# Resolution of the finite Markov moment problem ${ }^{\star}$ 

Laurent Gosse<br>IAC-CNR "Mauro Picone" (sezione di Bari)<br>Via Amendola 122/D - 70126 Bari, Italy<br>Olof Runborg*<br>NADA, KTH, 10044 Stockholm, Sweden


#### Abstract

We expose in full detail a constructive procedure to invert the so-called "finite Markov moment problem". The proofs rely on the general theory of Toeplitz matrices together with the classical Newton's relations.


Key words: Inverse problems, Finite Markov's moment problem, Toeplitz matrices.
1991 MSC: 65J22

## Résumé

Nous présentons en détail une procédure constructive pour inverser le "problème fini des moments de Markov". Les preuves reposent sur la théorie générale des matrices de Toeplitz et les classiques relations de Newton.

[^0]
## Version francaise abrégée

Afin d'inverser le systeme fini et mal conditionné (1), Koborov, Sklyar et Fardigola, $[9,13]$ ont proposé un algorithme recursif non-linéaire. Dans [7] nous avons prouvé un lemme le réduisant a l'extraction de valeurs propres généralisées, voir (4). Cette Note vise a expliquer en détail les raisons pour lesquelles cette procédure simplifiée résout le problème des moments de Markov. Apres avoir rappelé quelques éléments de la théorie des matrices de Toeplitz et les relations de Newton (Propositions 1 et 2), nous reformulons cet algorithme simplifié afin d'établir facilement certains lemmes techniques. Finalement, le Théorème 1 démontre le lien entre valeurs propres généralisées (4) et l'inversion de (1).

## 1 Introduction

We aim at inverting a moment system often associated with the prestigious name of Markov, as appearing in [2,3,6,11,12] in several fields of application; consult [5,10,14] for general background on moment problems. The original problem is the following. Given the moments $m_{k}$ for $k=1, \ldots, K$, find a bounded measurable density function $f$ and a real value $X>0$ such that
(1) $\int_{0}^{X} f(\xi) \xi^{k-1} d \xi=m_{k}, \quad k=1, \ldots, K$,
(2) $|f(\xi)|=1$ almost everywhere on $] 0, X[$,
(3) $f$ has no more than $K-1$ discontinuity points inside $] 0, X[$.

The solution is a piecewise constant function taking values in $\{-1,1\}$ a.e. on $] 0, X[$ and changing sign in at most $K-1$ points, which we denote $\left\{u_{k}\right\}$, ordered such that $0 \leq u_{1} \leq \cdots \leq u_{K}=X$. Finding $\left\{u_{k}\right\}$ from $\left\{m_{k}\right\}$ is an ill-conditioned problem when the $u_{k}$ values come close to each other; its Jacobian is a Vandermonde matrix and iterative numerical resolution routines require extremely good starting guesses. For less than four moments, however, a direct method based on solving polynomial equations was presented in [12]. Here we are concerned with an arbitrary number of moments $K \in \mathbb{N}$.

We consider a slightly modified version of the problem where $f$ takes values in $\{1,0\}$ instead of $\{-1,1\}$ and the moments are scaled as $m_{k} \rightarrow k m_{k}$. This is the precise setting for the applications in geometrical optics that we are interested in [7]. Moreover, to simplify the discussion we confine ourselves to the case when $K$ is even, setting $K=: 2 n$. The resulting problem can then be written as an algebraic system of nonlinear equations: Given $m_{k}$ find $u_{k}$ such that

$$
\begin{equation*}
m_{k}=\sum_{j=1}^{n} u_{2 j}^{k}-u_{2 j-1}^{k}, \quad k=1, \ldots, K=2 n \tag{1}
\end{equation*}
$$

An algorithm for solving this problem was presented by Koborov, Sklyar and Fardigola in $[9,13]$. It requires solving a sequence of high degree polynomial equations, constructed through a rather complicated process with unclear stability properties. In [7] we showed that the algorithm can be put in a more simple form that makes it much more suitable for numerical implementation. The simplified algorithm reads
(1) Construct the matrices $A$ and $B$ :

$$
A=\left(\begin{array}{cccc}
1 & &  \tag{2}\\
-m_{1} & 2 & \\
\vdots & \ddots & \ddots & \\
-m_{2 n-1} & \ldots & -m_{1} & 2 n
\end{array}\right), \quad B=\left(\begin{array}{cccc}
1 & & \\
m_{1} & 2 & & \\
\vdots & \ddots & \ddots & \\
m_{2 n-1} & \ldots & m_{1} & 2 n
\end{array}\right)
$$

