon/\lambda_0)$. (This formula is valid for $\epsilon \neq 0$ and a sufficiently high n, say $n > 5$.).

3.1.2. Multi-axial case

The discussion is about the same as the previous one; however, there is no simple formula like Eq. (49), the more general Eq. (47) must be used instead. If one considers the image centre, one obtains

$$R_n^*(\lambda;\epsilon,\lambda_0) = \cos^2\phi_\epsilon\cos^{4n}\phi + \sin^2\phi_\epsilon\sin^{4n}\phi, \quad (55)$$

$$R_n^{\oplus}(\lambda;\epsilon,\lambda_0) = r(\lambda)\left(\sin^2\phi_\epsilon\cos^{4n}\phi + \cos^2\phi_\epsilon\sin^{4n}\phi\right). \quad (56)$$

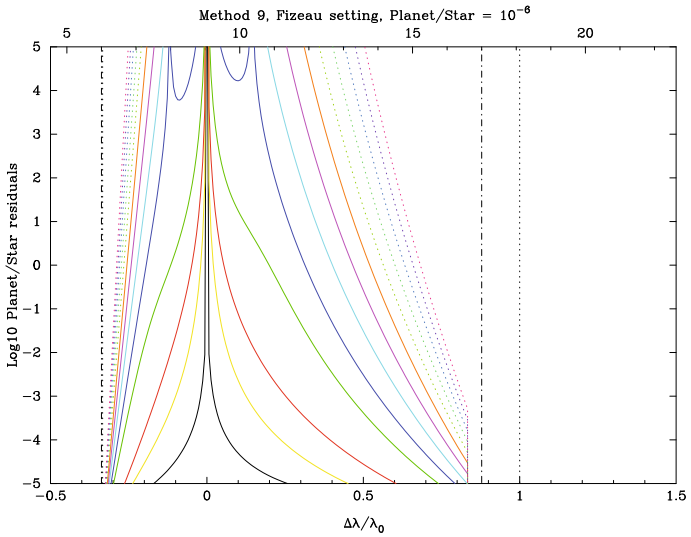

Fig. 9. Planet-to-starlight ratio spectrum. Same as Fig. 8 but for a multi-axial (Fizeau) setting. The integration has been performed by means of a single mode cylindrical fibre optics of radius λ_0/D, where D is separation of the two mirrors. The abrupt cut-off at approximately 16.5 μm is the long wavelength cut of the fibre optics.

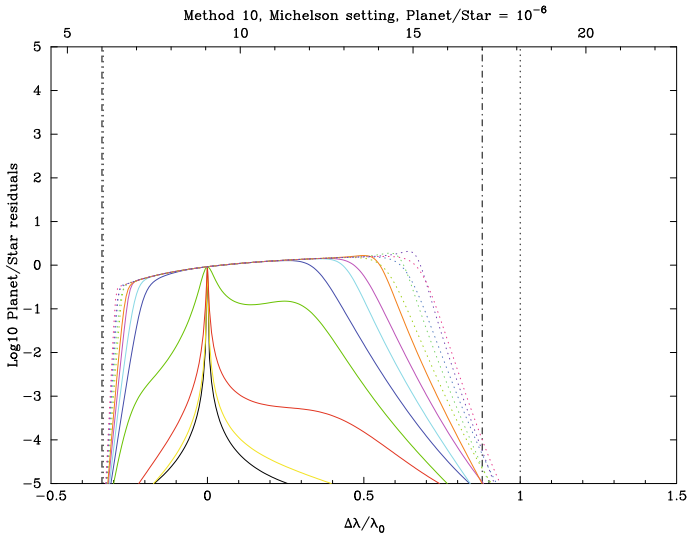

Fig. 10. Planet-to-starlight ratio spectrum. Same as Fig. 8, but suffering a piston of 3 nm or a random unbiased Gaussian piston of $\sigma = 3$ nm, while the device is tuned at $\lambda_0 = 9.05$ μm.

