Chebyshev approximation of the null function by an affine combination of complex exponential functions

Paul ARMAND†, Joël BENOIST† and Elsa BOUSQUET†

May 18, 2010

Abstract. We describe the theoretical solution of an approximation problem that uses a finite weighted sum of complex exponential functions. The problem arises in an optimization model for the design of a telescopes array occurring within optical interferometry for direct imaging in astronomy. The problem is to find the optimal weights and the optimal positions of a regularly spaced array of aligned telescopes, so that the resulting interference function approximates the zero function on a given interval. The solution is given by means of Chebyshev polynomials.

Key words. Chebyshev approximation, Chebyshev polynomials, complex exponential functions, interferometry, optimization

1 Introduction

This paper describes the solution of an approximation problem arising in the application of mathematical optimization for the design of a telescopes array that occurs within optical interferometry. Indeed, telescopes arrays will provide an efficient way for a direct imaging of exoplanets in a near future. The hypertelescope concept proposed by Labeyrie [2] (see also Reynaud and Delage [8], Vakili et al. [9], Patru et al. [7]), is an interferometric optical instrument that uses a specific conditioning of the beams light coming from the telescopes. Because the ratio of light intensities of a star and of an exoplanet can be more than one million or even one billion, it is absolutely necessary to design an instrument with a very high sensitivity, a property also called dynamic. In a recent paper, Armand et al. [1] have proposed an optimization model to compute an optimal positioning of telescopes and optimal amplitudes of light beams, to maximize the dynamic of a linear array of telescopes of same diameter. In the present paper, we consider the problem from a theoretical point of view.

Let us consider $N+1$ aligned telescopes of same diameter, those positions are identified by a vector $x = (x_0, \dots, x_N) \in \mathbb{R}^{N+1}$ and let $a = (a_0, \dots, a_N) \in \mathbb{R}^{N+1}$ be the vector of beams amplitudes through the telescopes. In the optical application, the light intensity of an image point is the product of the interference function,

†Laboratoire XLIM – Université de Limoges (France); email: paul.armand@unilim.fr, joel.benoist@unilim.fr and elsa.bousquet@etu.unilim.fr.

that is $t \mapsto |\sum_{n=0}^{N} a_n e^{ix_n t}|^2$ with another function called the diffraction envelope. See the book of Labeyrie et al. [3] for an introduction to optical interferometry. In the particular case of a temporal version of a hypertelescope [8], the maximization problem of the dynamic naturally leads to the following optimization problem:

$$\inf_{(a,x)\in\mathcal{A}\times\mathbb{R}^{N+1}} \sup_{t\in[\alpha,\beta]} |\sum_{n=0}^{N} a_n e^{ix_n t}|$$

where $\mathcal{A} = \{a \in \mathbb{R}^{N+1} : \sum_{n=0}^{N} a_n = 1\}$ is the affine hyperplane of $\mathbb{R}^{N+1}$ generated by the N-dimensional simplex and $[\alpha, \beta]$ is the interval in which we want that the light intensity is lowered as much as possible. This interval defines a dark zone where the main lobe of a secondary tiny source of light, e.g. an exoplanet, could be detected.

From a mathematical point of view, the above optimization problem is a challenging question. In particular, the existence of an optimal solution remains, to our knowledge, an open question. However, numerical simulations allow us to conjecture that at optimality, the components of the vector x should be periodic in the sense that the distance $x_n - x_{n-1}$ is constant for all $n = 1, \ldots, N$. For these both reasons, the present study is limited to the periodic case. Taking into account that for all $x_0 \in \mathbb{R}$ and $h \in \mathbb{R}$ we have

$$|\sum_{n=0}^{N} a_n e^{i(x_0+nh)t}| = |\sum_{n=0}^{N} a_n e^{inht}|,$$

we then formulate the periodic positioning problem under the general form

$$\inf_{(a,h)\in\mathcal{A}\times\mathbb{R}_+} \sup_{t\in[\alpha,\beta]} |\sum_{n=0}^{N} a_n e^{inht}|.$$

For a fixed value of the vector of positions, the objective function of the above minimization problems are convex functions of the variable $a \in \mathbb{R}^{N+1}$. This kind of problem appears in antenna array pattern synthesis and can be efficiently solved via convex optimization algorithms, see e.g. [5]. But when the positions are also variables of the optimization model, the problem is more difficult and, to our knowledge, did not received relevant solution. The objective of this paper is to solve explicitly the periodic positioning problem, that is to exhibit a pair $(a^*, h^*) \in \mathcal{A} \times \mathbb{R}_+$ that realizes the minimum of this problem.

The remainder of this paper is organized as follows. In the next section we describe the problem and state the main result. Section 3 introduces some linear isomorphisms that will be useful to reformulate the main problem into various forms. The proof of our theorem is fully developed in Section 4. In Section 5 we give a property of the optimal weights $a_0^*, \ldots, a_n^*$ and also an illustrative example of the approximation theorem. Note that the mathematical model that we propose is quite general, so that it can be applied to other domains of signal processing. In the last section we give a potential application to the optimization of the beam pattern of a linear array of antennas.

2 Notation, problem statements and main result

Let N denote a given positive integer and $M := \lfloor N/2 \rfloor$, i.e. $N = 2M$ if N is even and $N = 2M + 1$ if N is odd. Define the segment $I := [\alpha, \beta]$ for given constants $0 < \alpha < \beta$ and for any positive number h we define $I(h) := [\alpha h, \beta h]$. Let $\mathcal{A} := \{a \in \mathbb{R}^{N+1} : \sum_{n=0}^{N} a_n = 1\}$ be the affine hyperplane generated by the N-simplex of $\mathbb{R}^{N+1}$. Let $f : J \to \mathbb{C}$ be a continuous function defined on an interval $J \subset \mathbb{R}$. The infinity norm of f on J is defined by

$$\|f\|_J := \sup_{t \in J} |f(t)|.$$

The exponential function $e_h : \mathbb{R} \to \mathbb{C}$ parametrized by $h \in \mathbb{R}$ is defined by

$$e_h(t) := e^{iht}.$$

Let us define the weighted sum of exponential functions for $a \in \mathbb{R}^{N+1}$ and $h \in \mathbb{R}$ by

$$\sigma(a, h) := \sum_{n=0}^{N} a_n e_{nh}. \tag{2.1}$$

Our main problem can then be formulated as solving the infimum problem

$$\inf_{(a,h) \in \mathcal{A} \times \mathbb{R}_+} \|\sigma(a, h)\|_I. \tag{2.2}$$

Taking at first the infimum with respect to the variable a, we can rewrite the problem under the form

$$\inf_{h \in \mathbb{R}_+} \varphi(h), \tag{2.3}$$

where the function $\varphi : \mathbb{R} \to \mathbb{R}$ is defined by

$$\varphi(h) := \inf_{a \in \mathcal{A}} \|\sigma(a, h)\|_I. \tag{2.4}$$

Note that we have $\varphi(0) = 1$ and that $\varphi(h) \leq 1$ for all $h \in \mathbb{R}_+$.