(2) Let $\boldsymbol{m}=\left(m_{1}, m_{2}, \ldots, m_{2 n}\right)^{T}$ and solve the lower triangular Toeplitz linear systems

$$
\begin{equation*}
A \boldsymbol{a}=\boldsymbol{m}, \quad B \boldsymbol{b}=-\boldsymbol{m} \tag{3}
\end{equation*}
$$

to get $\boldsymbol{a}$ and $\boldsymbol{b}$.
(3) Construct the matrices $A_{1}, A_{2}$ from $\boldsymbol{a}=\left(a_{1}, a_{2}, \ldots, a_{2 n}\right)^{T}$ as

$$
A_{1}=\left(\begin{array}{cccc}
a_{1} & a_{2} & \ldots & a_{n} \\
a_{2} & a_{3} & \ldots & a_{n+1} \\
\vdots & \vdots & \ddots & \vdots \\
a_{n} & a_{n+1} & \ldots & a_{2 n-1}
\end{array}\right), \quad A_{2}=\left(\begin{array}{cccc}
a_{2} & a_{3} & \ldots & a_{n+1} \\
a_{3} & a_{4} & \ldots & a_{n+2} \\
\vdots & \vdots & \ddots & \vdots \\
a_{n+1} & a_{n+2} & \ldots & a_{2 n}
\end{array}\right)
$$

and the corresponding matrices $B_{1}, B_{2}$ from $\boldsymbol{b}$.
(4) Compute the generalized eigenvalues of the problems

$$
\begin{equation*}
A_{2} \boldsymbol{v}=u A_{1} \boldsymbol{v}, \quad B_{2} \boldsymbol{v}=u B_{1} \boldsymbol{v} \tag{4}
\end{equation*}
$$

The generalized eigenvalues of the left problem are precisely the even $u_{k}$-values, $\left\{u_{2 k}\right\}_{k=1}^{n}$, and the generalized eigenvalues of the right problem are the odd ones, $\left\{u_{2 k-1}\right\}_{k=1}^{n}$.

The forthcoming section is devoted to a complete justification of this algorithm. We recall that these inversion routines have been shown to be numerically efficient in the paper [7].

## 2 Analysis of the algorithm

We begin by stating two classical results of prime importance for the analysis.
Let $\mathbb{L}^{n} \subset \mathbb{R}^{n \times n}$ be the set of lower triangular $n \times n$ real Toeplitz matrices. We define the diagonal scaling matrix and the mapping $\mathcal{T}: \mathbb{R}^{n} \rightarrow \mathbb{L}^{n}$ as

$$
\left.\Lambda=\left(\begin{array}{llll}
0 & & & \\
& 1 & & \\
& & \ddots & \\
& & & n-1
\end{array}\right) ; \quad \mathcal{T}(\boldsymbol{x}):=\left(\begin{array}{ccc}
x_{1} & & \\
x_{2} & x_{1} & \\
\vdots & \ddots & \ddots \\
x_{n} & \ldots & x_{2}
\end{array}\right), x_{1}\right), \quad \boldsymbol{x}=\left(\begin{array}{c}
x_{1} \\
x_{2} \\
\vdots \\
x_{n}
\end{array}\right)
$$

The mapping $\mathcal{T}$ has the following properties, see e.g. [1]:
Proposition 1 Lower triangular Toeplitz matrices commute and $\mathbb{L}^{n}$ is closed under matrix multiplication and (when the inverse exists) inversion,

$$
\begin{equation*}
\mathcal{T}(\boldsymbol{x}) \mathcal{T}(\boldsymbol{y})=\mathcal{T}(\boldsymbol{y}) \mathcal{T}(\boldsymbol{x}) \in \mathbb{L}^{n}, \quad \mathcal{T}(\boldsymbol{x})^{-1} \in \mathbb{L}^{n} \tag{5}
\end{equation*}
$$