From this one gets the planet-to-starlight ratio

$$r_n(\lambda;\epsilon,\lambda_0) = r(\lambda)\frac{\tan^2\phi_\epsilon + \tan^{4n}\phi}{1+\tan^2\phi_\epsilon\tan^{4n}\phi}\,. \tag{57}$$

Again, if the piston ϵ is zero, the planet-to-starlight ratio goes to infinity like $\tan^{4n}\phi$, as n, the interferometer order, increases. This defines the many spectral orders: $\frac{2}{4k-1}\lambda_0 \le \lambda \le \frac{2}{4k-3}\lambda_0$, and the first one $\frac{2}{3}\lambda_0 \le \lambda \le 2\lambda_0$ is the most extented.

If the piston is not zero, the planet-to-starlight ratio is approximately multiplied by a factor $\cot^2\phi_\epsilon \approx (\lambda/\pi\epsilon)^2$ on approximately the same reduced bandwidth $\phi_1 \le \phi \le \pi - \phi_1$ as in the co-axial case. Figure 9 shows this ratio for a piston of zero.

3.2. Introduction of a random piston

Given the rather tight specifications on the value of the acceptable piston, the new concept seems more suited to a spatial project. However, with the advent of modern fringe-tracking techniques, it is possible to consider a ground-based instrument located on a site of excellent quality, such as in Antarctica. In that case, the rapid atmospheric fluctuations must be taken into account, and the residual differential piston ϵ between the P and Q pupils after correction by a fringe tracker, must be considered as a random variable. In this situation, the residual light function R_n^* must be replaced by its expectation $\mathrm{E}\{R_n^*\}$. To avoid repetition, we focus on the multi-axial case, because the co-axial one is very similar.

To compute $\mathrm{E}\{R_n^*\}$, we write

$$\begin{aligned} R_n^* &= \cos^2\phi_\epsilon\cos^{4n}\phi + \sin^2\phi_\epsilon\sin^{4n}\phi, &(58)\\ &= \frac{1}{2}\left(\cos^{4n}\phi + \sin^{4n}\phi\right) \\ &\quad + \frac{1}{2}\cos 2\phi_\epsilon\left(\cos^{4n}\phi - \sin^{4n}\phi\right), &(59)\end{aligned}$$

$$\begin{aligned} \mathrm{E}\{R_n^*\} &= \frac{1}{2}\left(\cos^{4n}\phi + \sin^{4n}\phi\right) + \frac{1}{2}\mathrm{E}\{\cos 2\phi_\epsilon\} \\ &\quad \times \left(\cos^{4n}\phi - \sin^{4n}\phi\right). &(60)\end{aligned}$$

To evaluate $\mathrm{E}\{\cos 2\phi_\epsilon\}$, we take advantage of the fact that, if X is a random variable, then $\mathrm{E}\{\cos X\}$ is the real part of its *characteristic function* $Z(\omega) = \mathrm{E}\{\mathrm{e}^{j\omega X}\}$ evaluated at $\omega = 1$. Furthermore, if X is a *Gaussian* random variable of mean m and standard deviation s, its characteristic function is $Z(\omega) = \mathrm{e}^{jm\omega - \frac{1}{2}s^2\omega^2}$. By setting $X = 2\pi\epsilon/\lambda$, we obtain $m = 2\pi\mu/\lambda$ and $s = 2\pi\sigma/\lambda$, where μ is the mean and σ the standard deviation of ϵ. It follows that

$$\mathrm{E}\{\cos 2\phi_\epsilon\} = \exp\left(-2\pi^2\sigma^2/\lambda^2\right)\cos(2\pi\mu/\lambda)\,. \tag{61}$$

It is convenient to express this result in terms of the *equivalent piston* $\tilde{\epsilon}$ defined as $\cos 2\phi_{\tilde{\epsilon}} = \mathrm{E}\{\cos 2\phi_\epsilon\}$. An important example is the unbiased case where $\mu = 0$. If σ/λ is small, one obtains $\mathrm{E}\{\cos 2\phi_\epsilon\} = \exp(-2\pi^2\sigma^2/\lambda^2) \approx 1 - 2(\pi\sigma/\lambda)^2$ and since $\cos 2\phi_{\tilde{\epsilon}} \approx 1 - 2(\pi\tilde{\epsilon}/\lambda)^2$, then the equivalent piston is equal to σ.