An example of graph of a function φ is depicted on Figure 1. The values have been obtained by numerical solutions of a sequence of problems (2.4). The representation shows that the function φ is nonconvex and that the problem (2.3) seems to be already difficult to solve from a theoretical point of view.

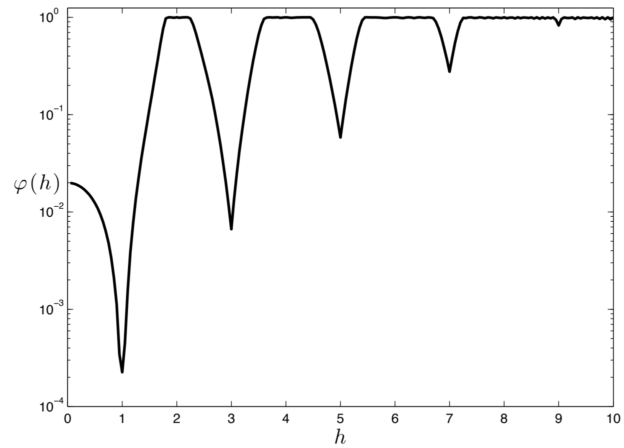

Figure 1: Semi-log graph of the function φ defined by (2.4) for the particular values $N = 4$, $[\alpha, \beta] = [\frac{9\pi}{10}, \frac{11\pi}{10}]$. The graph has been obtained with 200 discretization points of the h axis

Before stating our main result, we recall some well known facts about Chebyshev polynomials. Chebyshev polynomials are defined by the recurrence relation

$$T_{n+1}(x) + T_{n-1}(x) = 2xT_n(x) \quad \text{for all } n \geq 1, \tag{2.5}$$

with $T_0(x) = 1$ and $T_1(x) = x$ as initial conditions. We refer to [6] for the remarkable properties of this sequence of polynomials. In particular for all integer n the polynomial T_n has n simple roots in the open interval $]-1, 1[$. Let $\mathcal{P}_n$ denote the set of polynomials of degree at most n. The recurrence relation (2.5) allows to easily compute the coefficients of T_N in the canonical basis of $\mathcal{P}_N$. Since T_N is of the same parity as N, it can be written in the form (see e.g., [6, (2.16)])

$$T_N(x) = \sum_{m=0}^{M} t_m x^{N-2m}. \tag{2.6}$$

Let us define the polynomial of degree N

$$T_N^*(x) := \frac{1}{T_N(\frac{1}{s})} \sum_{m=0}^{M} t_m x^m \left(\frac{1+x}{2s}\right)^{N-2m}, \tag{2.7}$$

where

$$s := \cos(\frac{\alpha\pi}{\alpha+\beta}).$$

We can now state our main result.

Theorem 2.1 *The infimum in (2.3) is achieved at*

$$h^* := \frac{2\pi}{\alpha + \beta}$$

and we have the equalities

$$\varphi(h^*) = \|\sigma(a^*, h^*)\|_I = \frac{1}{T_N(\frac{1}{s})},$$

where $a^* = (a_0^*, \dots, a_N^*)$ *is the vector of coefficients of* T_N^**, that is*

$$T_N^*(x) = \sum_{n=0}^{N} a_n^* x^n.$$

3 Preliminaries

We denote $\mathbb{E} = \mathcal{P}_N$ the vector space of polynomials of degree at most N. For any $P \in \mathbb{E}$, we will use the notation $P = \sum_{n=0}^{N} p_n x^n$. The subspaces of polynomials of $\mathbb{E}$ which are even and odd functions of x are denoted by

$$\mathbb{F} := \{P \in \mathbb{E} : P(x) = P(-x)\} \quad \text{and} \quad \mathbb{F}' := \{P \in \mathbb{E} : P(x) = -P(-x)\}.$$

We have $\mathbb{E} = \mathbb{F} \oplus \mathbb{F}'$ and let $p_{\mathbb{F}} : \mathbb{E} \to \mathbb{F}$ be the projection along $\mathbb{F}'$ onto $\mathbb{F}$, that is

$$p_{\mathbb{F}}(\sum_{n=0}^{N} p_n x^n) = \sum_{m=0}^{M} p_{2m} x^{2m}.$$

We will also use the vector space

$$\mathbb{G} := \{P \in \mathbb{E} : P(x) = (-1)^N P(-x)\},$$

so that $\mathbb{G} = \mathbb{F}$ if N is even and $\mathbb{G} = \mathbb{F}'$ otherwise. Note that $T_n \in \mathbb{G}$ for all $n \in \mathbb{N}$.

Let us now define three linear isomorphisms. The first one is

$$\begin{array}{rccc} j: & \mathcal{P}_M & \to & \mathbb{G} \\ & P & \mapsto & x^N P(\frac{1}{x^2}) \end{array}$$

or equivalently defined by

$$j(\sum_{m=0}^{M} p_m x^m) = \sum_{m=0}^{M} p_m x^{N-2m}. \tag{3.1}$$

The second one is

$$\begin{array}{rccc} f: & \mathbb{E} & \to & \mathbb{E} \\ & A = \sum_{n=0}^{N} a_n x^n & \mapsto & B = \sum_{n=0}^{N} b_n x^n \end{array}$$

defined by

$$f(A) = (1-x)^N A\left(\frac{1+x}{1-x}\right), \tag{3.2}$$

or equivalently by

$$B = \sum_{n=0}^{N} a_n (1+x)^n (1-x)^{N-n}.$$

From (3.2), the inverse isomorphism satisfies

$$f^{-1}(B) = \left(\frac{x+1}{2}\right)^N B\left(\frac{x-1}{x+1}\right), \tag{3.3}$$

or equivalently

$$A = \sum_{n=0}^{N} b_n \left(\frac{x+1}{2}\right)^{N-n} \left(\frac{x-1}{2}\right)^n. \tag{3.4}$$

