©1997 M. Alfaro et al. (Eds.) Proceedings of the
International Workshop on Orthogonal Polynomials
in Mathematical Physics. Leganés, 24-26 June, 1996.

ENTROPY-CONVERGENCE, INSTABILITY IN STIELTJES AND HAMBURGER MOMENT PROBLEMS

M. Frontini and A. Tagliani

Abstract

The recovering of a positive density, of which a finite number of moments is assigned, is considered (in the Stieltjes and Hamburger moment problems). In the choice of the approximant the Maximum Entropy approach is adopted. Two main problems are taken into account.

1. A review of different criteria concerning the determinacy and indeterminacy of the infinite moment problem is presented. This review includes criteria based on the moments themselves or on the coefficients of the three-terms recurrence relation characterizing the orthogonal polynomials arising from the moment sequence. A particular criterion, guaranteeing the determinacy of the moment problem, is endowed with a geometrical interpretation, which allows the proof of entropy convergence. More precisely, it is proved that, whenever the sequence of assigned moments arises a determinate moment problem, then the approximants converge in entropy to the recovering function.

2. The entropy convergence allows us to employ a large number of moments. Unstability of the recovering is proved when an increasing number of moments is used. Subsequently the condition number, the ratio between the relative error of the function and the relative error of the modified moment, is analyzed. We also prove that the condition number is related to the orthogonal polynomials generated by the sequence of moments.

1 Introduction

A typical case in the Physics framework consists in the knowledge of the first few expectation values of a positive function $f(x)$ and in trying to obtain information on the values of such a function. We are concerned with an indeterminate moment problem, for which different solutions can be used.

The Maximum Entropy (ME) method, on Information Theory, chooses among all the admissible approximants, the least biased of them which is compatible with the partial information represented by few known moments. The analytical form of the approximant is obtained by maximizing the information Shannon-entropy functional under the constraints provided by the $M+1$ moments $\mu_0, ..., \mu_M$

$$S(f) = -\int_R f(x)\ln f(x)dx - (\lambda_0 - 1)(\int_R f(x)dx - \mu_0) - \sum_{j=1}^{M} \lambda_j (\int_R x^j f(x)dx - \mu_j). \qquad (1.1)$$

The analytical form of the approximant $f_M(x)$ is given by (see [5],[6])

$$f_M(x) = \exp(-\sum_{j=0}^{M} \lambda_j x^j) \qquad (1.2)$$

where the Lagrange multipliers satisfy the constraints

$$\mu_n = \int_R x^n \exp(-\sum_{j=0}^{M} \lambda_j x^j), \ n = 0, ..., M. \tag{1.3}$$

Successive moments can be found by the recursive relation

$$(n+1)\mu_n - \sum_{j=1}^{M} j\lambda_j \mu_{n+j} = 0, \ n \geq 1. \tag{1.4}$$

The entropy $H[f_M]$ of $f_M(x)$ is given by

$$H[f_M] = -\int_R f_M(x) \ln f_M(x) dx = \sum_{j=0}^{M} \lambda_j \mu_j. \tag{1.5}$$

Necessary and sufficient conditions for the existence of $f_M(x)$ when $R \equiv [0, +\infty)$ (Stieltjes moment problem) and $R \equiv (-\infty, +\infty)$ (Hamburger moment problem) are already known in the literature ([4]).
A further interesting question is connected with the convergence. A necessary condition for the convergence is the determinacy of the moment problem. For this pourpose different criteria are available in the literature. The criteria are provided in terms of the moments or in terms of the coefficient of the three-term recurrence relation characterizing orthogonal polynomials associated with the sequence of moments or the orthogonal polynomials themselves.
The aim of the paper is twofold.
1. To prove that whenever the assigned moments uniquely determine a function $f(x)$ then the ME approximants converge in entropy to $f(x)$.
2. If an increasing number of moments is taken into account then the λ_j evaluation, and consequently $f_M(x)$, is unstable rendering the problem computationally unsolvable. The condition number is related to orthogonal polynomials generated by the sequence of moments.

2 Known results about Stieltjes and Hamburger moment problem

We recall some classical results concerning the Stieltjes and Hamburger moment problem which will be used in the convergence proof. Such results are drawn essentially from [8].
Given a set of moments μ_j, $j = 0, 1, ...$ the Hankel determinants are defined as

$$\Delta_0 = \mu_0, \ \Delta_1 = \begin{vmatrix} \mu_0 & \mu_1 \\ \mu_1 & \mu_2 \end{vmatrix}, \ ..., \Delta_n = \begin{vmatrix} \mu_0 & \cdots & \mu_n \\ \vdots & \vdots & \vdots \\ \mu_n & \cdots & \mu_{2n} \end{vmatrix}, ... \tag{2.1}$$

$$\Delta_0^{(1)} = \mu_1, \ \Delta_1^{(1)} = \begin{vmatrix} \mu_1 & \mu_2 \\ \mu_2 & \mu_3 \end{vmatrix}, ..., \Delta_n^{(1)} = \begin{vmatrix} \mu_1 & \cdots & \mu_{n+1} \\ \vdots & \vdots & \vdots \\ \mu_{n+1} & \cdots & \mu_{2n+1} \end{vmatrix}, ... \tag{2.2}$$

i) Existence of $f_M(x)$
Except in some few cases the necessary and sufficient conditions for the existence of ME solution are identical to Hamburger or Stieltjes conditions ([4]). These conditions are given by the positivity of the Hankel determinants.