Moreover, $\mathcal{T}$ is linear and

$$
\begin{equation*}
\mathcal{T}(\boldsymbol{x}) \boldsymbol{y}=\mathcal{T}(\boldsymbol{y}) \boldsymbol{x}, \quad \mathcal{T}(\boldsymbol{x}) \mathcal{T}(\boldsymbol{y})=\mathcal{T}(\mathcal{T}(\boldsymbol{x}) \boldsymbol{y}) \tag{6}
\end{equation*}
$$

The $\Lambda$ matrix has the property

$$
\begin{equation*}
\mathcal{T}(\Lambda \boldsymbol{x})=\Lambda \mathcal{T}(\boldsymbol{x})-\mathcal{T}(\boldsymbol{x}) \Lambda \tag{7}
\end{equation*}
$$

Another result that we rely heavily upon is the classical Newton relations, see e.g. [8]:
Proposition 2 (Newton's relations) Let $P$ be the $n$-degree polynomial,

$$
P(x)=c_{0}+c_{1} x+\cdots+c_{n} x^{n}=: c_{n}\left(x-x_{1}\right) \cdots\left(x-x_{n}\right) .
$$

Set $S_{0}=n$ and define $S_{k}$ for $k>0$ as the sum of the roots of $P$ taken to the power $k, S_{k}=\sum_{j=1}^{n} x_{j}^{k}$. Then, the $n+1$ following relations hold:

$$
\begin{equation*}
c_{k} S_{0}+c_{k+1} S_{1}+\cdots+c_{n} S_{n-k}=k c_{k}, \quad k=0, \ldots, n \tag{8}
\end{equation*}
$$

### 2.1 Reformulation of the simplified algorithm

We want to write the equation $A \boldsymbol{a}=\boldsymbol{m}$ using the mapping $\mathcal{T}$ : hence we augment the $\boldsymbol{m}$ and $\boldsymbol{a}$-vectors with a zero and one element, respectively, to get $\tilde{\boldsymbol{m}}=(0, \boldsymbol{m})^{T}$ and $\tilde{\boldsymbol{a}}=(1, \boldsymbol{a})^{T}$, both in $\mathbb{R}^{K+1}$. We observe that the $A$-matrix in (2) is the lower right $K \times K$ block of $\Lambda-\mathcal{T}(\tilde{\boldsymbol{m}})$. Therefore,

$$
(\Lambda-\mathcal{T}(\tilde{\boldsymbol{m}})) \tilde{\boldsymbol{a}}=-\tilde{\boldsymbol{m}}+\binom{0}{A \boldsymbol{a}}=\binom{0}{A \boldsymbol{a}-\boldsymbol{m}}
$$

Thus the equation $A \boldsymbol{a}=\boldsymbol{m}$ in (3) is equivalent to

$$
\begin{equation*}
\mathcal{T}(\tilde{\boldsymbol{m}}) \tilde{\boldsymbol{a}}=\Lambda \tilde{\boldsymbol{a}} \tag{9}
\end{equation*}
$$

By the same argument, $B \boldsymbol{b}=-\boldsymbol{m}$ in (3) is equivalent to

$$
\begin{equation*}
\mathcal{T}(\tilde{\boldsymbol{m}}) \tilde{\boldsymbol{b}}=-\Lambda \tilde{\boldsymbol{b}} \tag{10}
\end{equation*}
$$

