Revista de la
Unión Matemática Argentina
Vol. 35 1990.

ON THE MAXIMUM ENTROPY SOLUTION OF THE CARATHEODORY-FEJER PROBLEM

RODRIGO AROCENA

A la memoria de Julio Rey Pastor, pionero de la matemática en el Río de la Plata

Abstract. An approach to the Carathéodory-Féjer problem based on the method of unitary extensions of an isometry, which gives a geometric interpretation of Schur parameters, was presented in a previous paper. In this note we complement that approach by showing that innovation and entropy type considerations lead to select a particular solution of that problem, as it happens in the trigonometric moment problem.

Innovation in each step

We want to show that some relations between maximum entropy and the moment problem, as discussed by Landau [L], also appear naturally in the context of the Carathéodory-Féjer problem when it is solved by the method sketched in [A-1]. So we shall use the notation and results of the last paper without stating them again.

Let $\{c_0, c_1, ..., c_n\} \subset C$ be such that $F_n = \{f \in H^\infty(T): \|f\|_\infty \leq 1 .\ \hat{f}(k) = c_k, 0 \leq k \leq n\}$ has more than one element. (See [A -1], theorem A). Set $K_n = \{\hat{f}(n+1) : f \in F_n\}$. A classical statement says that K_n is a closed disk with radius equal to $\prod\{(1-|\gamma_k|^2): 0 \leq k \leq n\}$, where $\gamma_0, \gamma_1, ..., \gamma_n$ are the Schur parameters corresponding to $c_0, c_1, ..., c_n$ [Sch]. We start by giving a proof of that result. More precisely:

Proposition A

$K_n = \{z \in C: \|z - a_n\| \leq \rho(n)\}$, *with* $\rho(n) = \prod\{(1-|\gamma_k|^2). 0 \leq k \leq n\}$, $a_n =$

$= \langle VP_{D(n)} V^n d_1, d_2 \rangle_{H(n)}$, *and* $P_{D(n)}$ *the orthogonal projection of* $H(n)$ *onto* $D(n)$.

Proof

It $f \in F_n$ there exists $(U,G) \in \mathbf{U}$ such that $\hat{f}(n+1) = \langle U^{n+1} d_1, d_2 \rangle_G =$

$$\langle VP_{D(n)} V^n d_1, d_2 \rangle_{H(n)} + \langle UP_{N(n)} V^n d_1, d_2 \rangle_G = a_n + \langle UP_{N(n)} V^n d_1, P_{M(n)} d_2 \rangle_G$$

Now, $\rho(n) = \text{dist}^2 [V^n d_1, D(n)] = \text{dist}^2 [d_2, R(n)]$, so $\rho(n) = \| P_{N(n)} V^n d_1 \|^2 =$

$= \|P_{M(n)} d_2\|^2$; thus, $\| \hat{f}(n+1) - a_n \| \leq \rho(n)$.

It remains to see that, if $|z| \leq \rho(n)$, there exists $(U,G) \in \mathbf{U}$ such that $\langle UP_{N(n)}V^n d_1, d_2 \rangle_G = z$. Let N'be the span of a unit vector v, set $H' = H(n) \oplus N'$ and call V' the isometry acting in H' that extends V to H(n) and verifies $V'(P_{N(n)}V^n d_1) = [z/\rho(n)]P_{M(n)}d_2 + \{[\rho(n)^2 - |z|^2] / \rho(n)\}^{1/2} v$. Take $(U,G) \in \mathbf{U}$ such that $U|_{H(n)} = V'$. Then $\langle UP_{N(n)}V^n d_1, d_2 \rangle_G =$
$= \langle V'P_{N(n)}V^n d_1, d_2 \rangle_G = [z / \rho(n)] \langle P_{M(n)} d_2, d_2 \rangle_{H(n)} = z$

Q.E.D.