The main result of this section is that, in the case of Gaussian random fluctuations of the piston around zero, the expected residual light function is equal to its deterministic value for the standard deviation of the piston; that is,

$$\mathrm{E}\{R_n^*(u,v,\lambda;\lambda_0,\epsilon)\} = R_n^*(u,v,\lambda;\lambda_0,\sigma). \tag{62}$$

Figures 10 and 11 show the planet-to-starlight ratio (or its expectation) for a piston of 3 nm.

4. Conclusion

This paper presented a more complete analysis of a very efficient nulling interferometer, which appears as an *iterative* version of the well known Bracewell nulling interferometer. With the help of extended analytical computations of the amplitude expression, we show that the device acts as an optical differential operator on the 3D light distribution, i.e., spatially and chromatically.

We considered two kinds of systems: the (P_n, Q_n) chessboards designed for the Michelson setting and the (P_n^*, Q_n^*) chessboards designed for the Fizeau one. The (P_n, Q_n) cannot be used in the Fizeau case because their PSF are not axisymmetric. The (P_n^*, Q_n^*), however, may be considered for a Michelson interferometer because their performances are just slightly inferior to their Fizeau equivalents and present the advantage of an axisymmetric PSF.

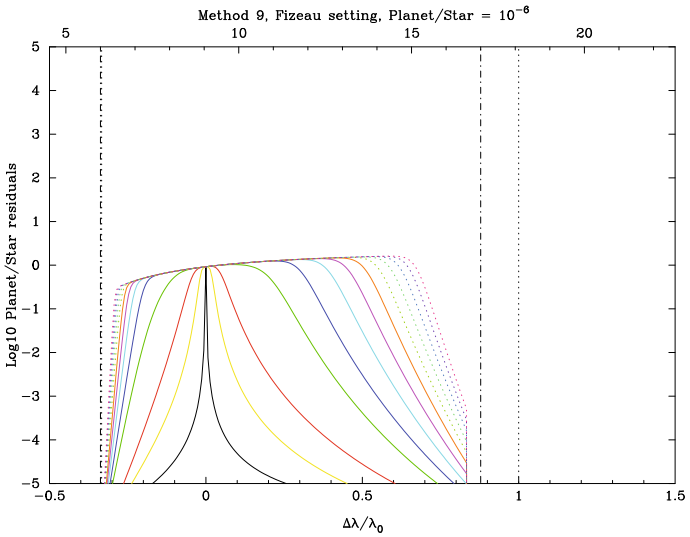

Fig. 11. Planet-to-starlight ratio spectrum. Same as Fig. 10 but for a multi-axial Fizeau setting.

The net effect of the device discussed in this paper is to "push away" the stellar light from the centre of the image, leaving room for an "out of phase" object to appear in the area now left free. This takes place for a very broad range of wavelengths if one considers the planet-to-starlight ratio. A design of order $n = 7$ tuned at $\lambda_0 = 9.05\,\mu$m can meet the DARWIN space project specifications, i.e., to cancel the starlight by a factor of 10^5 over a wavelength range of 6–17 μm. Since the planet light is also diminished, the order n should not be pushed far beyond the point where the planet-to-starlight ratio is equal to one. For example, with $n = 7$, the starlight residuals are reduced below the planet light on most of the DARWIN bandwidth, while one still has more than 50% of the planet light over 20% of the bandwidth. Therefore, the full nominal DARWIN bandwidth can be covered for the planet by adjusting the device successively to five different wavelengths λ_0.

For an effective implementation of this device on the sky, a detailed study of the signal-to-noise ratio should be made prior to decide to what order the interferometer should be constructed. This choice must result from a compromise between the depth of the nulling and the effective planet-light bandwidth.

Imperfections in design have been considered in the form of a deterministic or a random piston between the two pupils of the interferometer. A systematic piston or equivalently a standard deviation of 3 nm seems to lead to acceptable performances in the case of the DARWIN project but also for a ground-based interferometer.