with $\tilde{\boldsymbol{b}}=(1, \boldsymbol{b})^{T}$.
We can then directly also show that $\mathcal{T}(\tilde{\boldsymbol{a}})$ and $\mathcal{T}(\tilde{\boldsymbol{b}})$ are in fact each other's inverses.
Lemma 1 When $\tilde{\boldsymbol{a}}$ are $\tilde{\boldsymbol{b}}$ given by (9, 10) then $\mathcal{T}(\tilde{\boldsymbol{a}}) \mathcal{T}(\tilde{\boldsymbol{b}})=I$.
Proof: By Proposition 1 and (9, 10),
$\Lambda \mathcal{T}(\tilde{\boldsymbol{a}}) \tilde{\boldsymbol{b}}=\mathcal{T}(\Lambda \tilde{\boldsymbol{a}}) \tilde{\boldsymbol{b}}+\mathcal{T}(\tilde{\boldsymbol{a}}) \Lambda \tilde{\boldsymbol{b}}=\mathcal{T}(\mathcal{T}(\tilde{\boldsymbol{m}}) \tilde{\boldsymbol{a}}) \tilde{\boldsymbol{b}}+\mathcal{T}(\tilde{\boldsymbol{a}}) \Lambda \tilde{\boldsymbol{b}}=\mathcal{T}(\tilde{\boldsymbol{a}})[\mathcal{T}(\tilde{\boldsymbol{m}}) \tilde{\boldsymbol{b}}+\Lambda \tilde{\boldsymbol{b}}]=0$.
Thus $\mathcal{T}(\tilde{\boldsymbol{a}}) \tilde{\boldsymbol{b}}$ lies in the nullspace of $\Lambda$ which is spanned by the vector $\mathbf{1}:=(1,0, \ldots, 0)^{T}$. Moreover, $(\mathcal{T}(\tilde{\boldsymbol{a}}) \tilde{\boldsymbol{b}})_{1}=1$ since the first elements of $\tilde{\boldsymbol{a}}$ and $\tilde{\boldsymbol{b}}$ are both one and we must in fact have $\mathcal{T}(\tilde{\boldsymbol{a}}) \tilde{\boldsymbol{b}}=\mathbf{1}$. The lemma then follows from (6) and the definition of $\mathcal{T}$.

To understand the workings of the algorithm we need to introduce some new quantities and determine how they relate to $\tilde{\boldsymbol{a}}, \tilde{\boldsymbol{b}}$ and $\tilde{\boldsymbol{m}}$.

Let us start with some notation: we set $x_{j}=u_{2 j}$ and $y_{j}=u_{2 j-1}$ for $j=1, \ldots, n$. Furthermore, we introduce the sums

$$
\begin{equation*}
X_{k}=\sum_{j=1}^{n} x_{j}^{k}, \quad Y_{k}=\sum_{j=1}^{n} y_{j}^{k}, \quad k=1,2, \ldots, K=2 n, \tag{11}
\end{equation*}
$$

and define $X_{0}=Y_{0}=K$. In the even case, it then holds that

$$
m_{k}=\sum_{j=1}^{n} x_{j}^{k}-\sum_{j=1}^{n} y_{j}^{k}=X_{k}-Y_{k}, \quad k=1, \ldots, K=2 n
$$

We also define the two polynomials

$$
\begin{equation*}
p(x)=\left(x-x_{1}\right) \cdots\left(x-x_{n}\right)=: c_{0}+c_{1} x+\cdots+c_{n-1} x^{n-1}+c_{n} x^{n} \tag{12}
\end{equation*}
$$

and

$$
\begin{equation*}
q(x)=\left(x-y_{1}\right) \cdots\left(x-y_{n}\right)=: d_{0}+d_{1} x+\cdots+d_{n-1} x^{n-1}+d_{n} x^{n} . \tag{13}
\end{equation*}
$$

We note here that by construction $c_{n}=d_{n}=1$.
By applying (8) to $x^{n} p(x)$ with $k=0, \ldots, K$ we get

$$
\left(\begin{array}{cccc}
X_{0} & & & \\
X_{1} & X_{0} & & \\
\vdots & \ddots & \ddots & \\
X_{K} & \ldots & X_{1} & X_{0}
\end{array}\right)\left(\begin{array}{c}
c_{n} \\
c_{n-1} \\
\vdots \\
c_{0} \\
0 \\
\vdots \\
0
\end{array}\right)=\left(\begin{array}{c}
K c_{n} \\
(K-1) c_{n-1} \\
\vdots \\
(K-n) c_{0} \\
0 \\
\vdots \\
0
\end{array}\right)
$$

The analogous system of equations holds also for $Y_{k}$ and $d_{k}$. We introduce now some shorthand notation to write these equations in a concise form. First we set $\overline{\boldsymbol{c}}=$ $\left(c_{n}, \ldots, c_{0}\right)^{T} \in \mathbb{R}^{n+1}$ and $\overline{\boldsymbol{d}}=\left(d_{n}, \ldots, d_{0}\right)^{T} \in \mathbb{R}^{n+1}$. We then construct the larger vectors, padded with zeros: $\boldsymbol{c}=(\overline{\boldsymbol{c}}, \mathbf{0})^{T}$ and $\boldsymbol{d}=(\overline{\boldsymbol{d}}, \mathbf{0})^{T}$, both in $\mathbb{R}^{K+1}$. Finally, we let $\boldsymbol{X}=\left(X_{0}, \ldots, X_{K}\right)^{T}$ and $\boldsymbol{Y}=\left(Y_{0}, \ldots, Y_{K}\right)^{T}$. Using $\mathcal{T}$ and $\Lambda$ we can state the systems of equations above as follows:

$$
\begin{equation*}
\mathcal{T}(\boldsymbol{X}) \boldsymbol{c}=(K I-\Lambda) \boldsymbol{c}, \quad \mathcal{T}(\boldsymbol{Y}) \boldsymbol{d}=(K I-\Lambda) \boldsymbol{d} \tag{14}
\end{equation*}
$$