a) Hamburger moment problem ($M = 2n$, obviously)

$$\Delta_0 > 0, \ \ \Delta_1 > 0, ..., \Delta_n > 0. \tag{2.3}$$

b) Stieltjes moment problem
$M = 2n$

$$\Delta_0 > 0, \ \Delta_0^{(1)} > 0, \ \Delta_1 > 0, \ \Delta_1^{(1)} > 0, ..., \Delta_{n-1}^{(1)} > 0, \ \Delta_n > 0, \tag{2.4}$$

$M = 2n + 1$

$$\Delta_0 > 0, \ \Delta_0^{(1)} > 0, \ \Delta_1 > 0, \ \Delta_1^{(1)} > 0, ..., \Delta_n > 0, \ \Delta_n^{(1)} > 0. \tag{2.5}$$

ii) Determinacy
The determinacy of Stieltjes and Hamburger moment problems, unlike Hausdorff's moment problem, is not guaranted by the existence conditions of the finite moment problem when the number of moments increases. This fact, well known in the literature (see [8],[1]), is due to the non compactness of the support of the function $f(x)$ corresponding to the assigned sequence $\{\mu_k\}$. As a consequence the set of orthonormal polynomials generated by $\{\mu_k\}$ should be not complete in $L^2_{dF}(R)$, where $dF(x) = f(x)dx$.
For this reason several sufficient or necessary and sufficient conditions for the determinacy have been formulated, by analyzing the sequence $\{\mu_k\}$ of assigned moments, or the coefficients of three-term recurrence relation characterizing the orthonormal polynomials or the orthonormal polynomials themselves. Among these we recall, the ones characterized by the moment sequence $\{\mu_k\}$:
Hamburger case. A sufficient condition for the Hamburger moment problem to be determined is that the sequence $\{a_n\}$ $n = 1, 2, \ldots$ with $a_n = \frac{1}{n} \sqrt[2n]{\mu_{2n}}$ is not divergent [8].
Stieltjes case. A sufficient condition for the Stieltjes moment problem to be determined is that the sequence $\{a_n\}$ $n = 1, 2, \ldots$ with $a_n = \frac{1}{n^2} \sqrt[n]{\mu_n}$ is not divergent [8].
Concerning the three-term recurrence relation characterizing the monic orthogonal polynomials, the following criterion holds:
Carleman's criterion: The Hamburger moment problem is determined if $\sum\limits_{n=1}^{\infty} \lambda_n^{-\frac{1}{2}} = \infty$,
where

$$\begin{aligned} P_n(x) &= (x - c_n)P_{n-1}(x) - \lambda_n P_{n-2}(x) \\ P_{-1}(x) &= 0; \quad P_0(x) = 1; \quad c_n \in I\!R, \quad \lambda_n > 0, \quad n \geq 1 \end{aligned} \tag{2.6}$$

represents the sequence of monic orthogonal polynomials generated by $\{\mu_k\}$ ([2]), where other conditions in terms of c_n and λ_n are given.
The link between moment problem determinacy and orthonormal polynomials is emphasized by the following result.
Given $f(x) \in L^2_{dF}(R)$ the set of orthonormal polynomials $\{\tilde{P}_n(x)\}_{n \in N}$ is said to be complete if $\forall \epsilon > 0 \ \exists m = m(\epsilon)$ so that:

$$\|f(x) - p_m(x)\|^2 = \int_{-\infty}^{\infty} |f(x) - p_m(x)|^2 dF(x) < \epsilon$$

where $p_m(x) = \sum\limits_{k=0}^{m} b_k \tilde{P}_k(x)$.
The completeness is equivalent to the *Parseval* identity:

$$\sum_{k=0}^{\infty} |b_k|^2 = \int_{-\infty}^{\infty} |f(x)|^2 dF(x)$$

where

$$b_k = \int_{-\infty}^{\infty} f(x)\tilde{P}_k(x)dF(x), \qquad k = 0, 1, 2, \ldots$$

holds $\forall f(x) \in L^2_{dF}(-\infty, \infty)$.
The following two theorems connect the determinacy of the moment problem and the completeness of the corresponding sequence of orthonormal polynomials.
Theorem 2.1: A sufficient condition for the set $\{P_k(x)\}$ to be complete in $L^2_{dF}(-\infty, \infty)$ is that the corresponding moment problem be determined [9].
Theorem 2.2: A necessary and sufficient condition for the set $\{P_k(x)\}$ to be complete in $L^2_{dF}(-\infty, \infty)$ is that the moment problem corresponding to $dF^* = \frac{dF}{1+x^2}$, whose moments μ_k^* are related to μ_k moments through the relationship $\mu_k^* + \mu_{k+2}^* = \mu_k$, $k = 0, 1, 2, \ldots$, be determined [3].
To the sequence $\{\mu_k\}$ we can associate the following minimum problem