At last, the third isomorphism is

$$\begin{array}{cccc} g: & \mathbb{F} & \to & \mathbb{G} \\ & C = \sum_{m=0}^{M} c_m x^{2m} & \mapsto & D = \sum_{m=0}^{M} d_m x^{N-2m} \end{array}$$

defined by

$$g(C) = x^N C\left(\frac{\sqrt{x^2-1}}{x}\right), \tag{3.5}$$

with the convention that $\sqrt{-1} = i$, or equivalently defined by

$$D = \sum_{m=0}^{M} c_m (x^2-1)^m x^{N-2m}. \tag{3.6}$$

Making the substitution $Y = \frac{1}{x^2}$ in this last equality, we have

$$\sum_{m=0}^{M} d_m Y^m = \sum_{m=0}^{M} c_m (1-Y)^m,$$

from which we deduce that the inverse isomorphism of g satisfies

$$g^{-1}(D) = j^{-1}(D)(1-x^2). \tag{3.7}$$

Finally, we can define the composite linear maps

$$\ell = g \circ p_{\mathbb{F}} \circ f \quad \text{and} \quad \ell' = f^{-1} \circ g^{-1}.$$

We then have

$$\ell \circ \ell' = id_{\mathbb{G}}, \tag{3.8}$$

the identity map of $\mathbb{G}$.

The following Lemma will be useful in the sequel.

Lemma 3.1 *Define* $\mathbb{E}_1 = \{A \in \mathbb{E} : A(1) = 1\}$ *and* $\mathbb{G}_1 = \{D \in \mathbb{G} : D(1) = 1\}$.

(*i*) *For all* $A \in \mathbb{E}$, *we have* $A(1) = \ell(A)(1)$.

(*ii*) *We have the inclusion* $\ell'(\mathbb{G}_1) \subset \mathbb{E}_1$ *and the equality* $\ell(\mathbb{E}_1) = \mathbb{G}_1$ *.*

Proof. (*i*) Let $A \in \mathbb{E}$. From definition (3.2) the polynomial $B := f(A)$ satisfies

$$B(0) = A(1).$$

Now from definition of $p_{\mathbb{F}}$ the polynomial $C := p_{\mathbb{F}}(B)$ satisfies

$$C(0) = B(0).$$

Finally, from definition (3.5) the polynomial $D := g(C) = \ell(A)$ satisfies

$$D(1) = C(0).$$

Combining these three equalities we conclude $A(1) = \ell(A)(1)$.

(*ii*) Let us prove the first inclusion. Let $A \in \ell'(\mathbb{G}_1)$. There exists $D \in \mathbb{G}_1$ such that $A = \ell(D)$. Using successively assertion (*i*) and (3.8) we have the chain of equalities

$$A(1) = \ell(A)(1) = (\ell \circ \ell')(D)(1) = D(1) = 1$$

which implies that $A \in \mathbb{E}_1$.

Let us now prove the equality. Composing the first inclusion by the linear map ℓ we obtain $\mathbb{G}_1 \subset \ell(\mathbb{E}_1)$. The reverse inclusion is a straightforward consequence of assertion (*i*). □

4 Proof of the main result

Before giving a proof of Theorem 2.1, we propose to establish three Lemmas.

In a first lemma, we give a new formulation of the function φ defined by (2.4) and of the periodic positioning problem (2.2)

Lemma 4.1 *For all* $h \geq 0$

$$\varphi(h) = \inf_{A \in \mathbb{E}_1} \|A \circ e_1\|_{I(h)} \tag{4.1}$$

and problem (2.2) can be formulated as

$$\inf_{(A,h) \in \mathbb{E}_1 \times \mathbb{R}_+} \|A \circ e_1\|_{I(h)}. \tag{4.2}$$

Proof. Consider the change of variables $\tau : \mathcal{A} \to \mathbb{E}_1$ defined for all $a \in \mathcal{A}$ by

$$\tau(a) := \sum_{n=0}^{N} a_n x^n.$$

Then, for all $(a, h) \in \mathcal{A} \times \mathbb{R}_+$, defining $A := \tau(a)$ we have the chain of equalities

$$\begin{aligned} \|\sigma(a,h)\|_I &= \sup_{t\in I} |\sum_{n=0}^{N} a_n\, e^{inht}| \\ &= \sup_{t\in I} |A(e^{iht})| \\ &= \sup_{\theta\in I(h)} |A(e^{i\theta})| \\ &= \|A \circ e_1\|_{I(h)}. \end{aligned}$$

According to (2.4), we obtain (4.1) and thanks to this change of variables, our general problem (2.2) can be reformulated as (4.2). □

In the next lemma we propose again a new expression of the function φ.

Lemma 4.2 *Let $h \in]0, \frac{2\pi}{\beta}]$ and consider the interval included in $]-1, 1[$*

$$J(h) := [r(h), s(h)],$$

where $r(h) := \cos(\frac{\beta h}{2}) < s(h) := \cos(\frac{\alpha h}{2})$.

- *(i) For all $A \in \mathbb{E}$, we have $\|A \circ e_1\|_{I(h)} \geq \|\ell(A)\|_{J(h)}$, with equality whenever $A \in f^{-1}(\mathbb{F})$.*
- *(ii) We have*

$$\varphi(h) = \inf_{D\in\mathbb{G}_1} \|D\|_{J(h)}. \tag{4.3}$$

Moreover, if this infimum is achieved for some $D^* \in \mathbb{G}_1$, then

$$\varphi(h) = \|\ell'(D^*) \circ e_1\|_{I(h)}.$$

Proof. (i) Let us prove that for all $(\theta, A) \in \mathbb{R} \times \mathbb{E}$, we have

$$|A(e^{i\theta})| \geq |\ell(A)(\cos\frac{\theta}{2})|, \tag{4.4}$$

with equality whenever $A \in f^{-1}(\mathbb{F})$. Let $A = \sum_{n=0}^{N} a_n x^n \in \mathbb{E}$, $\theta \in \mathbb{R}$ and set $\theta' = \frac{\theta}{2}$. Define

$$B := f(A) = \sum_{n=0}^{N} b_n x^n, \quad C := p_{\mathbb{F}}(B) = \sum_{m=0}^{M} b_{2m} x^{2m} \quad \text{and} \quad D := g(C).$$

We then have $D = \ell(A)$. According to (3.4), we have

$$A(e^{i\theta}) = \sum_{n=0}^{N} b_n \left(\frac{e^{i\theta}+1}{2}\right)^{N-n} \left(\frac{e^{i\theta}-1}{2}\right)^{n}$$
$$= e^{iN\theta'} \sum_{n=0}^{N} i^n b_n (\cos\theta')^{N-n} (\sin\theta')^n.$$

By taking modulus on both sides and knowing that $|z| \geq |\Re(z)|$ for any complex number z, we deduce that

$$|A(e^{i\theta})| \geq |\sum_{m=0}^{M} (-1)^m b_{2m} (\cos\theta')^{N-2m} (\sin\theta')^{2m}|, \tag{4.5}$$

with equality whenever $B \in \mathbb{F}$ or equivalently when $A \in f^{-1}(\mathbb{F})$. According to definition (3.6), we also have

$$D(\cos\theta') = \sum_{m=0}^{M} b_{2m} (\cos^2\theta' - 1)^m (\cos\theta')^{N-2m}. \tag{4.6}$$

From (4.5) and (4.6) we conclude that

$$|A(e^{i\theta})| \geq |D(\cos\theta')|$$

with equality whenever $A \in f^{-1}(\mathbb{F})$, which proves (4.4). It suffices to take the supremum on both sides of the inequality (4.4) to conclude the proof of the first assertion.