We also clearly have $\tilde{\boldsymbol{m}}=\boldsymbol{X}-\boldsymbol{Y}$.
Before we can relate $\boldsymbol{c}$ and $\boldsymbol{d}$ with $\tilde{\boldsymbol{a}}$ we need the following lemma:
Lemma 2 Let $\boldsymbol{f}: \mathbb{R}^{n} \rightarrow \mathbb{R}^{n}$ be defined by $\boldsymbol{f}(\boldsymbol{x}):=\mathcal{T}(\boldsymbol{x})^{-1} \Lambda \boldsymbol{x}$ for $\boldsymbol{x}$ with a non-zero first element. Then $\boldsymbol{f}\left(\boldsymbol{x}_{1}\right)=\boldsymbol{f}\left(\boldsymbol{x}_{2}\right)$ implies that $\boldsymbol{x}_{1}=\alpha \boldsymbol{x}_{2}$ for some non-zero $\alpha \in \mathbb{R}$.

Proof: Suppose $\boldsymbol{f}(\boldsymbol{x})=\boldsymbol{y}$. Then $\mathcal{T}(\boldsymbol{x}) \boldsymbol{y}=\Lambda \boldsymbol{x}$ and by Proposition $1,(\mathcal{T}(\boldsymbol{y})-\Lambda) \boldsymbol{x}=0$. Hence $\boldsymbol{f}\left(\boldsymbol{x}_{1}\right)=\boldsymbol{f}\left(\boldsymbol{x}_{2}\right)$ implies that $\boldsymbol{x}_{1}$ and $\boldsymbol{x}_{2}$ both lie in the nullspace of $\mathcal{T}(\boldsymbol{y})-\Lambda$. Since the top left element of $\Lambda$ is zero, it follows that the first element of $\boldsymbol{y}$ is zero and therefore the diagonal of $\mathcal{T}(\boldsymbol{y})$ is zero. Consequently, the nullspace of $\mathcal{T}(\boldsymbol{y})-\Lambda$ has the same dimension as that of $\Lambda$, which is one.

We can now merge together and express the general structure from $(12,13)$ and $(9$, 10) in the most concise way.

Lemma 3 Suppose $\boldsymbol{c}$, $\boldsymbol{d}$ are defined by $(12,13)$ and $\tilde{\boldsymbol{a}}, \tilde{\boldsymbol{b}}$ by $(9,10)$. Then

$$
\mathcal{T}(\tilde{\boldsymbol{a}}) \boldsymbol{c}=\boldsymbol{d}, \quad \mathcal{T}(\tilde{\boldsymbol{b}}) \boldsymbol{d}=\boldsymbol{c}
$$

Proof: We only need to prove the left equality. The right one follows immediately from Lemma 1. Let $\boldsymbol{v}=\mathcal{T}(\tilde{\boldsymbol{a}}) \boldsymbol{c}=\mathcal{T}(\boldsymbol{c}) \tilde{\boldsymbol{a}}$. We want to show that $\boldsymbol{v}=\boldsymbol{d}$. We note first that by (14)
$\mathcal{T}(\boldsymbol{X}) \boldsymbol{c}=(K I-\Lambda) \boldsymbol{c} \quad \Rightarrow \quad \mathcal{T}(\boldsymbol{c}) \boldsymbol{X}=(K I-\Lambda) \boldsymbol{c} \quad \Rightarrow \quad \boldsymbol{X}=K \mathcal{T}(\boldsymbol{c})^{-1} \boldsymbol{c}-\mathcal{T}(\boldsymbol{c})^{-1} \Lambda \boldsymbol{c}$,
where $\mathcal{T}(\boldsymbol{c})$ is invertible since $c_{n}=1$. Moreover, it is clear that $\mathcal{T}(\boldsymbol{y}) \mathbf{1}=\boldsymbol{y}$ for all $\boldsymbol{y}$. Hence, $\boldsymbol{X}=K \mathbf{1}-\mathcal{T}(\boldsymbol{c})^{-1} \Lambda \boldsymbol{c}$. In the same way we also obtain $\boldsymbol{Y}=K \mathbf{1}-\mathcal{T}(\boldsymbol{d})^{-1} \Lambda \boldsymbol{d}$. Then,