$$\begin{aligned} \rho_n(z_0) &= \min_{p_n} \int_{-\infty}^{\infty} \mid p_n(x) \mid^2 dF(x) \\ p_n(z_0) &= 1, \end{aligned} \tag{2.7}$$

where $p_n(x)$ is a polynomial of degree $\leq n$ which assumes the value 1 at $x = z_0$ in the form $p_n(x) = \sum_{k=0}^{n} b_k \pi_k(x)$, where $\pi_k(x)$ indicates the orthonormal polynomial of degree k associated with $F(x)$.
It may be proved that [8]

$$\frac{1}{\rho_n(z_0)} = \sum_{k=0}^{n} |\pi_k(z_0)|^2. \tag{2.8}$$

From the Christoffel-Darboux's formula

$$\rho_n(z_0) = \frac{\Delta_n}{\begin{vmatrix} \mu_0 & \mu_1 & \cdots & \mu_n & 1 \\ \mu_1 & \mu_2 & \cdots & \mu_{n+1} & z_0 \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ \mu_n & \mu_{n+1} & \cdots & \mu_{2n} & z_0^n \\ 1 & \overline{z}_0 & \cdots & \overline{z}_0^n & 0 \end{vmatrix}} \tag{2.9}$$

holds. If $\rho(z_0) = \lim_{n \mapsto \infty} \rho_n(z_0)$ we have
Theorem 2.3: In order that the Hamburger moment problem may be determined it is necessary and sufficient that $\rho(z_0) = 0 \ \forall z_0 = x + iy$ with $y > 0$ [8].
Another criterion for the determinacy is obtained by considering the quantities (see [8]) $\rho(0) = \lim_{n \mapsto \infty} \rho_n(0)$, $\rho^{(1)}(0) = \lim_{n \mapsto \infty} \rho_n^{(1)}(0)$ and $\rho^{(2)}(0) = \lim_{n \mapsto \infty} \rho_n^{(2)}(0)$ corresponding to sequences $\{\mu_0, \mu_1, \ldots\}$, $\{\mu_1, \mu_2, \ldots\}$ and $\{\mu_2, \mu_3, \ldots\}$ respectively.
Relation (2.9) may be written, at $z_0 = 0$, through Hankel's determinants

$$\rho_n(0) = \frac{\Delta_n}{\begin{vmatrix} \mu_2 & \cdots & \mu_{n+1} \\ \vdots & \vdots & \vdots \\ \mu_{n+1} & \cdots & \mu_{2n} \end{vmatrix}}, \tag{2.10}$$

and similarly for $\rho^{(1)}(0)$ and $\rho^{(2)}(0)$ corresponding to the sequence $\{\mu_1, \mu_2, \ldots\}$ and $\{\mu_2, \mu_3, \ldots\}$ respectively. The following theorems hold
Theorem 2.4: A necessary and sufficient condition for the Hamburger moment problem be determined is that at least one of the quantities

$$\rho(0) = \lim_{n \mapsto \infty} \rho_n(0), \qquad \rho^{(2)}(0) = \lim_{n \mapsto \infty} \rho_n^{(2)}(0) \tag{2.11}$$

be equal to zero [8].

Theorem 2.5: A necessary and sufficient condition for the Stieltjes moment problem be determined is that at least one of the quantities

$$\rho(0) = \lim_{n\mapsto\infty} \rho_n(0), \qquad \rho^{(1)}(0) = \lim_{n\mapsto\infty} \rho_n^{(1)}(0) \tag{2.12}$$

be equal to zero [8].
Remark 2.1: Notice that whenever $\rho(0) = 0$ then both the Hamburger and Stieltjes moment problems, corresponding to the same moment sequence, are determined. On the other hand if $\rho(0) > 0$ and $\rho^{(1)}(0) = 0$, then the Stieltjes moment problem, corresponding to the given moment sequence, is determined, unlike the Hamburger moment problem (if $\rho^{(2)}(0) > 0$ too). Thus Theorems 2.4 and 2.5 can by formulated in terms of Hankel's determinants.
We are in position to give a geometrical meaning to the existence and determinacy conditions of the moment problems.