(*ii*) According to the equality (4.1) and assertion (*i*), we have

$$\varphi(h) \geq \inf_{A \in \mathbb{E}_1} \|\ell(A)\|_{J(h)}.$$

By virtue of Lemma 3.1-(*ii*), we then deduce that

$$\varphi(h) \geq \inf_{D \in \mathbb{G}_1} \|D\|_{J(h)}. \tag{4.7}$$

Let us prove now the reverse inequality. Let $D_1 \in \mathbb{G}_1$ and set $A_1 := \ell'(D_1)$. On the one hand, by using the definition ℓ' we have $A_1 = f^{-1}(g^{-1}(D_1)) \in f^{-1}(\mathbb{F})$ and by using Lemma 3.1-(*ii*) we have

$$A_1 \in \mathbb{E}_1. \tag{4.8}$$

On the other hand, according to (3.8) we have

$$\ell(A_1) = D_1. \tag{4.9}$$

By using the assertion (i) just proved, property (4.8) and equality (4.9), we deduce that

$$\|D_1\|_{J(h)} = \|\ell(A_1)\|_{J(h)} = \|A_1 \circ e_1\|_{I(h)} \geq \inf_{A \in \mathbb{E}_1} \|A \circ e_1\|_{I(h)} = \varphi(h).$$

By taking the infimum over $D_1 \in \mathbb{G}_1$, we obtain the reverse inequality to (4.7) and thus the equality holds.

At last, assume that there exists $D^* \in \mathbb{G}_1$ such that $\varphi(h) = \|D^*\|_{J(h)}$. We obviously have $\varphi(h) = \|A^* \circ e_1\|_{I(h)}$, where $A^* := \ell'(D^*)$. □

In the third lemma we propose a relaxation of the definition (4.3) of φ which will be useful to establish the achieved value of φ at minimum. This lemma follows from a standard result in approximation theory, but we give a detailed proof for completeness.

Lemma 4.3 *Let $h \in]0, \frac{2\pi}{\beta}]$ and consider the following relaxed form of problem (4.3)*

$$\varphi_1(h) := \inf_{D \in \mathbb{E}_1} \|D\|_{J(h)}. \tag{4.10}$$

Then the polynomial

$$D_h := \frac{T_N\left(\frac{2x - s(h) - r(h)}{s(h) - r(h)}\right)}{T_N\left(\frac{2 - s(h) - r(h)}{s(h) - r(h)}\right)}$$

of $\mathbb{E}_1$ satisfies the following equalities

$$\varphi_1(h) = \|D_h\|_{J(h)} = \frac{1}{T_N\left(\frac{2 - s(h) - r(h)}{s(h) - r(h)}\right)}.$$

Proof. Let us define $\mathbb{E}_0 = \{Q \in \mathbb{E} : Q(1) = 0\}$. Making the change of variable. $D = x^N - Q$ in the relaxed problem (4.10), we get

$$\varphi_1(h) = \inf_{Q \in \mathbb{E}_0} \|x^N - Q\|_{J(h)}$$

which represents, in the vector space $\mathbb{E}$, the distance with respect to the infinity norm of the polynomial $x^N \in \mathbb{E}$ to the hyperplane $\mathbb{E}_0$.

We shall apply the Chebyshev alternation Theorem (see e.g., [4, Theorem 3.5.2]). At first, let us remark that any nonzero $Q \in \mathbb{E}$ satisfying $Q(1) = 0$ has at most $N-1$ roots in the interval $J(h)$, which means that $\mathbb{E}_0$ satisfies the Haar condition (see [4, Definition 3.4.1]). Consider now the affine transformation

$$\Gamma : J(h) = [r(h), s(h)] \to [-1, 1]$$

defined by $\Gamma(x) = \frac{2x-s(h)-r(h)}{s(h)-r(h)}$, for all $x \in J(h)$, so that $\Gamma(r(h)) = -1$ and $\Gamma(s(h)) = 1$.

On one hand, the polynomial T_N equioscillates $N+1$ times in $[-1,1]$ and

$$\|T_N\|_{[-1,1]} = 1.$$

It follows that $T_N \circ \Gamma$ equioscillates $N+1$ times in $J(h)$ and that

$$\|T_N \circ \Gamma\|_{J(h)} = 1.$$

On the other hand, recalling that $J(h) \subset]-1,1[$ and that Γ is nondecreasing, we have $\Gamma(1) = \frac{2-s(h)-r(h)}{s(h)-r(h)} > 1$ which implies that $T_N \circ \Gamma(1) > 0$ since all the roots of T_N are in $]-1,1[$ and that $T_N(1) = 1$. Therefore the polynomial D_h which can be rewritten

$$D_h = \frac{T_N \circ \Gamma}{T_N \circ \Gamma(1)}$$

equioscillates $N+1$ times in $J(h)$ and

$$||D_h||_{J(h)} = \frac{1}{T_N\left(\frac{2-s(h)-r(h)}{s(h)-r(h)}\right)}.$$

Now, applying the Chebyshev alternation Theorem, we conclude that the polynomial $Q_h = x^N - D_h$ which belongs to $\mathbb{E}_0$ is the unique projection of x^N onto $\mathbb{E}_0$ and that

$$\varphi_1(h) = ||D_h||_{J(h)},$$

which completes the proof. □

We are now is position to complete the proof of the main result of this paper.

Proof of Theorem 2.1. Let us recall that $h^* := \frac{2\pi}{\alpha+\beta}$ and that $s := \cos\left(\frac{\alpha\pi}{\alpha+\beta}\right)$.

