A MOMENT APPROACH TO ANALYZE ZEROS
OF TRIANGULAR POLYNOMIAL SETS

JEAN B. LASSERRE

Abstract. Let \( I = \langle g_1, \ldots, g_n \rangle \) be a zero-dimensional ideal of \( \mathbb{R}[x_1, \ldots, x_n] \) such that its associated set \( G \) of polynomial equations \( g_i(x) = 0 \) for all \( i = 1, \ldots, n \) is in triangular form. By introducing multivariate Newton sums we provide a numerical characterization of polynomials in \( \sqrt{I} \). We also provide a necessary and sufficient (numerical) condition for all the zeros of \( G \) to be in a given set \( K \subset \mathbb{C}^n \), without explicitly computing the zeros. In addition, we also provide a necessary and sufficient condition on the coefficients of the \( g_i \)'s for \( G \) to have (a) only real zeros, (b) to have only real zeros, all contained in a given semi-algebraic set \( K \subset \mathbb{R}^n \). In the proof technique, we use a deep result of Curto and Fialkow (2000) on the \( K \)-moment problem, and the conditions we provide are given in terms of positive definiteness of some related moment and localizing matrices depending on the \( g_i \)'s via the Newton sums of \( G \). In addition, the number of distinct real zeros is shown to be the maximal rank of a related moment matrix.

1. Introduction

In this paper we consider an ideal \( I := \langle g_1, \ldots, g_n \rangle \subset \mathbb{R}[x_1, \ldots, x_n] \) generated by the real-valued polynomials \( g_i \in \mathbb{R}[x_1, \ldots, x_n] \). Let us call \( G := \{g_1, \ldots, g_n\} \) a polynomial set and let a term ordering of monomials with \( x_1 < x_2 < \cdots < x_n \) be given.

We assume that the system of polynomials equations \( \{g_i(x) = 0, i = 1, \ldots, n\} \) is in the following triangular form:

\[
g_i(x) = p_i(x_1, \ldots, x_{i-1})x_i^r + h_i(x_1, \ldots, x_i), \quad i = 1, \ldots, n,
\]

by which we mean the following:

(i) \( x_i \) is the main variable and \( p_i(x_1, \ldots, x_{i-1})x_i^r \) is the leading term of \( g_i \).

(ii) for every \( i = 2, \ldots, n \), every zero in \( \mathbb{C}^n \) of the polynomial system \( G_{i-1} := \{g_1, \ldots, g_{i-1}\} \) is not a zero of the leading coefficient \( \text{in}(g_i) := p_i(x_1, \ldots, x_{i-1}) \) of \( g_i \).

The set \( G \) is called a triangular set. From (i)-(ii), it follows that \( I \) is a zero-dimensional ideal. Conversely, any zero-dimensional ideal can be represented by a finite union of specific triangular sets (see e.g. Aubry et al. [1], Lazard [10]). For various definitions (and results) related to triangular sets (e.g. due to Kalkbrener, Lazard, Wu) the interested reader is referred to Lazard [10], Wang [7] and the many references therein; see also Aubry and Maza [2] for a comparison of symbolic algorithms related to triangular sets.

Received by the editors April 10, 2002.

2000 Mathematics Subject Classification. Primary 12D10, 26C10, 30E05.

Key words and phrases. System of polynomial equations, triangular sets, moment problem.
For instance, there are symbolic algorithms that, given \( I \) as input, generate a finite set of triangular systems in the specific form \( g_i(x) = x_i - f_i(x_1) \) for all \( i = 2, \ldots, n \). Triangular sets in the latter form are particularly interesting to develop efficient symbolic algorithms for counting and computing real zeros of polynomials sets (see e.g. Becker and Wörmann [5] and the recent work of Rouillier [11]).

The goal of this paper is to show that a triangular polynomial set \( G \) as in \((1.1)\) also has several advantages from a numerical point of view. Indeed, it also permits us to define multivariate Newton sums, the multivariate analogue of Newton sums for univariate polynomials (which can be used for counting real zeros as in Gantmacher [8, Chap. 15, p. 200]). We shall see that indeed the same is true for multivariate polynomials systems in triangular form \((1.1)\). Namely, we show that:

(a) With a triangular system \( G \) as in \((1.1)\) we may associate real moment matrices \( M_p(y) \) depending on the (known) multivariate Newton sums of \( G \) (to be defined later) and on an unknown vector \( y \). The condition \( M_p(y) \geq 0 \) for some specific \( p = r_0 \) (meaning \( M_p(y) \) positive semidefinite) defines a unique solution \( y^* \), the vector of all moments (up to order \( 2p \)) of a probability measure \( \mu^* \) supported on all the zeros of \( G \) in \( \mathbb{C}^n \). As a consequence, a polynomial of degree less than \( 2p \) is in \( \sqrt{T} \) if and only if its vector of coefficients \( f \) satisfies the linear system of equations \( M_p(y^*)f = 0 \).

(b) Moreover, given a set

\[ K := \{ z \in \mathbb{C}^n \mid w_j(z_1, \ldots, z_n, \overline{z_1}, \ldots, \overline{z_n}) \geq 0, \ j = 1, \ldots, m \} \subset \mathbb{C}^n, \]

defined by some polynomials \( \{ w_j \} \) in \( \mathbb{C}[z, \overline{z}] \) (which can be viewed as a semi-algebraic set in \( \mathbb{R}^{2n} \)), one may also check whether the zero set of \( G \) is contained in \( K \), by solving a convex semidefinite program for which efficient software packages are now available. The necessary and sufficient conditions state that the system of LMI (Linear Matrix Inequalities)

\[ M_{r_0}(y) \geq 0, \quad M_{r_0}(w_i y) \geq 0, \quad i = 1, \ldots, m, \]

for some appropriate moment matrix \( M_{r_0}(y) \) and localizing matrices \( M_{r_0}(w_i y) \) (depending on the Newton sums of \( G \)) must have a solution, which is then unique, i.e. \( y = y^* \) with \( y^* \) as in (a). In fact, it suffices to solve the single inequality \( M_p(y) \geq 0 \), which yields the unique solution \( y^* \), and then check afterwards whether \( M_{r_0}(w_i y^*) \geq 0 \), for all \( i = 1, \ldots, m \). For an introduction to semidefinite programming, the interested reader is referred to Vandenbergh and Boyd [13].