$$
\begin{aligned}
\mathcal{T}(\tilde{\boldsymbol{m}}) \tilde{\boldsymbol{a}} & =\mathcal{T}(\tilde{\boldsymbol{a}}) \tilde{\boldsymbol{m}}=\mathcal{T}(\tilde{\boldsymbol{a}})(\boldsymbol{X}-\boldsymbol{Y}) \\
& =-\mathcal{T}(\tilde{\boldsymbol{a}}) \mathcal{T}(\boldsymbol{c})^{-1} \Lambda \boldsymbol{c}+\mathcal{T}(\tilde{\boldsymbol{a}}) \mathcal{T}(\boldsymbol{d})^{-1} \Lambda \boldsymbol{d}
\end{aligned}
$$

We now note that by Proposition 1,

$$
\begin{aligned}
\mathcal{T}(\tilde{\boldsymbol{a}}) \mathcal{T}(\boldsymbol{c})^{-1} \Lambda \boldsymbol{c} & =\mathcal{T}(\boldsymbol{c})^{-1} \mathcal{T}(\tilde{\boldsymbol{a}}) \Lambda \boldsymbol{c}=\mathcal{T}(\boldsymbol{c})^{-1} \mathcal{T}(\Lambda \boldsymbol{c}) \tilde{\boldsymbol{a}}=\mathcal{T}(\boldsymbol{c})^{-1} \Lambda \mathcal{T}(\boldsymbol{c}) \tilde{\boldsymbol{a}}-\Lambda \tilde{\boldsymbol{a}} \\
& =\mathcal{T}(\boldsymbol{c})^{-1} \Lambda \boldsymbol{v}-\Lambda \tilde{\boldsymbol{a}}
\end{aligned}
$$

Since also, $\mathcal{T}(\boldsymbol{c}) \mathcal{T}(\tilde{\boldsymbol{a}})=\mathcal{T}(\boldsymbol{v})$ we get

$$
\begin{aligned}
\mathcal{T}(\tilde{\boldsymbol{m}}) \tilde{\boldsymbol{a}} & =-\mathcal{T}(\boldsymbol{c})^{-1} \Lambda \boldsymbol{v}+\Lambda \tilde{\boldsymbol{a}}+\mathcal{T}(\tilde{\boldsymbol{a}}) \mathcal{T}(\boldsymbol{d})^{-1} \Lambda \boldsymbol{d} \\
& =\mathcal{T}(\tilde{\boldsymbol{a}})\left[\mathcal{T}(\boldsymbol{d})^{-1} \Lambda \boldsymbol{d}-\mathcal{T}(\boldsymbol{v})^{-1} \Lambda \boldsymbol{v}\right]+\Lambda \tilde{\boldsymbol{a}}
\end{aligned}
$$

Consequently, by (9), $\mathcal{T}(\boldsymbol{d})^{-1} \Lambda \boldsymbol{d}=\mathcal{T}(\boldsymbol{v})^{-1} \Lambda \boldsymbol{v}$ and by Lemma 2, $\boldsymbol{v}=\alpha \boldsymbol{d}$; for some $\alpha \in \mathbb{R}$. But for the first element in $\boldsymbol{v}$ we then have $v_{1}=c_{n}=\alpha d_{n}$ and we get $\alpha=1$ since $c_{n}=d_{n}=1$.