As we know, the Fourier coefficients of each $f \in F_n$ are obtained in a step by step extension of $V^n d_1$. In the step that gives $c_{n+1} = \hat{f}(n+1) = \langle U^{n+1} d_1, d_2 \rangle_G$ we put $U^{n+1} d_1 = \mu + \beta$, with $\mu = VP_{D(n)} V^n d_1 + [z/\rho(n)] P_{M(n)} d_2$ belonging to $H(n)$ and $\beta = \{[\rho(n)^2 - |z|^2]/\rho(n)\}^{1/2} v$ orthogonal to that space. So we may say that c_{n+1} is obtained by using in μ the information we already had (i.e., the one given by $c_0, c_1, \ldots, c_n$) and innovating in β.

Now, $[\rho(n)^2 - |c_{n+1} - a_n|^2]/\rho(n) = \|\beta\| = \mathrm{dist}^2 [U^{n+1} d_1, H(n)] = \rho(n)(1 - |\gamma_{n+1}|^2)$ $= \rho(n+1)$

Since $\rho(n)$ is determined by $\{c_0, c_1, \ldots, c_n\}$, we may think of $(1 - |\gamma_{n+1}|^2$ as giving a measure of the innovation in step $(n+1)$ of the construction of the sequence $\{c_k\}$ of Fourier coefficients of f.

From $|c_{n+1} - a_n|^2 = \rho(n)|\gamma_{n+1}|^2$ and $\rho(n+1) = [\rho(n)^2 - |c_{n+1} - a_n|^2]/\rho(n)$ we see that the closer we choose c_{n+1} to the centre a_n of the disk K_n, the bigger the innovation will be and the larger the disk K_{n+1} where we shall have to choose c_{n+2} in the next step. In particular, when innovation in step (n+1) equals zero, f is completely determined by $\{c_0, c_1, \ldots, c_{n+1}\}$ while maximamum innovation corresponds to $\gamma_{n+1} = 0$.

On the entropy integral

In this context the following result, due to Boyd [B], is remarkable:

1 .Proposition

Let a sequence $\{c_k\} \subset \mathbf{C}$ *be such that there exists an* $f \in H^\infty(\mathbf{T})$ *that satisfies* $\|f\|_\infty \leq 1$ and $\hat{f}(k) = c_k$ *for every* k, *with corresponding Schur parameters* $\{\gamma_k\}$. *Then*

$$\lim_{n\to\infty} \Pi \{(1 - |\gamma_k|^2) : 0 \leq k \leq n\} = \exp\{(1/2\pi) \int \log(1 - |f|^2)\, dt\}$$

This property is closely related to the fundamental results we now recall.

2. Theorem. *Let* $d\eta = w\,dt + d\eta_s$ *be a finite positive measure on* $\mathbf{T}$ *such that* $d\eta_s$ *is singular to* dt *and* D_n *the determinant of the Toeplitz matrix* $(\hat{\eta}(j-k))_{0 \leq j,k \leq n}$. *Set* $e_j(t) = e^{ijt}$. *Then:*

a) *the distance in* $L^2(\eta)$ *of* e_0 *to the span of* $\{e_j : j \geq 1\}$ *equals*

$$\exp\{(1/2\pi) \int \log w\, dt\};$$

b) $\lim_{n\to\infty} (D_n/D_{n-1}) = \lim_{n\to\infty} D_n^{(1/n+1)} =$

$$= \exp\{(1/2\pi) \int \log w\, dt\}.$$

Property (2.a) is Szegö - Kolmogorov- Krein theorem (See[G-S], p.44 or [G],p.144, for example) and the proof (1) is based on it. Property (2.b) is Szegö's limit theorem ([G-S], p.65). The following extension of Boyd's proposition can be seen as translation of theorem (2) from the context of the trigonometric moment problem to the one of the Carathéodory-Féjer problem.

If $\alpha = (\alpha_{jk})_{j,k=12}$ is a positive matrix of measures on $\mathbf{T}$ call $L^2(\alpha)$ the Hilbert space generated by the linear span of $\{(e_j, e_k), j,k \in \mathbf{Z}\}$ and the scalar product given by $\langle (e_j,e_k), (e_{j'},e_{k'}) \rangle =$ $\hat{\alpha}_{11}(j-j') + \hat{\alpha}_{12}(j-k') + \hat{\alpha}_{21}(k-j') + \hat{\alpha}_{22}(k-k')$.