3 Geometrical interpretation

We limit ourselves to consider in detail the Stieltjes moment problem. The Hamburger problem may be similarly treated.
i) From existence.
Let $\mu_{0,n}^-$ the value of μ_0 so that $\Delta_n = 0$. Applying the Laplace rule to the first row

$$\Delta_n = \mu_0 \begin{vmatrix} \mu_2 & \cdots & \mu_{n+1} \\ \vdots & \vdots & \vdots \\ \mu_{n+1} & \cdots & \mu_{2n} \end{vmatrix} + c, \tag{3.1}$$

(where c is a constant depending of $\mu_1, ..., \mu_{2n}$);

$$0 = \mu_{0,n}^- \begin{vmatrix} \mu_2 & \cdots & \mu_{n+1} \\ \vdots & \vdots & \vdots \\ \mu_{n+1} & \cdots & \mu_{2n} \end{vmatrix} + c, \tag{3.2}$$

from which

$$\rho_n(0) = \Delta_n \Big/ \begin{vmatrix} \mu_2 & \cdots & \mu_{n+1} \\ \vdots & \vdots & \vdots \\ \mu_{n+1} & \cdots & \mu_{2n} \end{vmatrix} = \mu_0 - \mu_{0,n}^-, \tag{3.3}$$

is the distance between μ_0 and its lower bound $\mu_{0,n}^-$.
Similar relations for $\rho_n^{(1)}(0)$ and $\rho_n^{(2)}(0)$ hold.
Quantities $\rho_n(0)$, $\rho_n^{(1)}(0)$ and $\rho_n^{(2)}(0)$ benefit from an interesting geometrical interpretation because they represent respectively the distance between μ_0 and $\mu_{0,n}^-$, μ_1 and $\mu_{1,n}^-$ and between μ_2 and $\mu_{2,n}^-$, where $\mu_{0,n}^-$, $\mu_{1,n}^-$ and $\mu_{2,n}^-$ are the allowable lower bounds of μ_0, μ_1 and μ_2, when first $2n$ or $2n+1$ moments are assigned. More precisely

$$\rho_n(0) = \mu_0 - \mu_{0,n}^-, \qquad \rho_n^{(1)}(0) = \mu_1 - \mu_{1,n}^-, \qquad \rho_n^{(2)}(0) = \mu_2 - \mu_{2,n}^-. \tag{3.4}$$

Considering the sequence of given moments starting from μ_k, generalizing (2.11) and (3.4) we get:

$$\rho_n^{(k)}(0) = \mu_k - \mu_{k,n}^-, \tag{3.5}$$

where:

$$\rho_n^{(k)}(0) = \begin{vmatrix} \mu_k & \cdots & \mu_{n+k} \\ \vdots & \vdots & \vdots \\ \mu_{n+k} & \cdots & \mu_{2n+k} \end{vmatrix} \Big/ \begin{vmatrix} \mu_{k+2} & \cdots & \mu_{n+k+1} \\ \vdots & \vdots & \vdots \\ \mu_{n+k+1} & \cdots & \mu_{2n+k} \end{vmatrix}. \tag{3.6}$$

To summarize, given a sequence of moments $\{\mu_0, \mu_1, \ldots\}$, there exists the lower bound $\mu_{k,n}^-$, depending on n and the sequence itself, of every μ_k.
We will now prove that an upper limit for every moment μ_k exists when the other moments $\mu_{k-1}, \mu_{k+1}, \ldots, \mu_{2n+k-1}$ are fixed. For $k = 1$ consider again the moments $\{\mu_0, \ldots, \mu_{2n}\}$, with $\mu_0, \mu_2, \ldots, \mu_{2n}$ fixed, while μ_1 is varied continuously. $\Delta_n(\mu_1) = 0$ is an algebraic quadratic equation whose leading coefficient is negative, with absolute value provided by the determinant of the Hankel matrix obtained by deleting the first two rows and columns of Δ_n. This equation has two real and distinct roots μ_1^{nl} and μ_1^{nr}, with $\mu_1^{nl} < \mu_1^{nr}$. Because of the positivity of the determinants Δ_k, $\Delta_k^{(1)}$, $k = 0, 1, \ldots, n-1$, the roots μ_1^{nl} and μ_1^{nr} satisfy the inequalities

$$\mu_1^{nl} < \mu_{1,n}^- < \mu_1 < \mu_1^{nr} \leq \mu_1^{(n-1)r} \leq \ldots \leq \mu_1^{1r}. \tag{3.7}$$

Then $\mu_{1,n}^+ = \mu_1^{nr}$ represents the upper bound of μ_1 when moments $\mu_0, \mu_2, \ldots, \mu_{2n}$ are fixed. Fixed $\mu_0, \mu_2, \ldots, \mu_{2n}$ the interval of admissible values of μ_1 for which the determinant defined by (2.1), (2.2) are positive, is given by $[\mu_{1,n}^-, \mu_{1,n}^+]$.
In a similar way, from (2.2) we obtain $\Delta_n^{(1)}(\mu_2) = 0$ for the value $\mu_{2,n}^+$ and so the interval $[\mu_{2,n}^-, \mu_{2,n}^+]$ of the admissible values of μ_2, when the remaining moments are fixed.
Analogously, by using the sequence $\mu_{k-1}, \mu_k, \ldots, \mu_{2n+k-1}$ it is possible to show that every moment μ_k has an interval of admissible values:

$$[\mu_{k,n}^-, \mu_{k,n}^+]. \tag{3.8}$$

Remark 3.1 The moments μ_0 and μ_M admit only lower bounds $\mu_{0,n}^-$ and $\mu_{M,n}^-$ respectively given by $\Delta_n(\mu_0) = 0$, $\Delta_n(\mu_M) = 0$.
ii) From unicity
Unicity conditions imply, in the moment space, the relationships

$$\lim_{n\to\infty} \mu_{0,n}^- = \mu_0 \text{ or } \lim_{n\to\infty} \mu_{1,n}^- = \mu_1 \tag{3.9}$$