(c) As a consequence, we also provide a necessary and sufficient condition (only in terms of the Newton sums of \( G \)) for all the zeros of \( G \) to be real, and also for these real zeros to be in a given semi-algebraic set

\[ K_1 := \{ x \in \mathbb{R}^n \mid u_i(x_1, \ldots, x_n) \geq 0, \quad i = 1, \ldots, m \} \subset \mathbb{R}^n, \]

for some polynomials \( \{ u_i \} \) in \( \mathbb{R}[x_1, \ldots, x_n] \). In this case, the moment matrix is completely known and depends only on the Newton sums of \( G \). This latter result extends to the multivariate case a previous result of the same vein by the author for the univariate case [9].

(d) Finally, it is shown that the number of (distinct) real zeros of \( G \) is the maximal rank of a positive semidefinite moment matrix \( M_{r_0}(y) \), that is, a \( y \) that maximizes this rank is the vector of moments of a probability measure with support on all the real zeros of \( G \). This also provides a characterization of the ideal \( I(V_{\mathbb{R}}(I)) \) in terms of moment matrices.
The basic technique that we use relies on a deep result of Curto and Fialkow [6] for the K-moment problem.

2. Notation, definitions and preliminary results

Some of the material in this section is from Curto and Fialkow [6]. Let \( P_r \) be the space of polynomials in \( \mathbb{C}[z_1, \ldots, z_n, \overline{z}_1, \ldots, \overline{z}_n] \) (for short \( \mathbb{C}[z, \overline{z}] \)) of degree at most \( r \in \mathbb{N} \). Now, following notation as in Curto and Fialkow [6], a polynomial \( \theta \in \mathbb{C}[z, \overline{z}] \) is written

\[
\theta(z, \overline{z}) = \sum_{\alpha,\beta} \theta_{\alpha\beta} z^\alpha \overline{z}^\beta = \sum_{\alpha,\beta} \theta_{\alpha\beta} z_{1}^{\alpha_1} \cdots z_{n}^{\alpha_n} \overline{z}_{1}^{\beta_1} \cdots \overline{z}_{n}^{\beta_n},
\]

in the usual basis of monomials (e.g. ordered lexicographically)

\[
(2.1) \quad 1, z_1, \ldots, z_n, \overline{z}_1, \ldots, \overline{z}_n, z_1^2, z_1z_2, \ldots.
\]

We here identify \( \theta \in \mathbb{C}[z, \overline{z}] \) with its vector of coefficients \( \theta := \{ \theta_{\alpha\beta} \} \) in the basis (2.1).

Given an infinite sequence \( \{ y_{\alpha\beta} \} \) indexed in the basis (2.1), we also define the linear functional on \( \mathbb{C}[z, \overline{z}] \)

\[
\theta \mapsto \Lambda(\theta) := \sum_{\alpha,\beta} \theta_{\alpha\beta} y_{\alpha\beta} = \sum_{\alpha,\beta} \theta_{\alpha\beta} y_{\alpha_1 \ldots \alpha_n, \beta_1 \ldots \beta_n}.
\]

2.1. The moment matrix. Given \( p \in \mathbb{N} \) and an infinite sequence \( \{ y_{\alpha\beta} \} \), let \( M_p(y) \) be the unique square matrix such that

\[
\langle M_p(y) f, h \rangle = \Lambda(f \overline{h}) \quad \forall f, h \in P_p
\]

(see e.g. Curto and Fialkow [6] p. 3).

To fix ideas, in the two-dimensional case, the moment matrix \( M_1(y) \) is given by

\[
M_1(y) = \begin{bmatrix}
1 & y_{0010} & y_{0001} & y_{1000} & y_{0100} \\
y_{0010} & y_{0100} & y_{1001} & y_{2000} & y_{1100} \\
y_{0001} & y_{0101} & y_{1011} & y_{1110} & y_{2010} \\
y_{0100} & y_{0101} & y_{1101} & y_{1110} & y_{2100} \\
y_{0011} & y_{0111} & y_{1111} & y_{2011} & y_{2110}
\end{bmatrix}.
\]

Thus, the entry of the moment matrix \( M_p(y) \) corresponding to column \( z^\gamma \overline{z}^\eta \) and row \( z^\delta \overline{z}^\zeta \) is \( y_{\alpha+\gamma, \beta+\eta} \), and if \( y \) is the moment vector of a measure \( \mu \) on \( \mathbb{C}^n \), then

\[
\langle M_p(y) f, f \rangle = \Lambda(|f|^2) = \int |f|^2 \, d\mu \geq 0 \quad \forall f \in P_p,
\]

which shows that \( M_p(y) \) is positive semidefinite (denoted \( M_p(y) \succeq 0 \)).

2.2. Localizing matrices. Let \( \{ y_{\alpha\beta} \} \) be an infinite sequence and let \( \theta \in \mathbb{C}[z, \overline{z}] \). Define the localizing matrix \( M_p(\theta y) \) to be the unique square matrix such that

\[
\langle M_p(\theta y) f, g \rangle = \Lambda(\theta fg) \quad \forall f, g \in P_p.
\]