3. Propositon B *For* $f \in H^\infty(\mathbf{T})$ *such that* $\|f\|_\infty \leq 1$ *let* $\{\gamma_k\}$ *be the sequence of its Schur parameters,* Γ_n *the Toeplitz matrix* $(\hat{f}(j-k))_{0 \leq j,k \leq n}$ *and* $\alpha = (\alpha_{jk})_{j,k=12}$ *given by* $\alpha_{11} = \alpha_{22} = 1$, $\alpha_{12} = \bar{\alpha}_{21} = f$. *Then the distance in* $L^2(\alpha)$ *of* $(0, e_0)$ *to the span of*

$\{(e_j, e_k): j \geq 0, k > 0\}$ *is equal to*

$$\lim_{n\to\infty} [\det(1-\Gamma_n^* \Gamma_n)/\det(1-\Gamma_{n-1}^* \Gamma_{n-1})] =$$

$$= \lim_{n\to\infty} \Pi\{(1-|\gamma_k|^2): 0 \leq k \leq n\} = \exp\{(1/2\pi) \int \log(1-|f|^2)\, dt\}.$$

Proof

With $\hat{f}(k) = c_k$ set $H = V\{H(n): n \geq 0\}$. Thus, with obvious notation, H is generated by $\{V^j d_1, V^k d_2: j, k \in \mathbf{Z}\}$ and $L(V^j d_1 + V^k d_2) = (e_j, e_k)$ defines a unitary isomorphism of H onto $L^2(\alpha)$ that takes d_2 to $(0, e_0)$ and $V\{V^j d_1, V^k d_2: j \geq 0, k > 0\}$ onto $V\{e_j, e_k): j \geq 0, k > 0\}$. We know [A-1] that $\rho(n) = \det(1-\Gamma_n^* \Gamma_n)/\det(1-\Gamma_{n-1}^* \Gamma_{n-1}) = \Pi\{(1-|\gamma_k|^2): 0 \leq k \leq n\}$ is equal to the distance in H(n) from d_2 to the span of $\{V^j d_1, V^k d_2: n \geq j \geq 0, n \geq k > 0\}$. Since $H(n) \subset H(n+1)$ for every n, $\lim_{n\to\infty} \{(\det(1-\Gamma_n^* \Gamma_n)/\det(1-\Gamma_{n-1}^* \Gamma_{n-1})\} = \lim_{n\to\infty} \Pi\{(1-|\gamma_k|^2): 0 \leq k \leq n\}$ is the distance in H from d_2 to $V\{V^j d_1, V^k d_2: j \geq 0, k \geq 0\}$, ie, the distance in $L^2(\alpha)$ from $(0, e_0)$ to $V(e_j, e_k): j \geq 0, k > 0\}$. So the result follows from(1).

Q.E.D.

Now, with the notation of (2) if n is the spectral measure of a zero-mean Gaussian stationary process $X = \{X_j: j \in \mathbf{Z}\}$, the *entropy rate* H(X) of X is such that [L]

$$H(X) = \lim_{n\to\infty} (1/2) \log [2\pi e D_n^{(1/n+1)}]$$

so

$$H(X) = (1/2)\log[2\pi e] + (1/4\pi) \int \log w\, dt.$$

Thus, the association to each f as in (3) of a Gaussian stationary process X with spectral measure $d\eta = (1-|f|^2)\, dt$ gives an entropy meaning to the integral

$\{(1/2\pi) \int \log(1-|f|^2) dt\}$. Proposition (1) says that the sum of the logarithms

of the step by step innovations converge to that entropy integral.