Let us first prove that

$$\forall h \in]0, h^*] \quad \varphi(h^*) = \varphi_1(h^*) \leq \varphi_1(h) \leq \varphi(h). \tag{4.11}$$

The rightmost inequality is due to (4.3), (4.10) and to the inclusion $\mathbb{G}_1 \subset \mathbb{E}_1$. Recalling the definitions of $s(h)$ and $r(h)$ in Lemma 4.2, we have $s(h^*) = -r(h^*) = s$. It follows that

$$\frac{2x - s(h^*) - r(h^*)}{s(h^*) - r(h^*)} = \frac{x}{s} \tag{4.12}$$

and thus the polynomial D_h defined in Lemma 4.3 satisfies

$$D_{h^*} = \frac{T_N\left(\frac{x}{s}\right)}{T_N\left(\frac{1}{s}\right)} \tag{4.13}$$

By Lemma 4.3, the fact that $D_{h^*} \in \mathbb{G}_1$ and (4.3), we have

$$\begin{aligned}
\varphi_1(h^*) &= \|D_{h^*}\|_{J(h^*)} \\
&\geq \inf_{D \in \mathbb{G}_1} \|D\|_{J(h^*)} \\
&= \varphi(h^*) \\
&\geq \varphi_1(h^*).
\end{aligned}$$

All the inequalities are thus equalities and so $\varphi(h^*) = \varphi_1(h^*)$. To conclude the proof of (4.11), it remains to show that $\varphi_1(h^*) \leq \varphi_1(h)$ for all $h \in]0, h^*]$. Let $h \in]0, h^*[$. Define

$$c := \frac{\beta}{\alpha} > 1 \quad \text{and} \quad k := \frac{\alpha h}{4} \in]0, \frac{\pi}{2(1+c)}[$$

The first derivative of the function $t \mapsto \frac{\sin ct}{\sin t}$ is negative on $]0, \frac{\pi}{2c}[$ if and only if $c \tan t < \tan ct$ for all $t \in]0, \frac{\pi}{2c}[$. This last inequality is a consequence of the strict convexity of the tangent function on $]0, \frac{\pi}{2}[$. It follows that the previous ratio of sinus functions is decreasing on $]0, \frac{\pi}{2(1+c)}[$. We then deduce the following inequalities,

$$\frac{\sin ck}{\sin k} \geq \frac{\sin \frac{\pi c}{2(1+c)}}{\sin \frac{\pi}{2(1+c)}} = \cot \frac{\pi}{2(1+c)} \geq 1$$

and

$$1 + \frac{2}{(\frac{\sin ck}{\sin k})^2 - 1} \leq 1 + \frac{2}{\cot^2 \frac{\pi}{2(1+c)}} = \frac{1}{\cos \frac{\pi}{1+c}}.$$

We then obtain

$$\begin{aligned}
\frac{2 - s(h) - r(h)}{s(h) - r(h)} &= \frac{(1 - \cos \frac{\beta h}{2}) + (1 - \cos \frac{\alpha h}{2})}{(1 - \cos \frac{\beta h}{2}) - (1 - \cos \frac{\alpha h}{2})} \\
&= \frac{\sin^2 \frac{\beta h}{4} + \sin^2 \frac{\alpha h}{4}}{\sin^2 \frac{\beta h}{4} - \sin^2 \frac{\alpha h}{4}} \\
&= 1 + \frac{2}{(\frac{\sin ck}{\sin k})^2 - 1} \\
&\leq \frac{1}{\cos \frac{\pi}{1+c}} \\
&= \frac{1}{\cos \frac{\pi \alpha}{\alpha + \beta}} \\
&= \frac{1}{s}
\end{aligned}$$

All the roots of T_N are contained in $]-1, 1[$. Rolle Theorem implies that the same property holds for the derivative of T_N. Since $T_N(1) = 1$, the polynomial T_N is then increasing on $[1,+\infty[$. Therefore, the last inequality implies that

$$T_N\left(\frac{2-s(h)-r(h)}{s(h)-r(h)}\right) \le T_N\left(\frac{1}{s}\right).$$

By virtue of Lemma 4.3 and (4.12), this last inequality reads $\varphi_1(h^*) \le \varphi_1(h)$, which completes the proof of (4.11).

Let us prove now that

$$\forall h \in [h^*,+\infty[\quad \varphi(h^*) \le \varphi(h). \tag{4.14}$$

Let $h > h^*$. We shall distinguish two cases.

First case: $\beta h < 2\pi$. Let $h_1 := \frac{2\pi-\beta h}{\alpha}$. Since $h > h^* = \frac{2\pi}{\alpha+\beta}$, we clearly have

$$h_1 \in]0,h^*[\quad \text{and} \quad I(h_1) \subset [2\pi-\beta h, 2\pi-\alpha h].$$

For all $A \in \mathbb{E}_1$, we deduce that

$$\begin{aligned} \|A \circ e_1\|_{I(h)} &= \|A \circ e_1\|_{[2\pi-\beta h,2\pi-\alpha h]} \\ &\ge \|A \circ e_1\|_{I(h_1)}. \end{aligned}$$

By taking the infimum over all $A \in \mathbb{E}_1$ and by using the property (4.1) of φ, then using (4.11) we deduce that $\varphi(h) \ge \varphi(h_1) \ge \varphi(h^*)$.

Second case: $\beta h \ge 2\pi$. Let $n := \left\lfloor \frac{\beta h}{2\pi} \right\rfloor$, so that $2n\pi \le \beta h < 2n\pi + 2\pi$. If $\alpha h \le 2n\pi$, then for all $A \in \mathbb{E}_1$ we have $\|A \circ e_1\|_{I(h)} \ge 1$, which implies the desired conclusion $\varphi(h) \ge 1 = \varphi(0) \ge \varphi(h^*)$. So, we can now assume that $\alpha h > 2n\pi$, meaning that $[\alpha h, \beta h] \subset]2n\pi, 2(n+1)\pi[$. Let $h_2 := h - \frac{2n\pi}{\beta}$. We clearly have

$$\beta h_2 \in]0,2\pi[\quad \text{and} \quad I(h_2) \subset [\alpha h - 2n\pi, \beta h - 2n\pi].$$

For all $A \in \mathbb{E}_1$, we deduce that

$$\begin{aligned} \|A \circ e_1\|_{I(h)} &= \|A \circ e_1\|_{[\alpha h-2n\pi,\beta h-2n\pi]} \\ &\ge \|A \circ e_1\|_{I(h_2)}. \end{aligned}$$

By taking again the infimum over all $A \in \mathbb{E}_1$ and by using the result just proved in the first case, we deduce that $\varphi(h) \ge \varphi(h_2) \ge \varphi(h^*)$ and completes the proof of (4.14).