Finally, we also establish the following lemma.
Lemma 4 Let $V$ and $W$ be the Vandermonde matrices corresponding to $\left\{x_{j}\right\}_{j=0}^{n}$ and $\left\{y_{j}\right\}_{j=0}^{n}$ respectively, with $x_{0}=y_{0}=0$. Then

$$
V^{T} R \mathcal{T}(\overline{\boldsymbol{c}}) V=\operatorname{diag}\left(\left\{\tilde{p}^{\prime}\left(x_{k}\right)\right\}_{k=0}^{n}\right), \quad W^{T} R \mathcal{T}(\overline{\boldsymbol{d}}) W=\operatorname{diag}\left(\left\{\tilde{q}^{\prime}\left(x_{k}\right)\right\}_{k=0}^{n}\right)
$$

where $\tilde{p}(x):=x p(x), \tilde{q}(x):=x q(x)$ and $R=\left\{\delta_{n+2-i-j}\right\} \in \mathbb{R}^{n+1 \times n+1}$ is the reversion matrix.

Proof: We have

$$
\left(V^{T} R \mathcal{T}(\overline{\boldsymbol{c}}) V\right)_{i j}=\sum_{r=0}^{n} \sum_{\ell=0}^{r} c_{r} x_{i-1}^{\ell} x_{j-1}^{r-\ell}= \begin{cases}\frac{\tilde{p}\left(x_{i-1}\right)-\tilde{p}\left(x_{j-1}\right)}{x_{i-1}-x_{j-1}}, & i \neq j, \\ \tilde{p}^{\prime}\left(x_{i-1}\right), & i=j,\end{cases}
$$

showing the left equality since $\left\{x_{j}\right\}_{j=0}^{n}$ are the roots of $\tilde{p}(x)$. The right equality follows in the same way.

### 2.3 Conclusion

We can now conclude and show that the unknown values $u_{j}$ in (1) are indeed the generalized eigenvalues of (4).

Theorem 1 Suppose $K=2 n$; let $\boldsymbol{a}, \boldsymbol{b}$ be defined by (3). If all values $\left\{x_{j}\right\} \cup\left\{y_{j}\right\}$ are distinct, then $\left\{x_{j}\right\},\left\{y_{j}\right\}$ are the generalized eigenvalues of (4).

Proof: Let $\tilde{\boldsymbol{a}}, \tilde{\boldsymbol{b}}$ be defined by (9,10), which is equivalent to (3). Also define $\boldsymbol{c}$ and $\boldsymbol{d}$ as before by $(12,13)$. By Lemma 3 we have $\mathcal{T}(\tilde{\boldsymbol{a}}) \boldsymbol{c}=\boldsymbol{d} \in \mathbb{R}^{2 n+1}$, i.e.

Clearly, the lower left block of the matrix multiplied by $\overline{\boldsymbol{c}}$ is zero, i.e. $\sum_{i=0}^{n} c_{i} a_{i+k}=0$ for $k=1, \ldots, n$. Now, let $v_{i}$ be the coefficients of the polynomial $v(x):=p(x) /\left(x-x_{j}\right)$ for some fixed $j$. Hence, by the special structure of (12),

$$
c_{0}+c_{1} x+\ldots+c_{n} x^{n}=:\left(v_{1}+v_{2} x+\ldots+v_{n} x^{n-1}\right)\left(x-x_{j}\right),
$$

and, for $i=0, \ldots, n$,

$$
c_{i}=\left\{\begin{array}{lr}
-x_{j} v_{i+1}, & i=0, \\
v_{i}-x_{j} v_{i+1}, 1 \leq i \leq n-1, \\
v_{i}, & i=n
\end{array}\right.
$$

Thus we deduce,

$$
0=-x_{j} v_{1} a_{k}+\sum_{i=1}^{n-1}\left(v_{i}-x_{j} v_{i+1}\right) a_{i+k}+v_{n} a_{n+k}=\sum_{i=1}^{n} v_{i} a_{i+k}-x_{j} \sum_{i=1}^{n} v_{i} a_{i+k-1}
$$

which is the componentwise statement of $A_{2} \boldsymbol{v}=x_{j} A_{1} \boldsymbol{v}$. It remains to show that the rightmost sum is non-zero for at least some $k$, so that $x_{j}$ is indeed a well-defined generalized eigenvalue. Let $\overline{\boldsymbol{a}}=\left(1, a_{1}, \ldots, a_{n}\right)^{T}$ and $\overline{\boldsymbol{v}}=\left(0, v_{1}, \ldots, v_{n}\right)^{T}$. Then (15) gives $\mathcal{T}(\overline{\boldsymbol{a}}) \overline{\boldsymbol{c}}=\overline{\boldsymbol{d}}$ and, using Lemma 4 while taking $k=1$ we have the sum