Acknowledgements. We are very grateful to Raphaël Galicher for a careful reading of the manuscript and to the referee whose comments have greatly helped clarify and improve this paper. This work received the support of PHASE, the high angular resolution partnership between ONERA, Observatoire de Paris, CNRS, and University Denis Diderot Paris 7.

Appendix A: Constraints on the number of cells

Here we try to find a lower limit to the number of cells S constituting the chessboards for a given even nulling order N. We recall that the numeric values contained in the P_n and Q_n chessboards are even and odd integers, respectively.

We find that our design is optimal for $N = 2$ and 4, but the question remains open for $N \geq 6$.

Cancelling up to the second nulling order. If n_s and m_s are the P and Q matrices elements, we must have

$$\sum_{s=1}^{S} n_s^2 = \sum_{s=1}^{S} m_s^2 = \Sigma, \tag{A.1}$$

where Σ is the common sum. Since the n_s are even, Σ must be divisible by 4. Because m_s is odd, we have $m_s = 2k_s + 1$ and

$$\sum_{s=1}^{S}(2k_s+1)^2 \equiv 0 \pmod 4, \tag{A.2}$$

$$4\sum_{s=1}^{S} k_s^2 + 4\sum_{s=1}^{S} k_s + S \equiv 0 \pmod 4, \tag{A.3}$$

$$S \equiv 0 \pmod 4. \tag{A.4}$$

Therefore the number of cells in each chessboard must be a non-zero multiple of 4. We have shown by the construction of the P_1 and Q_1 chessboards that $S = 4$ was sufficent, and this is the minimum number of cells required to cancel the chromatic term up to second order.

The above result is a consequence of the property that $(2k + 1)^2 \equiv 1 \pmod 4$. In fact one can show that $(2k + 1)^2 \equiv 1 \pmod 8$ that, in turn, is a particular case with $p = 1$ of the more general result

$$(2k+1)^{2^p} \equiv 1 \pmod{2^{p+2}}, \tag{A.5}$$

which one can easily establish by induction.

Cancelling up to the fourth nulling order. From (A.5) with $p = 2$, we obtain $(2k + 1)^4 \equiv 1 \pmod{16}$. Then we deduce that $S \equiv 0 \pmod{16}$. Again, we show that the 16 cells in P_2 and Q_2 were sufficent to cancel the chromatic term up the fourth order, and it is the minimum.

Cancelling beyond the fourth nulling order. One can use (A.5) to sets limits on the sixth order cancelled by P_3 and Q_3, but they are too loose. (Unless one can prove that, if 2^{2n} is a divisor of $\sum_{s=1}^{S}(2k_s + 1)^{2n}$ then, for $n > 1$, it is also a divisor of $\sum_{s=1}^{S}(2k_s + 1)^{2^{2n-2}}$). All that can be said is that the number of cells must be a multiple of 16, while P_3 and Q_3 possesses 64. Up to now, as soon as the nulling order $N \geq 5$, no tighter limits than those deduced from (A.5) are known to the authors regarding the number of cells necessary to cancel the chromatic term up to that order.

References

Allouche, J.-P., & Mendès-France, M. 2008, Monatshefte für Mathematik, 155, 301
Bracewell, R. N. 1978, Nature, 274, 780
Buisset, C., Rejeaunier, X., Rabbia, Y., et al. 2006, in Advances in Stellar Interferometry, ed. J. D. Monnier, M. Schöller, & W. C. Danchi, SPIE Conf. Proc., 6268, 626819
Prouhet, E. 1851, Comptes Rendus des Séances de l'Académie des Sciences, 33, 225
Ratcliffe, J. G. 1994, Foundations of Hyperbolic Manifold, 1st edn. (New York: Springer-Verlag)
Rouan, D., & Pelat, D. 2008, A&A, 484, 581
Wallner, O., Perdigues Armengol, J. M., & Karlsson, A. L. 2004, in New Frontiers in Stellar Interferometry, ed. W. A. Traub. (Bellingham, WA: The International Society for Optical Engineering), SPIE Conf. Proc., 5491, 798