Thus, if \( \theta(z, \overline{z}) = \sum_{\alpha,\beta} \theta_{\alpha\beta} z^\alpha \overline{z}^\beta \) and \( M_p(y)(i, j) = y_{\gamma \eta} \), then

\[
M_p(\theta y)(i, j) = \sum_{\alpha,\beta} \theta_{\alpha\beta} y_{\alpha+\gamma, \beta+\eta}.
\]
For instance, with $z \mapsto \theta(z, \overline{z}) := 1 - \overline{z}_1 z_1$, $M_1(\theta y)$ reads
\[
\begin{bmatrix}
1 - y_{1010} & y_{0100} - y_{1020} & y_{0010} - y_{1011} & y_{1000} - y_{2010} & y_{1000} - y_{1110} \\
y_{1000} - y_{2010} & y_{1010} - y_{2020} & y_{1001} - y_{2011} & y_{2000} - y_{2021} & y_{1100} - y_{2110} \\
y_{1000} - y_{1110} & y_{1010} - y_{2020} & y_{1001} - y_{2110} & y_{2000} - y_{2021} & y_{2000} - y_{2011} \\
y_{0010} - y_{1020} & y_{0020} - y_{1030} & y_{0011} - y_{1021} & y_{1010} - y_{2020} & y_{1010} - y_{1110} \\
y_{0001} - y_{1011} & y_{0011} - y_{1021} & y_{0002} - y_{1012} & y_{1001} - y_{2011} & y_{1001} - y_{1111}
\end{bmatrix}.
\]

It follows that if $y$ is the moment vector of some measure $\mu$ on $\mathbb{C}^n$, supported on the set $\{z \in \mathbb{C}^n \mid \theta(z, \overline{z}) \geq 0\}$, we then have
\[
\langle M_p(\theta y) f, f \rangle = \Lambda(|f|^2) = \int \theta |f|^2 \, d\mu \geq 0 \quad \forall f \in \mathcal{P}_p,
\]
so that $M_p(\theta y) \succeq 0$.

### 2.3. Multivariate Newton sums.

With $x_1 < x_2 < \cdots < x_n$ and given a fixed term ordering of monomials, consider a triangular polynomial system $\mathbb{G} = \{g_1, \ldots, g_n\}$ as in (1.1), that is,
\[
g_i(x) = p_i(x_1, \ldots, x_{i-1}) x_i^{r_i} + h_i(x_1, \ldots, x_i) = 0 \quad \forall i = 1, \ldots, n
\]
(with $p_i \in \mathbb{R}$), and the $p_i$'s are such that for all $i = 2, 3, \ldots, n$,
\[
g_k(z) = 0 \quad \forall k = 1, \ldots, i-1 \Rightarrow p_i(z) \neq 0.
\]

For each $i = 1, \ldots, n$, $p_i(x_1, \ldots, x_{i-1}) x_i^{r_i}$ is the leading term of $g_i$. In the terminology used in e.g. Wang [7] Definitions 2.1, $\mathbb{G}$ is a triangular set.

In view of the assumption on the $g_i$'s, it follows that $\mathbb{G}$ has exactly $s := \prod_{i=1}^n r_i$ zeros $\{z(i)\}_{i=1}^s$ in $\mathbb{C}^n$ (counting their multiplicity) so that $I = \langle g_1, \ldots, g_n \rangle$ is a zero-dimensional ideal and the affine variety $V(\mathbb{G}(I)) \subset \mathbb{C}^n$ is a finite set of cardinality $s_G \leq s$.

For every $\alpha \in \mathbb{N}^n$ define $s_{\alpha}$ to be the real number
\[
s_{\alpha} := s^{-1} \sum_{i=1}^s z(i)^\alpha = \sum_{i=1}^s z_1^{\alpha_1} z_2^{\alpha_2} \cdots z_n^{\alpha_n} (i)
\]
which we call the (normalized) $\alpha$-Newton sum of $\mathbb{G}$ by analogy with the Newton sums of a univariate polynomial (see e.g. Gantmacher [3] p. 199).

**Remark 2.1.** Note that the Newton sums $s_{\alpha}$ depend on $\mathbb{G}$ and not only on the zeros $\{z(i)\}$ because we take into account the possible multiplicities.

**Proposition 2.2.** Let the $g_i$'s be as in (2.6)-(2.7) and let $s_{\alpha}$ be as in (2.8). Then each $s_{\alpha}$ is a rational fraction in the coefficients of the $g_i$'s and can be computed recursively.

For a proof see [5.1]

**Example 2.3.** Consider the elementary example with $\mathbb{G} := \{x_1^2 + 1, x_1 x_2 + x_2 + 1\}$. Then, $s_{\alpha}$ is just the usual (normalized) $i$-Newton sum of $x_1 \mapsto x_1^2 + 1$. For instance, it follows that $s_{01} = 0$, $s_{02} = 0$. Similarly, $s_{11} = -1/2$, $s_{21} = 0$, $s_{22} = 1/2$, etc.

Interestingly, given a polynomial $t \in \mathbb{R}[x_1, \ldots, x_n]$, Rouillier [11] §3 also defines extended Newton sums of what he calls a multi-ensemble associated with a set of points of $\mathbb{C}^n$. He then uses these extended Newton sums to obtain a certain triangular representation of zero-dimensional ideals.
3. Main result

In this section we assume that we are given a polynomial set $\mathbb{G} := \{g_1, \ldots, g_n\}$ in the triangular form (2.6)-(2.7).

3.1. The associated moment matrix. The idea in this section is to build up the moment matrices (defined in 2.1) associated with a particular measure $\mu^*$ on $\mathbb{C}^n$ whose support is on all the zeros of the polynomial set $\mathbb{G}$. That is, let $\{z(i)\}$ be the collection of $s := \prod_{j=1}^{\infty} r_j$ zeros in $\mathbb{C}^n$ of $\mathbb{G}$ (counting their multiplicity) and let $\mu^*$ to be the probability measure on $\mathbb{C}^n$ defined by

$$\mu^* := s^{-1} \sum_{i=1}^{s} \delta_{z(i)},$$

where $\delta_z$ stands for the Dirac measure at the point $z \in \mathbb{C}^n$.

By definition of $\mu^*$, its moments $\int z^\alpha \, d\mu^*$ are just the normalized $\alpha$-Newton sums (2.8).

Indeed,

$$s_\alpha := \int z^\alpha \, d\mu^* = s^{-1} \sum_{i=1}^{s} z(i)^\alpha.$$ (3.2)

If we write

$$y^*_\alpha \beta := \int z^\alpha \, d\mu^*, \quad \alpha, \beta \in \mathbb{N}^n,$$ (3.3)

we have

$$s_\alpha = y^*_\alpha 0 = y^*_0 \alpha, \quad y^*_\alpha \beta = y^*_\beta \alpha, \quad \alpha, \beta \in \mathbb{N}^n.$$ (3.4)

3.2. Construction of the moment matrix of $\mu^*$. With $\mu^*$ as in (3.1) let $\{s_\alpha, y^*_\alpha \beta\}$ defined in (3.2)-(3.3) be the infinite sequence of all its moments.