If the sequence of moments $\{\mu_k\}$ determines a unique function $f(x)$, fixed an integer k, also the family of functions

$$f_{S,k}(x) = x^k f(x), \tag{3.10}$$

arises a sequence of moments to which corresponds a determined Stieltjes problem. Thus, analogously to (2.12) considering (3.6) we get

$$\lim_{n\to\infty} \rho_n^{(k)}(0) = 0. \tag{3.11}$$

By combining (3.5) and (3.11) it is possible to show that, fixed an integer k, with $k < n$ and k independent from n, the moment μ_k, as n increases, satisfies the relation

$$\lim_{n\to\infty} \mu_{k,n}^- = \mu_k, \tag{3.12}$$

so that the interval defined by (3.8) closes from the left on μ_k.
We show now that the interval (3.8) closes from the right on μ_k, by using the following auxiliary functions.
For a fixed k, let consider the family of positive functions

$$g_{S,k}(x) = (1 - x + x^2)x^{k-1} f(x) \qquad x \geq 0 \tag{3.13}$$

whose moments η_j^k are related to the moments of $f(x)$ by

$$\eta_j^k = \mu_{k-1+j} - \mu_{k+j} + \mu_{k+1+j}, \qquad j = 0, 1, 2, \ldots \tag{3.14}$$

Functions (3.13) define an infinite sequence of moments to which corresponds a determined Stieltjes problem. Let us consider the moments $\mu_{k-1}, ..., \mu_{2n+k-1}$.
By fixing the moments $\mu_{k-1}, \mu_{k+1}, \mu_{k+2}, ..., \mu_{2n+k-1}$, while μ_k varies continuously. We consider the sequence of moments $\{\eta_0^k, \eta_1^k, ..., \eta_{2n}^k\}$ and (3.6), with μ_j substituted by η_j^k. Relations (3.11), arising by the determinacy of Stieltjes moment problem, and (3.14) imply that upper bound of the admissible values of μ_k, indicated by $\mu_{k,n}^+$, represents a decreasing monotonic sequence having μ_k as its limit when n increases. So, analogously as for the lower bound of μ_k, fixed an integer k, the moment μ_k, as n increases, satisfies the relation

$$\lim_{n\to\infty} \mu_{k,n}^+ = \mu_k. \tag{3.15}$$

As a consequence, from the determinacy of the moment problem, the interval defined by (3.8) closes from the left and right on μ_k.
Similar results may be obtained in the Hamburger case if we substitute (3.10) with

$$\begin{aligned} f_{H,2k}(x) &= x^{2k} f(x), \qquad \text{and} \\ f_{H,2k}^*(x) &= (1+x+x^2)x^{2k} f(x), \end{aligned} \tag{3.16}$$

in such a way that

$$\begin{aligned} \lim_{n\to\infty} \mu_{2k,n}^- &= \mu_{2k}, \\ \lim_{n\to\infty} \mu_{2k+1,n}^- &= \mu_{2k+1}, \end{aligned} \tag{3.17}$$

respectively. By substituting (3.13) with

$$\begin{aligned} g_{H,2k}(x) &= (1-x+x^2)x^{2k} f(x), \qquad \text{and} \\ g_{H,2k}^*(x) &= (1-x^2+x^4)x^{2(k-1)} f(x), \end{aligned} \tag{3.18}$$

we prove that

$$\begin{aligned} \lim_{n\to\infty} \mu_{2k+1,n}^+ &= \mu_{2k+1}, \\ \lim_{n\to\infty} \mu_{2k,n}^+ &= \mu_{2k}, \end{aligned} \tag{3.19}$$

respectively.

4 Entropy-convergence

It is evident that entropy-convergence implies determinacy of the moment problem. Furthermore entropy-convergence means $\lim\limits_{M\to\infty} H[f_M] = H[f]$.
i) Hamburger case ($M = 2n$)
Let us consider two cases separately.
a) $\rho(0) = \lim\limits_{n\to\infty} \rho_n(0) = 0$.
Let $\mu_1, ..., \mu_M$ be assigned and μ_0 varies continuously. Then from (1.2) the Lagrange multipliers λ_j are functions of μ_0. By differentiating (1.3) with respect to μ_0 we have the linear differential system

$$\begin{bmatrix} \mu_0 & \cdots & \mu_M \\ \vdots & \vdots & \vdots \\ \mu_M & \cdots & \mu_{2M} \end{bmatrix} \cdot \begin{bmatrix} d\lambda_0/d\mu_0 \\ d\lambda_1/d\mu_0 \\ \vdots \\ d\lambda_M/d\mu_0 \end{bmatrix} = - \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{4.1}$$