Properties (4.11) and (4.14), and the fact that $\varphi(0) = 1 \ge \varphi(h^*)$, imply that the infimum of (2.3) is achieved at h^*, which proves the first assertion of Theorem 2.1.

In summary, we proved that

$$\inf_{h\in[0,+\infty[} \varphi(h) = \varphi(h^*) = \|D_{h^*}\|_{J(h^*)} = \frac{1}{T_N(\frac{1}{s})},$$

where $D_{h^*} \in \mathbb{G}_1$ is defined by (4.13). To complete the proof of Theorem 2.1, it remains to show that

$$\|D_{h^*}\|_{J(h^*)} = \|T_N^* \circ e_1\|_{I(h^*)}, \tag{4.15}$$

where T_N^* is defined by (2.7).

Since $D_{h^*} \in \mathbb{G}_1$, by virtue of Lemma 4.2 we have

$$\|D_{h^*}\|_{J(h^*)} = \|\ell'(D_{h^*}) \circ e_1\|_{I(h^*)}.$$

It then suffices to show that $\ell'(D_{h^*}) = T_N^*$ to get (4.15). From (4.13) and the expression (2.6) of the Chebyshev polynomial T_N, we have

$$D_{h^*}(x) = \frac{1}{T_N(\frac{1}{s})} \sum_{m=0}^{M} \frac{t_m}{s^{N-2m}} x^{N-2m}.$$

According to (3.1) we deduce that

$$j^{-1}(D_{h^*})(x) = \frac{1}{T_N(\frac{1}{s})} \sum_{m=0}^{M} \frac{t_m}{s^{N-2m}} x^{m}.$$

From expression (3.7) of g^{-1} we obtain

$$g^{-1}(D_{h^*})(x) = \frac{1}{T_N(\frac{1}{s})} \sum_{m=0}^{M} \frac{t_m}{s^{N-2m}} (1-x^2)^{m}.$$

Finally, recalling that $\ell' = f^{-1} \circ g^{-1}$ and using the expression (3.3) of f^{-1} we conclude that

$$\begin{aligned} \ell'(D_{h^*})(x) &= \frac{(\frac{x+1}{2})^N}{T_N(\frac{1}{s})} \sum_{m=0}^{M} \frac{t_m}{s^{N-2m}} \left(1 - \left(\frac{x-1}{x+1}\right)^2\right)^m \\ &= T_N^*(x), \end{aligned}$$

which ends the proof. □

5 Optimal weights and illustrative example

In this section we illustrate Theorem 2.1 by means of an example, but before we give a simple property of the optimal coefficients and an explicit formula for the function $\sigma(a^*, h^*)$ defined by (2.1).

Proposition 5.1 *The coefficients of the polynomial $T_N^*(x) = \sum_{n=0}^{N} a_n^* x^n$ defined by (2.7) are symmetric and positive, that is*

$$a_n^* = a_{N-n}^* \quad \text{and} \quad a_n^* \geq 0$$

for all $n = 0, \ldots, N$.

Proof. Let us first note that T_N^* defined by (2.7) can be written under the form

$$T_N^*(x) := \frac{1}{T_N(\frac{1}{s})} \sqrt{x}^N T_N \left(\frac{1}{2s} \left(\sqrt{x} + \frac{1}{\sqrt{x}} \right) \right). \tag{5.1}$$

From formula (5.1) we have $x^N T_N^*(\frac{1}{x}) = T_N^*(x)$, which implies that the coefficients of T_N^* are symmetric.

Let us prove now that these coefficients are positive. Let us define the polynomial

$$Q_N = T_N(\tfrac{1}{s})\, T_N^*.$$

According to (2.5) and (5.1), the sequence of polynomials $\{Q_n\}_{n\in\mathbb{N}}$ satisfies the recurrence relation

$$Q_{n+1}(x) = \frac{1}{s}(x+1)Q_n - xQ_{n-1} \quad \text{for all } n \geq 1, \tag{5.2}$$

with $Q_0(x) = 1$ and $Q_1(x) = \frac{1}{2s}(1+x)$ as initial conditions. In particular, we have $Q_2(x) = \frac{1}{2s^2}(1 + 2(1-s^2)x + x^2)$.

We will prove that for all integer $k = 0, \ldots, n$ the kth derivative of Q_n at 0, denoted by $Q_n^{(k)}(0)$, is positive. To do that, let us prove by induction on $n \in \mathbb{N}$ that the following property holds:

$$\forall k \in \{0, \ldots, n-2\}\ Q_n^{(k)}(0) \geq \frac{1}{s} Q_{n-1}^{(k)}(0) \quad \text{and} \quad \forall k \in \{0, \ldots, n\}\ Q_n^{(k)}(0) \geq 0. \tag{5.3}$$

From above it is true for $n = 0$, $n = 1$ and $n = 2$. Assume now that (5.3) holds for all integer $m \leq n$ for a given integer $n \geq 2$. Let us prove that it also holds for $n+1$.

At first, from (5.2) and the induction hypothesis we trivially have

$$Q_{n+1}^{(0)}(0) = \frac{1}{s} Q_n^{(0)}(0) \geq 0.$$

Now, let $k \in \{1, \ldots, n-1\}$. Using the recurrence relation (5.2), we see that the kth derivative of Q_{n+1} is of the form

$$Q_{n+1}^{(k)} = \frac{1}{s} Q_n^{(k)} + \frac{k}{s}(Q_n^{(k-1)} - s\, Q_{n-1}^{(k-1)}) + \frac{x}{s}(Q_n^{(k)} - s\, Q_{n-1}^{(k)}).$$

We then deduce that

$$\begin{aligned}
Q_{n+1}^{(k)}(0) &= \frac{1}{s} Q_n^{(k)}(0) + \frac{k}{s}(Q_n^{(k-1)}(0) - s\, Q_{n-1}^{(k-1)}(0)) \\
&\geq \frac{1}{s} Q_n^{(k)}(0) + \frac{k}{s}(\frac{1}{s} Q_{n-1}^{(k-1)}(0) - s\, Q_{n-1}^{(k-1)}(0)) \\
&= \frac{1}{s} Q_n^{(k)}(0) + \frac{k(1-s^2)}{s^2} Q_{n-1}^{(k-1)}(0) \\
&\geq \frac{1}{s} Q_n^{(k)}(0) \\
&\geq 0,
\end{aligned}$$

where each inequality follows from the induction hypothesis. At last, since the coefficients of Q_{n+1} are symmetric and since we have proved that $Q^{(1)}_{n+1}(0) \geq 0$ and $Q^{(0)}_{n+1}(0) \geq 0$, we also have $Q^{(n)}_{n+1}(0) \geq 0$ and $Q^{(n+1)}_{n+1}(0) \geq 0$, which completes the proof. □