Calculation of the maximum entropy solution

So, when $\{c_0, c_1, ..., c_n\} \subset C$ is such that F_n has more than one element, we may say that the function $f \in F_n$ corresponding to the Schur parameters $\gamma_j = 0$ for every $j > n$ is the maximum entropy solution of the Carathéodory-Féjer problem. Of course, it can be obtained by means of Schur's algorithm [Sch]. Here we sketch an alternative method, based on the fact that every f $\in F_n$ is given by a unitary extension $(U,G) \in \mathbf{U}$ of a well defined isometry V and that the maximum entropy solution corresponds to the "most innovative" [A-2] element in **U**.

Since $\#(F_n) > 1$, the orthogonal complement $N \equiv N(n)$ of the domain D(n) of the isometry V in H(n) is one-dimensional and the same happens with the orthogonal complement $M \equiv M(n)$ of the range R(n) of V. Set $M_j \equiv M$ for every $j \leq 0$, $N_k \equiv N$ for every $k \geq 0$ and

$\tilde{G} = (\oplus\{M_j: j < 0\}) \oplus H(n) \oplus (\oplus\{N_k: k > 0\})$

Let S be the unilateral shift. Since $H(n) \oplus (\oplus\{N_k: k > 0\}) = D(n) \oplus (\oplus\{N_k: k \geq 0\})$

and $(\oplus\{M_j: j<0\}) \oplus H(n) = (\oplus\{M_j: j \leq 0\}) \oplus R(n)$, a unitary operator $\tilde{U}$ is defined by setting $U=S$ on $(\oplus \{ M_j : j < 0 \}) \oplus (\oplus \{ N_k : k \geq 0\})$ and $U = V$ on $D(n)$.

Then $(U,G) \in \mathbf{U}$ and, from the geometric enterpretation of Schur parameters, it follows that in this case $\gamma_k = 0$ for every $k > n$.

Consequently, $(\tilde{U},\tilde{G})$ gives the solution f we are interested in.

We shall use now the notation defined in the statement and proof of proposition B. Remark that, if for any trigonometric polynomial q we set $\rho(t) \equiv q(-t)$, then we have:

$$< (q_1, q_2), (q'_1, q'_2) >_{L^2(\alpha)} = \int_{\mathbf{T}} \{p_1 \bar{p}'_1 + p_1 \bar{p}'_2 f + p_2 \bar{p}'_1 \bar{f} + p_2 \bar{p}'_2\}.$$

Let $v \in N$ be a unit vector in N such that $< V^n d_1, d_2 > > 0$; thus, v is a weil defined linear combination of $\{V^j d_1, V^k d_2 : n \geq j \geq 0, n \geq k \geq 0\}$ and there exist two analytic trigonometric polynomials of degree at most n, q_1 and q_2, such that $Lv = (q_1, q_2) \in L^2(\alpha)$. Also, $L(S^j v) = (e_j q_1, e_j q_2)$, $\forall j \in \mathbf{Z}$. From $0 = < S^j v, v >$ for every $j > 0$ and $1 = <v, v>$, it follows that

$$p_1 \bar{p}_1 + p_1 \bar{p}_2 f + p_2 \bar{p}_1 \bar{f} + p_2 \bar{p}_2 \equiv 1 \text{ on } \mathbf{T}$$

Now, $p_1 \bar{p}_2 f = g + \phi$, with $\phi \in H^\infty$ and g an antianalytic polynomial of degree at most n completely determined by $\{c_0, c_1, ..., c_n\}$, such that

$$2 \operatorname{Re} \phi \equiv 1 - |p_1|^2 - |p_2|^2 - 2 \operatorname{Re} g \text{ on } \mathbf{T} \text{ and } \phi(0) = 0$$

Consequently, ϕ is also determined by $\{c_0, c_1, ..., c_n\}$ and f is given by the rational fraction

$$f = \{e_n g + e_n \phi\} / \{e_n p_1 \bar{p}2\}$$

In order to have explicit formulas note that

$$v = [1/(1-|\gamma_n|^2)] \, d_1(n) + [\gamma_n / (1 - |\gamma_n|^2)] \, d_2(n),$$

with $d_1(n)$ given by (16) in [A-1] and $d_2(n)$ by a similar expression.

Summing up.