(where unknown moments $\mu_{M+1}, ..., \mu_{2M}$ are obtained from (1.4)). From (1.5), by taking into account the first equation of (4.1),

$$\frac{d}{d\mu_0} H[f_M] = \lambda_0 + \sum_{j=0}^{M} \mu_j \frac{d\lambda_j}{d\mu_0} = \lambda_0 - 1, \tag{4.2}$$

and by (2.11),

$$\frac{d^2}{d\mu_0^2}H[f_M] = \frac{d\lambda_0}{d\mu_0} = -\frac{1}{\rho_M(0)} < 0 \tag{4.3}$$

hold. Thus $H[f_M]$ is a differentiable concave function of μ_0.
We are in position to prove that $\lim_{M\to\infty} H[f_M] = H[f]$.
Let us suppose that

$$\lim_{M\to\infty} H[f_M] = H_{lim} > H[f] \tag{4.4}$$

holds. When, for given μ_0, $M \to \infty$ we have from a geometrical interpretation of determinacy conditions $\lim_{M\to\infty} \mu_{0,n}^{-} = \mu_0$ while, with M fixed, by varying μ_0

$$\lim_{\mu_0 \to \mu_{0,n}^{-}} H[f_M] = -\infty \tag{4.5}$$

as $f_M(x)$ degenerates in a set of n masses.
Let us consider the following sequences of ME solutions
1. $f_M(x)$ with moments $\mu_0, ..., \mu_M$
2. $f_M^{\epsilon}(x)$ with moments $\mu_0 - \epsilon, \mu_1, ..., \mu_M$ with $\epsilon \in I\!R^+$ and $\mu_0 - \epsilon > \mu_{0,M}^{-}$ so that $f_M^{\epsilon}(x)$ exists and $H[f_M^{\epsilon}] \leq H[f]$ holds. As $M \to \infty$ then from (2.11) $\epsilon \to 0$ so that the moment μ_0 of $f_M(x)$ and $f_M^{\epsilon}(x)$ differ by an arbitrarily small quantity, while their entropies differ by a finite quantity greater than $H_{lim} - H[f]$.
Taking into account (4.4) we have reached a contradiction on the continuity of $H[f_M]$ as a function of μ_0. Then $H_{lim} = H[f]$ holds and so

$$\lim_{M\to\infty} H[f_M] = H[f] \tag{4.6}$$

is proved.
b) $\rho(0) = c > 0$, and $\rho^{(2)}(0) = \lim_{n\to\infty} \rho_n^{(2)}(0) = 0$.
Let $\mu_0, \mu_1, \mu_3 ..., \mu_M$ be assigned and μ_2 varies continuously. Then from (1.3) the Lagrange multipliers λ_j are functions of μ_2. By differentiating (1.3) with respect to μ_2 we have the linear differential system

$$\begin{bmatrix} \mu_0 & \cdots & \mu_M \\ \vdots & \vdots & \vdots \\ \mu_M & \cdots & \mu_{2M} \end{bmatrix} \cdot \begin{bmatrix} d\lambda_0/d\mu_2 \\ \vdots \\ d\lambda_M/d\mu_2 \end{bmatrix} = -\,[\,0\;0\;1\;0\;\ldots\;0\,]^T \tag{4.7}$$

(where the unknown moments $\mu_{M+1}, ..., \mu_{2M}$ are obtained from (1.4)). From (1.5), by taking into account the first equation of (4.7),

$$\frac{d}{d\mu_2}H[f_M] = \lambda_2 + \sum_{j=0}^{M} \mu_j \frac{d\lambda_j}{d\mu_2} = \lambda_2 \tag{4.8}$$

then

$$\frac{d^2}{d\mu_2^2}H[f_M] = \frac{d\lambda_2}{d\mu_2} = -\frac{\Delta_{M,2}}{\Delta_M} \tag{4.9}$$

holds. Here $\Delta_{M,2}$ is obtained from Δ_M by deleting the third row and the third column.
Being $\Delta_{M,2}$ the determinant of the principal matrix $P_{M,2}^T[\Delta_M]P_{M,2}$ of order M where $P_{M,2}^T$ is the permutation matrix interchanging, starting from the third row, every row with the successive one, is positive. Thus $H[f_M]$ is a differentiable concave function of μ_2. By taking into account the geometrical meaning of $\rho^{(2)}(0)$ the same conclusions like in the previous case can be drawn. The obtained result can be formulated by means of