Proposition 5.2 *The function $\sigma(a^*, h^*)$ defined by (2.1) satisfies for all $t \in \mathbb{R}$,*

$$\sigma(a^*, h^*)(2t) = e^{iNh^*t} \, \frac{T_N\left(\frac{1}{s}\cos(h^*t)\right)}{T_N(\frac{1}{s})}.$$

Proof. From (2.1), Theorem 2.1 and (5.1) we have for all $t \in \mathbb{R}$

$$\begin{aligned} \sigma(a^*, h^*)(2t) &= \sum_{n=0}^{N} a_n^* e^{2inh^*t} \\ &= T_N^*(e^{2ih^*t}) \\ &= \frac{1}{T_N(\frac{1}{s})} e^{iNh^*t} \, T_N\left(\frac{1}{s}\cos(h^*t)\right). \end{aligned}$$

□

We illustrate the approximation theorem with the following simple example. Let us consider the interval $[\alpha, \beta] = [1, 2]$. We then have

$$h^* = \frac{2\pi}{3} \quad \text{and} \quad s = \frac{1}{2}.$$

The function $\sigma(a^*, h^*)$, see definition (2.1), is a periodic function with period $\alpha + \beta = 3$. The exact value of the optimal weights are reported in Table 1 for $N = 1, \ldots, 9$. Figure 2 shows the corresponding graphs of the function $|\sigma(a^*, h^*)|$. Thanks to Proposition 5.2 we have

$$|\sigma(a^*, h^*)(t)| = \frac{T_N\left(2\cos(\frac{\pi}{3}t)\right)}{T_N(2)},$$

for all $t \in \mathbb{R}$.

6 Application to linear antenna array design

As we said in the introduction, the mathematical problem that we have considered comes from the design of a linear array of telescopes [1]. Our initial goal was to find the optimal position and contribution of each telescope to design an intrument with

N	$(a_0^* \cdots a_M^*)$	$\varphi(h^*)$
1	$\begin{pmatrix}1\end{pmatrix}/2$	$1/2$
2	$\begin{pmatrix}2 & 3\end{pmatrix}/7$	$1/7$
3	$\begin{pmatrix}4 & 9\end{pmatrix}/26$	$1/26$
4	$\begin{pmatrix}8 & 24 & 33\end{pmatrix}/97$	$1/97$
5	$\begin{pmatrix}16 & 60 & 105\end{pmatrix}/362$	$1/362$
6	$\begin{pmatrix}32 & 144 & 306 & 387\end{pmatrix}/1351$	$1/1351$
7	$\begin{pmatrix}64 & 336 & 840 & 1281\end{pmatrix}/5042$	$1/5042$
8	$\begin{pmatrix}128 & 768 & 2208 & 3936 & 4737\end{pmatrix}/18817$	$1/18817$
9	$\begin{pmatrix}256 & 1728 & 5616 & 11448 & 16065\end{pmatrix}/70226$	$1/70226$

Table 1: Exact values of the optimal weights for $[\alpha, \beta] = [1, 2]$ and infimum values of problem (2.3).

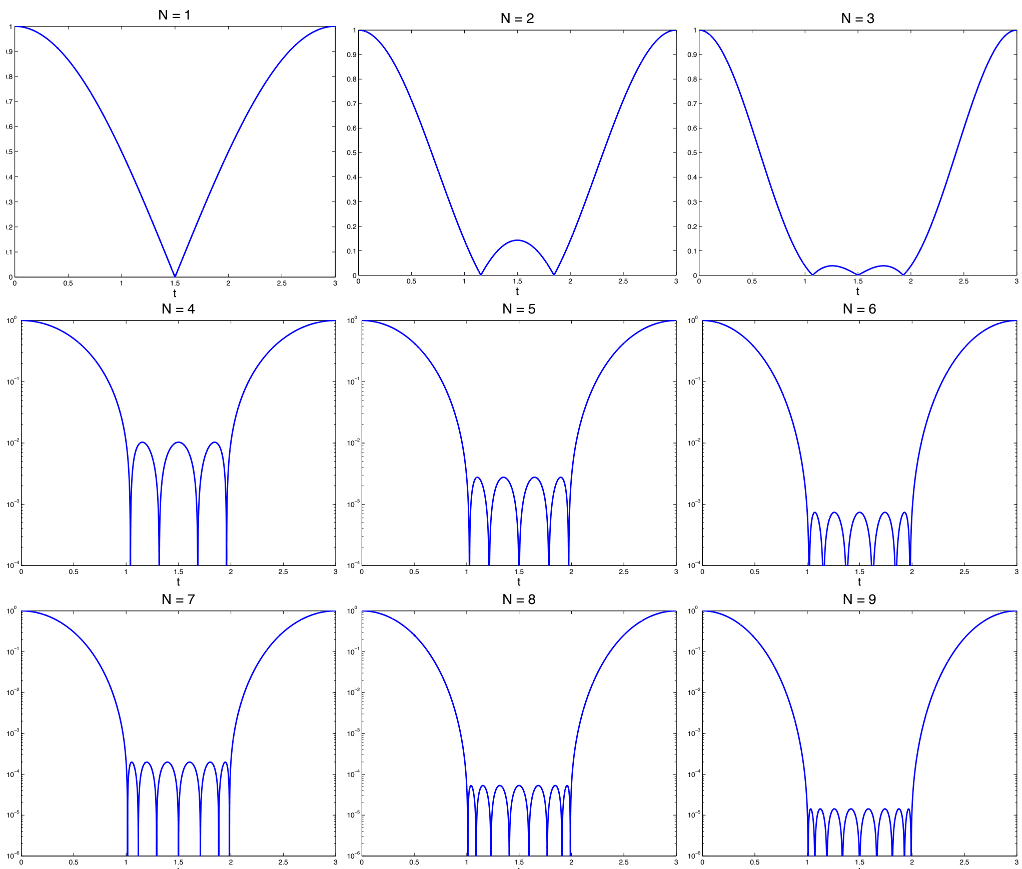

Figure 2: Graphs of $|\sigma(a^*, h^*)|$ defined by (2.1) for $[\alpha, \beta] = [1, 2]$. The three first graphs are plotted with a linear scale, the others with a logarithmic scale of the vertical axis.

a high dynamic, hence the idea of minimizing the norm of the function σ (2.1) on a given interval. In this section we would like to show how our approximation result could be also used as a beamforming technique for linear antenna arrays.