We then call $M_p(\mu^*)$ the moment matrix associated with $\mu^*$, that is, in $M_p(y)$ we replace the entries $y_{\alpha \alpha}$ or $y_{\alpha 0}$ by $s_\alpha$ and the other entries $y_{\alpha \beta}$ by $y^*_\alpha \beta$. By Proposition 2.2 the entries $s_\alpha$ are known and rational fractions of the coefficients of the polynomials $g_i$’s. They can be computed numerically or symbolically. On the other hand, moments $y^*_\alpha \beta$ do not have a closed form expression in terms of the coefficients of polynomials $g_i$’s.

Therefore, we next introduce a moment matrix $M_p(\mu^*, y)$ obtained from $M_p(\mu^*)$ by replacing the (unknown) entries $y^*_\alpha \beta$ by variables $y_{\alpha \beta}$ and look for conditions on this matrix $M_p(\mu^*, y)$ to be exactly $M_p(\mu^*)$. For instance, in the two-dimensional case, the moment matrix $M_1(\mu^*, y)$ reads

$$M_1(\mu^*, y) = \begin{bmatrix}
1 & s_{10} & s_{01} & s_{10} & s_{01} \\
10 & y_{101} & y_{1001} & s_{20} & s_{11} \\
01 & y_{011} & y_{0101} & s_{11} & s_{02} \\
10 & s_{20} & s_{11} & y_{1010} & y_{0110} \\
01 & s_{11} & s_{02} & y_{1001} & y_{0101}
\end{bmatrix}$$

(with $s_\alpha = y_{\alpha 0} = y_{\alpha 0}$). Moreover, from the definition of $\mu^*$, we may impose $M_p(\mu^*, y)$ to be symmetric for all $p \in \mathbb{N}$, because $y^*_\alpha \beta = y^*_\beta \alpha$ for all $\alpha, \beta \in \mathbb{N}^n$ (see (3.4)).

As $\mathbb{G}$ is a triangular polynomial system in the form (2.0)-(2.7), $I = \langle g_1, \ldots, g_n \rangle$ is a zero-dimensional ideal. Therefore, let $H := \{h_1, \ldots, h_m\}$ be a reduced Gröbner basis of $I$ with respect to (in short, w.r.t.) the term ordering already defined (e.g.}

License or copyright restrictions may apply to redistribution; see http://www.ams.org/journal-terms-of-use
the lexicographical ordering \( x_1 < x_2 < \cdots < x_n \). As \( I \) is zero dimensional, for every \( i = 1, \ldots, n \), we may label the first \( n \) polynomials \( h_j \) of \( H \) in such a way that \( x_i^{r_i} \) is the leading term of \( h_i \) (see e.g. Adams and Loustaunau [3, Theor. 2.2.7]).

**Proposition 3.1.** Let \( \mathbb{G} \) be the triangular polynomial system in \( \{2.6\} - \{2.7\} \) (with some term ordering), and let \( H = \{h_1, \ldots, h_m\} \) be its reduced Gröbner basis (with \( x_i^{r_i} \) the leading term of \( h_i \) for all \( i = 1, \ldots, n \).

Let \( \mu^* \) be the probability measure defined in (3.1). For every \( \alpha, \beta \in \mathbb{N}^n \) let

\[
y^*_{\alpha\beta} := \int \mathcal{P}^* z^\beta \, d\mu^*.
\]

Then, for every \( \gamma, \eta \in \mathbb{N}^n \), \( y^*_{\gamma\eta} \) is a linear combination of the \( y^*_{\alpha\beta} \)'s with \( \alpha_i, \beta_i < r_i \) for all \( i = 1, \ldots, n \), that is,

\[
y^*_{\gamma\eta} = \sum_{\alpha\beta} u_{\alpha\beta}(\eta, \gamma) y^*_{\alpha\beta}, \quad \alpha_i, \beta_i < r_i \quad \forall i = 1, \ldots, n,
\]

for some scalars \( \{u_{\alpha\beta}(\eta, \gamma)\} \).

**Proof.** Let \( H = \{h_1, \ldots, h_m\} \) be the reduced Gröbner basis of \( I \) w.r.t. the term ordering, with \( x_i^{r_i} \) the leading term of \( h_i \) for all \( i = 1, \ldots, n \leq m \).

For \( \gamma, \eta \in \mathbb{N}^n \), write

\[
z^\gamma = \sum_{i=1}^m q_i(z) h_i(z) + q_\eta(z), \quad \mathcal{P}^\gamma = \sum_{i=1}^m v_i(\mathcal{P}) h_i(\mathcal{P}) + v_\eta(\mathcal{P}),
\]

for some polynomials \( \{q_\eta, q_i\} \) and \( \{v_\gamma, v_i\} \) in \( \mathbb{R}[x_1, \ldots, n] \), that is, \( z^\gamma \) (resp. \( \mathcal{P}^\gamma \)) is reduced to \( q_\eta(z) \) (resp. \( v_\gamma(z) \)) w.r.t. \( H \). Due to the special form of \( H \), it follows that the monomials \( z^\alpha \) of \( q_\eta, v_\gamma \) satisfy \( \alpha_i < r_i \) for all \( i = 1, \ldots, n \). Hence,

\[
q_\eta(z) v_\gamma(\mathcal{P}) = \sum_{\alpha\beta} u_{\alpha\beta}(\eta, \gamma) \mathcal{P}^\alpha z^\beta, \quad \alpha_i, \beta_i < r_i \quad \forall i = 1, \ldots, n,
\]

for some scalars \( \{u_{\alpha\beta}(\eta, \gamma)\} \). Therefore, from the definition of \( \mu^* \),

\[
y^*_{\gamma\eta} = \int z^\gamma \mathcal{P}^\gamma \, d\mu^* = \int \left( \sum_{i=1}^m q_i(z) h_i(z) + q_\eta(z) \right) \left( \sum_{i=1}^m v_i(\mathcal{P}) h_i(\mathcal{P}) + v_\eta(\mathcal{P}) \right) \, d\mu^* = \int q_\eta(z) v_\gamma(\mathcal{P}) \, d\mu^* = \sum_{\alpha\beta} u_{\alpha\beta}(\eta, \gamma) \int \mathcal{P}^\alpha z^\beta \, d\mu^*
\]

\[
= \sum_{\alpha\beta} u_{\alpha\beta}(\eta, \gamma) y^*_{\alpha\beta}, \quad \alpha_i, \beta_i < r_i \quad \forall i = 1, \ldots, n.
\]

\[\square\]