Theorem 4.1: In the Hamburger case whenever the sequence of assigned moments determines uniquely a density function then ME approximants converge in entropy to $f(x)$.
ii) Stieltjes case
Once a moment sequence is assigned, the Stieltjes problem should be determined, unlike the Hamburger moment. Thus it is necessary to consider the two cases separately.
a) $\rho(0) = \lim_{n\to\infty} \rho_n(0) = 0$ (Stieltjes and Hamburger's problems determined).
The formal procedure is similar to the Hamburger case. By differentiating (1.3) with respect μ_0, (4.1)- (4.5) are obtained again and conclusion (4.6) is reached.
b) $\rho^{(1)}(0) = \lim_{n\to\infty} \rho_n^{(1)}(0)$ (Stieltjes problem determined and Hamburger problem indetermined).
It follows

$$\rho(0) = \lim_{n\to\infty} \rho_n(0) = c > 0, \qquad \rho^{(1)}(0) = \lim_{n\to\infty} \rho_n^{(1)}(0) = 0$$

where c is a positive constant. Let $\mu_0, \mu_2, ..., \mu_M$ be assigned and μ_1 varied continuously. Then from (1.3) the Lagrange multipliers λ_j are function of μ_1. By differentiating (1.3) with respect to μ_1 we have the differential linear system

$$\begin{bmatrix} \mu_0 & \cdots & \mu_M \\ \vdots & \vdots & \vdots \\ \mu_M & \cdots & \mu_{2M} \end{bmatrix} \cdot \begin{bmatrix} d\lambda_0/d\mu_1 \\ \vdots \\ d\lambda_M/d\mu_1 \end{bmatrix} = -\,[0\;1\;0\;\ldots\;0]^T \tag{4.10}$$

(where unknown moments $\mu_{M+1}, ..., \mu_{2M}$ are obtained from (1.4)). From (1.5), by taking into account the first equation of (4.10),

$$\frac{d}{d\mu_1} H[f_M] = \sum_{j=0}^{M} \mu_j \frac{d\lambda_j}{d\mu_1} + \lambda_1 = \lambda_1, \tag{4.11}$$

then

$$\frac{d^2}{d\mu_1^2} H[f_M] = \frac{d\lambda_1}{d\mu_1} = -\frac{\Delta_{M,1}}{\Delta_M}. \tag{4.12}$$

holds. Here $\Delta_{M,1}$ is obtained from Δ_M by deleting the second row and the second column.
Being $\Delta_{M,1}$ the determinant of the principal matrix $P_{M,1}^T[\Delta_M]P_{M,1}$ of order M where $P_{M,1}^T$ is the permutation matrix interchanging, starting from the second row, every row with the successive one, is positive. Thus $H[f_M]$ is a differentiable concave function of μ_1. By taking into account the geometrical meaning of $\rho^{(1)}(0)$ the same conclusions as in the Hamburger case can be drawn.
The obtained result can be formulated by means of
Theorem 4.2: In the Stieltjes case whenever the sequence of assigned moments uniquely determines a density function then ME approximants converge in entropy to $f(x)$.

5 Stability analysis

Let us study the stability, bounding ourselves to the case $M = 2n$, without loss of generality
1) of the parameters λ_j computation,
2) of the function $f_M(x)$ as the number M of assigned moments increases.
5.1 Multipliers $\lambda_0, ..., \lambda_M$ stability.
We now prove that, as the number M of moments increases, the λ_j multipliers computation determining $f_M(x)$, becomes unstable whenever the Stieltjes or Hamburger moment problems is determined.
Having proved that every moment μ_k varies into the interval $[\mu_{k,n}^-, \mu_{k,n}^+]$, and that, for a fixed k the moment μ_k is into an interval of admissible values going to zero as M increases. A consequence

of the decreasing range of μ_k admissible values is the unstability of the λ_k computation. In fact, fixed all the moments, except μ_k, we get, by varying μ_k continuously

$$\frac{d\lambda_k}{d\mu_k} = (-1)^{M-k+1}\frac{\Delta_{M,k}}{\Delta_M} \tag{5.1}$$

(the moments $\mu_{M+1}, \ldots, \mu_{2M}$ are computed by (1.4)) where $\Delta_{M,k}$ is obtained from Δ_M by deleting the $k+1$-th row and column.
Being $\Delta_{M,k}$ the determinant of the principal matrix $P^T_{M,k}[\Delta_M]P_{M,k}$ of order M where $P^T_{M,k}$ is the permutation matrix interchanging, starting from the $k+1$-th row, every row with the successive one, is positive. Hence λ_k are monotonic functions of μ_k in the interval $[\mu^-_{k,n}, \mu^+_{k,n}]$. As M increases the size of the interval $[\mu^-_{k,n}, \mu^+_{k,n}]$ tends to zero. As λ_k may assume all the values in $(-\infty, +\infty)$ when μ_k is varying in the interval $[\mu^-_{k,n}, \mu^+_{k,n}]$, λ_k must have higher gradient.
5.2 Behaviour of $f_M(x)$ as the last moment μ_M is modified.
Let $f_M(x;\lambda_j)$ the ME solution corresponding to the set of moments $\mu_0, \ldots, \mu_M$ and $f_M(x;\lambda_j+\Delta\lambda_j)$ the ME solution corresponding to the set of moments $\mu_0, \ldots, \mu_{M-1}, \mu_M + \Delta\mu_M$.
Let us consider the following quantity

$$\frac{f_M(x;\lambda_j+\Delta\lambda_j) - f_M(x;\lambda_j)}{f_M(x;\lambda_j)} := \epsilon[f_M(x;\Delta\mu_M)], \tag{5.2}$$