Consider an uniformly spaced one-dimensional array of $N+1$ isotropic antennas positioned along an axis at locations $x_n = nd$, $n = 0, \ldots, N$ where $d > 0$ is the distance between two elements. Assume that the antennas are supplied with the same phase and the amplitude values $a_n \in \mathbb{R}$, $n = 0, \ldots, N$. The well-known array factor or array beampattern is the function

$$A(\theta) = \sum_{n=0}^{N} a_n e^{\frac{2i\pi}{\lambda} nd \cos\theta}$$

where θ is the azimuth angle and λ the wavelength.

Several beamforming techniques allow to choose weights in some optimal sense, see for example [10]. Let us show that our theoretical result can be applied to a new beamforming method which could be used in antenna design. Let $0 \leq \theta_0 < \theta_1 < \frac{\pi}{2}$ be two azimuth angles. Consider the optimization problem

$$\begin{array}{ll} \text{minimize}_{(a,d)\in\mathbb{R}^{N+1}\times\mathbb{R}_+} & \max\{|A(\theta)| : \theta \in [\theta_0, \theta_1]\} \\ \text{subject to} & \sum_{n=0}^{N} a_n = 1. \end{array}$$

The goal is to find both spacing and weighting of the antennas, so that the maximum sidelobes level is as lower as possible in a predetermined region. Thanks to a change of variables $t = \cos\theta$ and $h = \frac{2\pi}{\lambda}d$, Theorem 2.1 directly applies to find the optimal spacing, that is

$$d^* = \frac{\lambda}{\cos\theta_0 + \cos\theta_1}$$

and the optimal weigths a_n^*, $n = 0, \ldots, N$.

Figure 3 shows the beam pattern, that is the polar plot of the function $|A(\cdot)|^2$, of an array with nine elements and for different choices of azimuth angles θ_0 and θ_1. Table 2 shows the corresponding optimal distance and attenuation (in dB) of secondary lobes in the angular region defined by θ_0 and θ_1.

θ_0	θ_1	d^*/λ	Δ
0°	60°	2/3	85
30°	60°	$\sqrt{3}-1$	101
75°	85°	2.89	56

Table 2: Optimal distance and attenuation $\Delta := \max\{10\log_{10}|A(\theta)|^2 : \theta \in [\theta_0, \theta_1]\}$

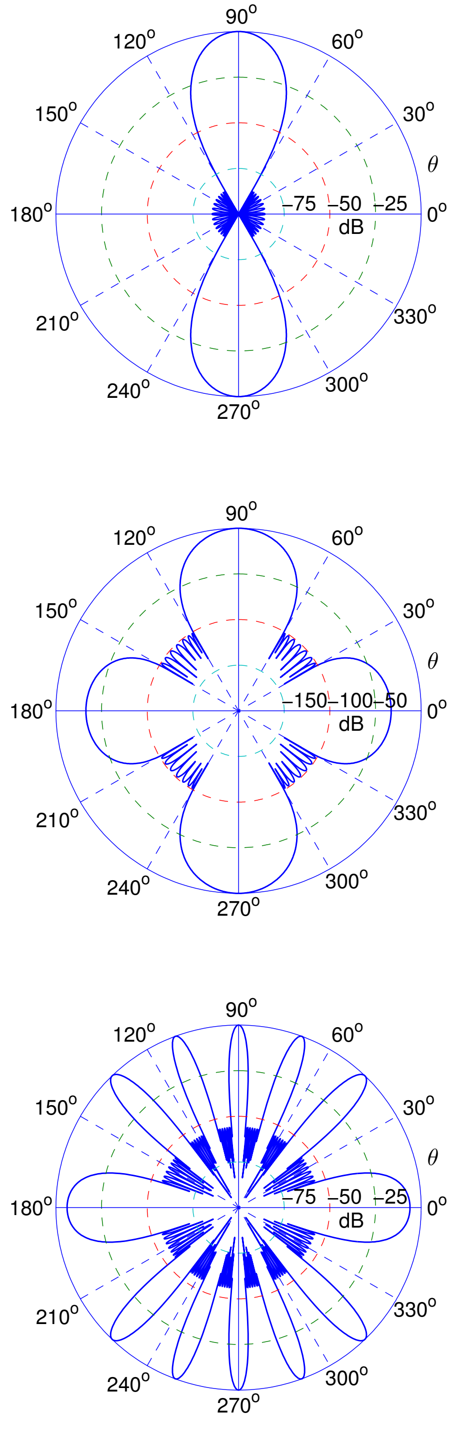

Figure 3: Azimuthal gain patterns of a 9-elements isotropic array

References

[1] P. Armand, J. Benoist, E. Bousquet, L. Delage, S. Olivier, and F. Reynaud, *Optimization of a one dimensional hypertelescope for a direct imaging instrument in astronomy*, European Journal of Operational Research, 195 (2009), pp. 519–527.

[2] A. Labeyrie, *Resolved imaging of extra-solar planets with future 10-100 km optical interferometric arrays*, Astronomy & Astrophysics, 118 (1996), pp. 517–524.

[3] A. Labeyrie, S. G. Lipson, and P. Nisenson, *An Introduction to Optical Stellar Interferometry*, Cambridge University Press, 2006.

[4] P.-J. Laurent, *Approximation et optimisation*, Hermann, Paris, 1972. Collection Enseignement des Sciences, No. 13.

[5] H. Lebret and S. Boyd, *Antenna array pattern synthesis via convex optimization*, IEEE Transaction on Signal Processing, 45 (1997), pp. 526–532.

[6] J. C. Mason and D. C. Handscomb, *Chebyshev polynomials*, Chapman & Hall/CRC, Boca Raton, FL, 2003.

[7] F. Patru, N. Tarmoul, D. Mourard, and O. lardière, *Direct imaging with highly diluted apertures – ii. properties of the point spread function of a hypertelescope*, Monthly Notices of the Royal Astronomical Society, 395 (2009), pp. 2363–2372.

[8] F. Reynaud and L. Delage, *Proposal for a temporal version of a hypertelescope*, Astronomy & Astrophysics, 465 (2007), pp. 1093–1097.

[9] F. Vakili, E. Aristidi, L. Abe, and B. Lopez, *Interferometric remapped array nulling*, Astronomy & Astrophysics, 465 (2007), pp. 1093–1097.

[10] H. L. Van Trees, *Optimum Array Processing. Part IV of Detection, Estimation, and Modulation Theory*, John Wiley and Sons, Inc., New York, 2002.