The \( y^*_{\alpha\beta} \)'s with \( \alpha_i, \beta_i < r_i \) correspond to the irreducible monomials \( x^\alpha \), \( x^\beta \) with respect to the Gröbner basis \( H \), which form a basis of \( \mathbb{R}[x_1, x_2, \ldots, x_n]/I \) viewed as a vector space over \( \mathbb{R} \). In fact, in view of the triangular form \( \{2.6\} - \{2.7\} \), the Gröbner basis \( \mathcal{H} \) of \( I \) w.r.t. to the lexicographical ordering \( x_1 < \cdots < x_n \) is such that \( r_i = r_i \) for all \( i = 1, \ldots, n \) and \( \mathcal{H} \) has exactly \( n \) terms (Rouillier [12]).

In view of Proposition 3.1 we may redefine the moment matrix \( M_p(\mu^*, y) \) in an equivalent form as follows.
**Definition 3.2** (Construction of $C_i$ in $\{H_i\}$). Let $H = \{h_1, \ldots, h_m\}$ be a reduced Gröbner basis of $I$ w.r.t. the given term ordering (with $x_i^r_i$ the leading term of $h_i$ for all $i = 1, \ldots, n$).

The moment matrix $M_p(\mu^*, y)$ is the moment matrix $M_p(y)$ defined in (2.1) and where:
- every entry $y_{\alpha 0}$ or $y_{0\alpha}$ of $M_p(y)$ is replaced with the (known) $\alpha$-Newton sum $s_\alpha$ of $G$.
- every entry $y_{\gamma \eta}$ in $M_p(y)$ is replaced with the linear combination (3.6) of $\{y_{\alpha \beta}\}$ with $\alpha_i, \beta_i < r_i'$ for all $i = 1, \ldots, n$.

Thus, in this equivalent formulation, only a finite number of variables $y_{\alpha \beta}$ are involved in $M_p(\mu^*, y)$, all with $\alpha_i, \beta_i < r_i'$ for all $i = 1, \ldots, n$.

**Remark 3.3.** The above definition of $M_p(\mu^*, y)$ depends on the reduced Gröbner basis $H$ of $G$, whereas the entries $s_\alpha$ only depend on the $g_i$'s.

**Example 3.4.** Let

$$G := \{x_1^2 + x_1, (x_1^2 + 3)x_2 - x_1^2x_2^2 + (x_1^2 - x_1 - 1)x_2 - x_1 + 1\}.$$ 

Then,

$$H = \{x_1^2 + x_1; 6x_2^3 - 3x_1^2x_2^2 + 4x_2x_1^2 - 3x_2x_1 - 2x_2 - x_1^2 - 3x_1 + 2\},$$

and, for instance, denoting "\(~\)" as the reduction process w.r.t. $H$,

$$z_2^3 \sim (3z_2^2z_1^2 - 4z_2z_1^3 + 3z_2z_1 + 2z_2 + z_1^2 + 3z_1 - 2)/6,$$

and as $z_1^2 \sim -z_1^2$, we have

$$y_{4003} = (-3y_{2022} + 4y_{2021} - 3y_{2011} - 2y_{2001} - y_{2020} - 3y_{2010} + 2y_{2000})/6,$$

and the latter expression can be substituted for every occurrence of $y_{4003}$.

**Theorem 3.5.** Let $G$ be a triangular polynomial system as in (2.6)-(2.7) and let $\{s_\alpha\}$ be the Newton sums of $G$ in (2.8). Let $M_p(\mu^*, y)$ be the moment matrix as in Definition 3.2 and let $r_0 := 2 \sum_{j=1}^n (r_j' - 1)$. Then:

(i) For all $p \geq r_0$, $M_p(\mu^*, y) = M_p(\mu^*)$ if and only if $M_p(\mu^*, y) \geq 0$.

(ii) For all $p \geq r_0$, $\text{rank}(M_p(\mu^*)) = \text{rank}(M_{r_0}(\mu^*))$, the number of distinct zeros in $\mathbb{C}^n$ of the polynomial system $G$.

(iii) Let $f \in \mathbb{C}[z, \overline{z}]$ be of degree less than $2p$. All the zeros in $\mathbb{C}^n$ of the polynomial system $G$ are zeros of $f$ if and only if

$$M_p(\mu^*)f = 0.$$ 

In particular, a polynomial $f \in \mathbb{R}[x_1, \ldots, x_n]$ of degree less than $2p$ is in $\sqrt{T}$ if and only if (3.7) holds.

The proof is postponed to (5.2).

**Remark 3.6.** (a) Theorem 3.5(iii) has an equivalent formulation as follows. Let $f \in \mathbb{C}[z, \overline{z}]$ be of degree at most $2p$ and let $\hat{f}$ be its reduction w.r.t. $H$, the Gröbner basis of $G$ defined in Proposition (5.1). Then the condition $M_p(\mu^*)f = 0$ is equivalent to $M_{r_0}(\mu^*)\hat{f} = 0$.

(b) Given a reduced Gröbner basis $H$ of $I$, the condition $M_{r_0}(\mu^*, y) \geq 0$ in Theorem 3.5(ii) is equivalent to the same condition for its submatrix $\hat{M}_{r_0}(\mu^*, y)$ whose indices of rows and columns in the basis (2.1) correspond to independent
monomials \( \{z^\alpha\} \) which form a basis of \( \mathbb{R}[x_1, \ldots, x_n]/I \), their conjugate \( \{\overline{z}^\alpha\} \) and the corresponding monomial products \( z^\alpha \cdot \overline{z}^\beta \). Indeed, the positive semidefinite condition on the latter is equivalent to the positive semidefinite condition on the former.