Theorem C. *Let* $\{c_0, c_1, ..., c_n\} \subset \mathbf{C}$ *be such that* $F_n = \{f \in H^\infty(\mathbf{T}) : \|f\|_\infty \leq 1, \hat{f}(k) = c_k, 0 \leq k \leq n\}$ *has more than one element. Set* $c_j = 0$ *for* $-n \leq j \leq -1$ *and* $e_j = e^{ijt}$ *for every* $j \in \mathbf{Z}$, *call* $\mathbf{P}$ *the linear span of* $\{e_0, e_1, ..., e_n\}$ *and in* $\mathbf{P} \times \mathbf{P}$ *define a scalar product by setting, with* $\delta(0) = 1$ *and* $\delta(j) = 0$ if $j \neq 0$,

$$< (e_j, e_k), (e_{j'}, e_{k'}) > = \delta(j-j') + c_{j-k'} + \bar{c}_{j'-k} + \delta(k-k').$$

Then there exists one and only one $(q_1, q_2) \in \mathbf{P} \times \mathbf{P}$ *such that* $< (q_1, q_2), (e_j, 0) > = < (q_1, q_2), (0, e_k) > = 0$ for $0 \leq j < n$, $0 \leq j \leq n$, and $< (q_1, q_2), (e_n, 0) > > 0$.

Set $\rho_1(t) \equiv q_1(-t)$, $\rho_2(t) \equiv q_2(-t)$; if $\rho_1 \bar{\rho}_2 \, \Sigma \{c_j e_j : 0 \leq j \leq n\} = \Sigma \{g_j e_j : -n \leq j \leq 2n\}$, call $g = \Sigma \{g_j e_j : -n \leq j \leq 0\}$; let

$\phi \in \mathbf{P}$ *be determined by* $2 \operatorname{Re} \phi = 1 - |p_1|^2 - |p_2|^2 - 2 \operatorname{Re} g$ *and* $\phi(0) = 0$. *Set* $f_0 = (e_n g + \phi) / (e_n p_1 p_2)$.

Then: $f_0 \in F_n$ *and, for any* $f \in F_n$ such that $f \neq f_0$,

$$\int \log(1-|f_o|^2)\,dt > \int \log(1-|f|^2)\,dt.$$

Finally, we want to point out that our approach to maximum entropy aims to establish the relations of some results of [C] and [D-G] with the method of unitary extensions of isometries.

References

[A-1] AROCENA, R. *Un a geometric interpretation of Schur parameters,* contribution to the volume of the "Revista de la Union Matematica Argentina" dedicated to Mischa Cotlar.

[A-2] IDEM. *Schur analysis of a class of translation invariant forms,* in "Analysis and Partial Differential Equations" (Marcel Dekker) Marcel Dekker 1990, p.355-369. (C. Sadosky ed.).

[B] BOYD, D.W. *Schur's algorithm for bounded holomorphic functions,* Bull. London Math. Soc 11 (1979), 145-150.

[C] CONSTANTINESCU, T. *A Maximum Entropy Principale for Contractive intertwining Dilations,* Op. Theory: Adv. and Appl.24 (1987), 69-85.

[D-H] DYM, H. & GOHBERG, I.*A Maximum Entropy Principle for Contractive Interpolants* J. Funct. Anal. 65 (1986), 83-125.

[G] GARNETT, J. *Bounded Analytic Functions* , (Academic Press, 1981).

[G-S] GRENADER, U. & SZEGO, G. *Toeplitz forms and their applications,* (University of California Press, 1958).

[L] LANDAU, H.J. *Maximum entropy and the moment problem,* Bull Amer. Math. Soc. 16, No.1 (1987), 47-77.

[Sch] SCHUR, I. *On power series which are bounded in the interior of the unit circle, I, II,* first published in German in 1918 -1919. English Translation in I. Schur Methods in Operator Theory and Signal Processing (I.Gohberg, ed), Op. Theory: Adv. and Appl. 18 (1986), 31-88.

Recibido por U.M.A. el 16 de mayo de 1989.