$$
\sum_{i=1}^{n} v_{i} a_{i}=\overline{\boldsymbol{a}}^{T} \overline{\boldsymbol{v}}=\left(\mathcal{T}(\overline{\boldsymbol{c}})^{-1} \overline{\boldsymbol{d}}\right)^{T} \overline{\boldsymbol{v}}=\left(V^{T} R \overline{\boldsymbol{d}}\right)^{T} \operatorname{diag}\left(\left\{\tilde{p}^{\prime}\left(x_{k}\right)^{-1}\right\}\right) V^{T} \overline{\boldsymbol{v}}=q\left(x_{j}\right),
$$

since $V^{T} R \overline{\boldsymbol{d}}=\left\{q\left(x_{k}\right)\right\}$ and $V^{T} \overline{\boldsymbol{v}}=\left\{x_{k} v\left(x_{k}\right)\right\}=\left\{\delta_{k-j} \tilde{p}^{\prime}\left(x_{k}\right)\right\}$. Hence, the sum is non-zero because $x_{j} \neq y_{i}$ for all $i, j>0$. The same argument can be used for any $j$,
which proves the theorem for $\left\{x_{j}\right\}$. The proof for $\left\{y_{j}\right\}$ is identical upon exchanging the roles of $\boldsymbol{c}, \tilde{\boldsymbol{a}}$ and $\boldsymbol{d}, \tilde{\boldsymbol{b}}$. This leads to (4).

## References

[1] D. Bini, Polynomial and matrix computations. I. Fundamental Algorithms, Birkhaüser Verlag.
[2] Y. Brenier, Équations de moment et conditions d'entropie pour des modèles cinétiques (French) [Moment equations and entropy conditions for kinetic models] Séminaire sur les Équations aux Dérivées Partielles, 1994-1995, Exp. No. XXII, 11 pp., École Polytech., Palaiseau, 1995.
[3] Y. Brenier and L. Corrias, A kinetic formulation for multibranch entropy solutions of scalar conservation laws, Ann. I.H.P. Nonlinear Anal. 15 (1998), 169-190.
[4] D. Czarkowski, D. \& G. Chudnovsky, I.W. Selesnick, Solving the optimal $P W M$ problem for single-phase inverters, IEEE Trans. Circ. Syst. I 49 (2002), 465-475.
[5] P. Diaconis, D. Friedman, The Markov moment problem and de Finetti's theorem, to appear in Math. Zeitschrift
[6] L. Gosse, Using K-branch entropy solutions for multivalued geometric optics computations, J. Comp. Phys. 180 (2002) 155-182.
[7] L. Gosse \& O. Runborg, Finite moment problems and applications to multiphase computations in geometric optics, Comm. Math. Sci. 3 (2005) 373-392.
[8] I.N. Hepstein, Topics in Algebra, Ginn Waltham MA (1964) page 208.
[9] V.I. Korobov, G. M. Sklyar, Time-optimality and the power moment problem (Russian) Mat. Sb. (N.S.) 134 (176) (1987) 186-206, 287; translation in Math. USSR-Sb. 62 (1989), no. 1, 185-206
[10] M.G. Krein, A.A. Nudel'man, The Markov moment problem and extremal problems, A.M.S. translations (1977).
[11] A.S. Lewis, Superresolution in the Markov moment problem, J. Math. Anal. Appl. 197 (1996) 774-780.
[12] O. Runborg, Some new results in multiphase geometrical optics, Math. Mod. Numer. Anal. 34 (2000), 1203 - 1231.
[13] G.M. Sklyar, L.V. Fardigola, The Markov power moment problem in problems of controllability and frequency extinguishing for the wave equation on a half-axis, J. Math. Anal. Appl. 276 (2002) 109-134.
[14] G. Talenti, Recovering a function from a finite number of moments, Inverse Problems 3 (1987) 501-517.


[^0]:    * Support by the Wolfgang Pauli Institute (Wien) in the frame of the "Nanoscience" Programme and the European network HYKE, funded by the EC as contract HPRN-CT-2002-00282 is acknowledged.
    * Corresponding Author

    Email addresses: 1.gosse@ba.iac.cnr.it (Laurent Gosse), olofr@nada.kth.se (Olof Runborg).