**Example 3.7.** Consider the trivial example \( \mathbb{G} := \{x^2 + 1\} \) so that \( V_C(I) = \{\pm i\} \). Then the condition \( \hat{M}_{r_0}(\mu^*, y) \succeq 0 \) (see Remark 3.6(b)) reads

\[
\hat{M}_2(\mu^*, y) = \begin{bmatrix}
1 & 0 & 0 & y_{11} \\
0 & y_{11} & -1 & 0 \\
0 & -1 & y_{11} & 0 \\
y_{11} & 0 & 0 & 1
\end{bmatrix} \succeq 0,
\]

which clearly implies \( y_{11} = 1 = \int z \, d\mu^* \). Moreover,

\[
\text{rank}(M_{r_0}(\mu^*, y)) = \text{rank}(\hat{M}_{r_0}(\mu^*, y)) = 2 = |V_C(I)|.
\]

Similarly, let \( \mathbb{G} := \{x_1^2 + 1, x_1x_2 + 1\} \) so that \( V_C(I) = \{(i, i), (-i, -i)\} \). We have \( r_0 = 2 \) and with the lexicographical ordering \( x_1 < x_2 \), \( \hat{H} := \{x_1^2 + 1, x_2 - x_1\} \) is a reduced Gröbner basis of \( I \). Hence, in the moment matrix \( M_{r_0}(\mu^*, y) \) every \( y_{\alpha_1\alpha_2\beta_1\beta_2} \) is replaced with \( y_{\alpha_1+\beta_1,0+\alpha_2+\beta_2} \). Moreover, we only need to consider \( \alpha_1, \beta_1 \leq 1 \). Therefore, we only need to consider the monomials \( \{z_1, z_1\overline{z}, \overline{z}_1\overline{z}\} \), and in view of Remark 3.6(b), the (equivalent) condition \( \hat{M}_{r_0}(\mu^*, y) \succeq 0 \) reads (denoting \( y_{1010} = y \))

\[
\begin{bmatrix}
1 & 0 & 0 & y \\
0 & y & -1 & 0 \\
0 & -1 & y & 0 \\
y & 0 & 0 & 1
\end{bmatrix} \succeq 0,
\]

which implies \( y = 1 = \int z \, d\mu^* \). Moreover,

\[
\text{rank}(M_{r_0}(\mu^*, y)) = \text{rank}(\hat{M}_{r_0}(\mu^*, y)) = 2 = |V_C(I)|.
\]

3.3. **Conditions for a localization of zeros of \( \mathbb{G} \).** Let \( w_i \in \mathbb{C}[z, \overline{z}], i = 1, \ldots, m \), be given polynomials and let \( \mathbb{K} = \mathbb{C}^m \) be the set defined by

\[
(3.8) \quad \mathbb{K} := \{z \in \mathbb{C}^n \mid w_i(z, \overline{z}) \geq 0, \quad i = 1, \ldots, m\}.
\]

We now consider the following issue:

**Under what conditions on the coefficients of the polynomials \( g_i \)'s are all the zeros of the triangular system \( \mathbb{G} \) contained in \( \mathbb{K} \)?**

Let \( M_p(w_i, y) \) be the localizing matrices (cf. §2.2) associated with the polynomials \( w_i \), for all \( i = 1, \ldots, m \). As we did for the moment matrix \( M_p(\mu^*, y) \) in Definition 3.2 we define \( M_p(\mu^*, w_i, y) \) to be the localizing matrix \( M_p(w_i, y) \) where the entries \( y_{\alpha_0} \) and \( y_{\alpha 0} \) are replaced with the \( \alpha \)-Newton sums \( s_\alpha \), and where all the \( y_{\alpha \beta} \) are replaced by the linear combinations (3.6) of the \( \{y_{\alpha \beta}\} \) with \( \alpha_1, \beta_1 < r_i' \) for all \( i = 1, \ldots, n \). Accordingly, \( M_p(\mu^*, w_i) \) is obtained from \( M_p(w_i, y) \) by replacing \( y \) with \( y^* \) as in Proposition 3.1.

**Theorem 3.8.** Let \( \mathbb{G} \) be the triangular system in (2.6), (2.7) and let \( M_{r_0}(\mu^*, y) \) be as in Theorem 3.5 Then, all the zeros of \( \mathbb{G} \) are in \( \mathbb{K} \) if and only if

\[
(3.9) \quad M_{r_0}(\mu^*, w_i) \succeq 0, \quad i = 1, \ldots, m.
\]
Equivalently, all the zeros of $\mathbb{G}$ are in $\mathbb{K}$ if and only if the system of linear matrix inequalities
\begin{equation}
M_{r_0}(\mu^*, y) \succeq 0, \quad M_{r_0}(\mu^*, w_i, y) \succeq 0, \quad i = 1, \ldots, m,
\end{equation}
has a solution $y$.

**Proof.** The necessity is obvious. Indeed, assume that all the zeros of $\mathbb{G}$ are in $\mathbb{K}$. Let $\mu^*$ be as in (3.11) and let $y^* := \{s_{\alpha}, y_{\alpha\beta}\}$ be the infinite sequence of moments of $\mu^*$. Then, of course, $M_p(\mu^*) \succeq 0$ and
\[M_p(\mu^*, w_i) = M_p(w_i y^*) \succeq 0, \quad i = 1, \ldots, m,
\]
for all $p \in \mathbb{N}$, is a necessary condition for $\mu^*$ to have its support in $\mathbb{K}$. Thus, $y := \{y_{\alpha\beta}\}$ is a solution of (3.10).