representing the relative error on $f_M(x)$ generated by the variation $\Delta\lambda_j$ on λ_j given by a variation of $\Delta\mu_M$ on μ_M.
Because of the particular analytical form of $f_M(x)$ we get:

$$\epsilon[f_M(x;\Delta\mu_M)] = e^{-\sum_{j=0}^{M}\Delta\lambda_j x^j} - 1 = e^{-\Delta\mu_M\sum_{j=0}^{M}\frac{\Delta\lambda_j}{\Delta\mu_M}x^j} - 1; \tag{5.3}$$

by differentiating (1.3) with respect to μ_M we obtain

$$\frac{d\lambda_j}{d\mu_M} = -\frac{\Delta^{\{j\}}_M}{\Delta_M}, \qquad j = 0, 1, \ldots, M, \tag{5.4}$$

where $[\Delta^{\{j\}}_M]$ is obtained from the matrix $[\Delta_M]$ when we substitute the j-th column by the vector $e_M = [0, \ldots, 0, 1]^T$ we obtain

$$\epsilon[f_M(x;\Delta\mu_M)] = e^{\frac{\Delta\mu_M}{\Delta_M}\sum_{j=0}^{M}\Delta^{\{j\}}_M x^j} - 1 = e^{\frac{\Delta\mu_M}{\Delta_M}P_M(x)} - 1; \tag{5.5}$$

where

$$P_M(x) = \begin{vmatrix} \mu_0 & \cdots & \mu_{M-1} & 1 \\ \mu_1 & \cdots & \mu_M & x \\ \vdots & \vdots & \vdots & \vdots \\ \mu_M & \cdots & \mu_{2M-1} & x^M \end{vmatrix} \tag{5.6}$$

is, up to a constant factor, the orthogonal polynomial of degree M defined by the first $2M-1$ moments of $f_M(x)$.
From (5.6), by well known properties of the zeros of orthogonal polynomials, there exist M points where $\epsilon[f_M(x;\Delta\mu_M)] = 0$ holds. ME solutions $f_{M-1}(x)$ and $f_M(x)$, having the same first $M-1$ moments, have exactly M common points. This property can be considered a characteristic of the ME solutions (eq.(1.2)) as, in general, two continuous functions having the same first $M-1$ moments have at least M common points (see [7]).
From (5.5), by using Taylor expansion, we get

$$\epsilon[f_M(x;\Delta\mu_M)] \simeq \frac{\Delta\mu_M}{\Delta_M}P_M(x) = \mu_M\frac{P_M(x)}{\Delta_M}\frac{\Delta\mu_M}{\mu_M} \tag{5.7}$$

Thus the relative error on μ_M is amplified by the factor

$$\mu_M \frac{| P_M(x) |}{\Delta_M} \tag{5.8}$$

representing a condition number of $f_M(x)$ when μ_M is varied.
From (5.5) notice that if the number of moments increases, then a perturbation on the last moment determines a zero variation of the function in an increasing number M of points, located exactly in the zeros of the orthogonal polynomial $P_M(x)$.

6 Conclusions

The classical Stieltjes and Hamburger's moment problem and sequence of maximum entropy approximants, whose first moments are assigned, have been considered and it has been proved that, whenever an infinite moment problem is determined, maximum entropy approximants converge in entropy to a function characterized by given moments. By resorting to geometrical meaning of determinacy, the unstability in the approximant evaluation has been proved and the condition number has been defined in terms of orthogonal polynomials arising from the sequence of assigned moments.

References

[1] N.I. Akhiezer: The classical moment problem, Oliver and Boyd, Edinburgh, 1965.

[2] T.S. Chihara: An introduction to orthogonal polynomials, Gordon and Breach Science Publishers, New York, 1978.

[3] G. Freud: Orthogonal polynomials, Pergamon Press, New York, 1966.

[4] M. Frontini, A. Tagliani *Maximum entropy in the finite Stieltjes and Hamburger moment problem* J. Math. Physics. **35**. (1994). 6748-6756.

[5] E.T. Jaynes: *Where do we stand on maximum entropy* in *The maximum entropy formalism* (R.D. Levine and M. Tribus Eds.) MIT Press Cambridge MA, (1978),15-118.

[6] J.N.Kapur: Maximum-Entropy Models in Science and Engineering, (John Wiley Sons, 1989).

[7] G. Polya and G. Szegő: Problems and Theorems in Analysis, **Vol. 1**, (Springer-Verlag, 1972).

[8] J.A. Shohat and J.D. Tamarkin: The problem of moments, (American Mathematical Society, Providence, RI, 1963).

[9] G. Szegő: Orthogonal polynomials, (American Mathematical Society, Providence, RI, 1975). Fourth edition.

Marco Frontini, Aldo Tagliani
Politecnico di Milano
Dipartimento di Matematica
Via Bonardi, 9
20133 Milano (ITALIA)
marfro@mate.polimi.it, aldtag@mate.polimi.it