Conversely, let $y$ be a solution of (3.10). From Theorem 4.5(i) \{s_{\alpha}, y_{\alpha\beta}\} is the moment vector of $\mu^*$, that is, \{y_{\alpha\beta}\} for all $\alpha, \beta$ with $\alpha_i, \beta_i < r_i'$, for all $i = 1, \ldots, n$. Then, all the other $y_{\alpha\beta}$ can be obtained from the former by (3.10). Therefore, in view of the construction of the localizing matrices $M_p(\mu^*, w_i, y)$, we have
\[M_p(\mu^*, w_i, y) = M_p(\mu^*, w_i, y^*) = M_p(\mu^*, w_i).
\]

Moreover, using the terminology of Curto and Fialkow [10] (see also the proof of Theorem 5.5, all the moment matrices $M_p(\mu^*, y) = M_p(\mu^*)$ (\(p > r_0\)) are flat positive extensions of $M_{r_0}(\mu^*, y) = M_{r_0}(\mu^*)$. As $M_{r_0}(\mu^*, w_i, y) = M_{r_0}(\mu^*, w_i) \succeq 0$, it follows from Theorem 1.6 in Curto and Fialkow [6] that $\mu^*$ has its support contained in $\mathbb{K}$. Hence, as $\mu^*$ is supported on all the zeros of $\mathbb{G}$, all the zeros of $\mathbb{G}$ are in $\mathbb{K}$. \hfill \square

### 3.4. Triangular systems with only real zeros.

In this section we are interested in conditions on the coefficients of the polynomials $g_i$’s for the triangular system $\mathbb{G}$ to have all its zeros real (i.e. in $\mathbb{R}^n$). One way to proceed is to apply Theorem 3.8 with the set $\mathbb{K}$ defined by $\mathbb{K} := \{z \in \mathbb{C}^n \mid w_i(z, \overline{z}) = 0, \ i = 1, \ldots, n\}$ with $z \mapsto w_i(z, \overline{z}) := z_i - \overline{z}_i$ for all $i = 1, \ldots, n$.

In this case, everything simplifies because the localizing conditions
\[M_p(\mu^*, w_i, y) = 0, \quad i = 1, \ldots, n, \quad p \in \mathbb{N},
\]
(necessary for $\mu^*$ to have its support on $\mathbb{K}$), simply mean that for every $\alpha, \beta \in \mathbb{N}^n$,
\[y_{\alpha\beta} = y_{\alpha + \beta, 0} = y_{0, \alpha + \beta} = s_{\alpha + \beta}.
\]

In other words, we only need to deal with the Newton sums $\{s_{\alpha}\}$ of $\mathbb{G}$. In particular, to define $M_p(\mu^*, y)$, we do not have to introduce the reduced Gröbner basis $H$ of $I$ in Definition 4.2. Thus, the moment matrix $M_p(\mu^*)$ simplifies, and we only need consider the basis of monomials
\begin{equation}
1, x_1, \ldots, x_n, x_1^2, x_1 x_2, \ldots, x_1^n, x_2^2, \ldots, x_n^2, \ldots
\end{equation}
(without their conjugates) for the real-valued polynomials in $\mathbb{R}[x_1, \ldots, x_n]$.

Therefore, with $\mu^*$ as in (3.11), the moment matrix $M_p(\mu^*)$ is now indexed in the basis (3.11) and is completely known. Indeed,
- if $M_p(\mu^*)(1, j) = s_{\alpha}$ if the column $j$ corresponds to the monomial $x_{\alpha}$ in the basis (3.11), and
- if $M_p(\mu^*)(1, j) = s_{\alpha}$ and $M_p(\mu^*)(i, 1) = s_{\beta}$, then $M_p(\mu^*)(i, j) = s_{\alpha + \beta}$.

In fact, as $M_p(\mu^*)$ is completely determined from the Newton sums $\{s_{\alpha}\}$ of $\mathbb{G}$, let us call $M_p(s)$ the moment matrices $M_p(\mu^*)$ for all $p \in \mathbb{N}$. 

---

License or copyright restrictions may apply to redistribution; see http://www.ams.org/journal-terms-of-use
Next, let $K_1 \subset \mathbb{R}^n$ be the semi-algebraic set defined by
\[ K_1 := \{ x \in \mathbb{R}^n \mid u_i(x) \geq 0, \ i = 1, \ldots, m \}, \]
for some given polynomials $u_i \in \mathbb{R}[x_1, \ldots, x_n], i = 1, \ldots, m$.

We also denote by $M_p(u_i, s)$ the localizing matrix $M_p(u_i y)$ indexed in the basis (3.11), and where all the entries $\{y_{i \alpha}\}$ are replaced with the corresponding Newton sums $\{s_{i \alpha}\}$. We obtain

**Theorem 3.9.** Let $G$ be the triangular system defined in (2.6)–(2.7) and let $\{s_{i \alpha}\}$ be the Newton sums of $G$ defined in (2.2). Let $r_0 := \sum^{n}_{i=1} (r'_j - 1)$ with $r'_j$ as in Theorem 3.8.

(i) All the zeros of $G$ are real if and only if
(\begin{equation}
M_{r_0}(s) \succeq 0.
\end{equation})

Moreover, the number of distinct zeros is $\text{rank}(M_{r_0}(s))$.

(ii) All the zeros of $G$ are real and in $K_1$ if and only if
(\begin{equation}
M_{r_0}(s) \succeq 0, \ M_{r_0}(u_i, s) \succeq 0, \ i = 1, \ldots, m.
\end{equation})

**Proof.** This is just a particular case of Theorem 3.8 where the simplification is due to the localizing constraints $M_p(u_i y) = 0$ for all $i = 1, \ldots, n$, which permits us to deal only with the Newton sums $\{s_{i \alpha}\}$ of $G$. Again, as in the proof of Theorem 3.8 one uses Theorem 1.6 of Curto and Fialkow [6], but this time for measures on $\mathbb{R}^n$ and not on $\mathbb{C}^n$. \qed

**Example 3.10.** Let $G := \{x^2_1 - 1, x_1 x^2_2 - 1\}$ so that $V_G(I) \neq V_R(I)$. In the matrix $M_{r_0}(s)$ we only need to consider its submatrix $\hat{M}_2(s)$ with rows and columns indexed by the monomials $\{1, x_1, x_2, x_1 x_2\}$ because $x^2_1$ and $x^2_2$ are linear combinations of those monomials (see Remark 3.6(b)). Therefore,
\begin{equation}
\hat{M}_2(s) = \begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 \\
0 & 0 & 1 & 0
\end{bmatrix}
\end{equation}

and obviously, $M_2(s) \succeq 0$ does not hold.

Now with $G := \{x^2_1 - 1, x_1 x_2 - 1\}$ we have $V_G(I) = V_R(I) = \{(1, 1), (-1, -1)\}$, and we obtain
\begin{equation}
\hat{M}_1(s) = \begin{bmatrix}
1 & 0 & 0 \\
0 & 1 & 1 \\
0 & 1 & 1
\end{bmatrix} \succeq 0,
\end{equation}

with $\text{rank}(\hat{M}_1(s)) = 2 = |V_R(I)|$.

Theorem 3.9 is the analogue in the multivariate case of the result in Lasserre [9] in which one obtains a similar necessary and sufficient condition on the Newton sums of a univariate polynomial $g$, for $g$ to have all its zeros real and in a prescribed interval $[a, b]$. In the univariate case, and with $G = \{g\}$ for a single univariate polynomial $g$ of degree $n + 1$, one may check that $(n + 1)M_n(s)$ is just the (Hankel) matrix associated with Hermite’s quadratic form $\text{Her}(g, 1)$ (see [4] p. 99]). Similarly, given another univariate polynomial $h$, $(n + 1)M_n(h, s)$ is the matrix associated with Hermite’s quadratic form $\text{Her}(g, h)$ and whose signature gives the number of real zeros of $g$ that satisfy $h(x) > 0$ minus the number of real zeros that satisfy $h(x) < 0$.
(4, Theorem 4.13). Both $M_n(s)$ and $M_n(h,s)$ are explicit in terms of standard Newton sums.

In the multivariate case, let $\widehat{M}_{r_0}(s)$ (resp. $\widehat{M}_{r_0}(u_i,s)$) be the submatrix obtained from $M_{r_0}(s)$ (resp. $M_{r_0}(u_i,s)$) by keeping only the rows and columns indexed by monomials $\{x^\alpha\}$ which form a basis of $\mathbb{R}[x_1,\ldots,x_n]/I$ as an $\mathbb{R}$-vector space. Then, one may check that (after scaling) $\widehat{M}_{r_0}(s)$ is the matrix associated with the multivariate Hermite’s quadratic form Her$(G,1)$ (see [4, p. 129]). Similarly (after scaling again), $\widehat{M}_{r_0}(u_i,s)$ is the matrix associated with the multivariate Hermite’s quadratic form Her$(G,u_i)$, and whose signature gives the number of real zeros of $G$ that satisfy $u_i(x) > 0$ minus the number of real zeros that satisfy $u_i(x) < 0$ ([4, Theorem 4.72]). Here, and as in the univariate case, both $\widehat{M}_{r_0}(s)$ and $\widehat{M}_{r_0}(u_i,s)$ are obtained explicitly in terms of generalized Newton sums, because of the triangular form of $G$. Note that in Theorem 3.9 we do not need to determine a basis of $\mathbb{R}[x_1,\ldots,x_n]/I$.

3.5. Counting real zeros. We still consider a triangular system $G$ as in (2.6)-(2.7) and now consider the issue of counting the real zeros of $G$.

As we did for $\mu^*$, we build up the moment matrix of a probability measure $\mu$ with support on the real zeros of $G$. This time, we cannot use the Newton sums $\{s_\alpha\}$ in (2.2) because some zeros of $G$ may not be real. Therefore, we replace $s_\alpha$ with the unknown $y_\alpha$. Namely, we define the moment matrix $M_\mu(y)$ as follows.

**Definition 3.11.** Let $H := \{h_1,\ldots,h_m\}$ be a reduced Gröbner basis of $I$ w.r.t. some term ordering (with $x_i^r$ the leading term of $h_i$ for all $i = 1,\ldots,n \leq m$). Then $M_\mu(y)$ is the moment matrix defined in (2.1) but now with rows and columns indexed in the basis (3.11), and where:

For every $\alpha \in \mathbb{N}^n$, the monomial $y_\alpha$ is replaced by a linear combination of the variables $y$ with $\beta_i < r_i'$ for all $i = 1,\ldots,n$ (see Proposition 3.11) coming from the reduction of the monomial $z^\alpha$ w.r.t. $H$. That is, if $z^\alpha \rightsquigarrow \sum \beta u_\beta(\alpha)z^\beta$ with $\beta_i < r_i'$ for all $i$, and for some scalars $\{u_\beta(\alpha)\}$, then $y_\alpha$ is replaced with $\sum \beta u_\beta(\alpha)y_\beta$.

We denote by $V_\mathbb{R}(I) \subset \mathbb{R}^n$ the set of real zeros of $G$ and $I(V_\mathbb{R}(I)) \subset \mathbb{R}[x_1,\ldots,x_n]$ the ideal generated by the variety $V_\mathbb{R}(I)$. Recall that a polynomial $f \in \mathbb{R}[x_1,\ldots,x_n]$ is identified with its vector (also denoted by $f$) of coefficients in the basis (3.11).

**Proposition 3.12.** Let $G$ be a triangular system as in (2.6)-(2.7) and let $M_\mu(y)$ be as in Definition 3.11, and let $r_0 := \sum_i (r_i' - 1)$. Then:

(a) The number $s_0$ of distinct real zeros of $G$ is given by the maximal rank of $M_{r_0}(y)$ over all possible solutions $y$ (if any) of $M_{r_0}(y) \geq 0$.

(b) Let $y$ be such that $M_{r_0}(y) \geq 0$ with rank($M_{r_0}(y)$) = $s_0$. Let $f \in \mathbb{R}[x_1,\ldots,x_n]$ be a polynomial of degree less than $r_0$. Then,

\begin{equation}
(3.14) \quad f \in I(V_\mathbb{R}(I)) \iff M_{r_0}(y)f = 0.
\end{equation}

**Proof.** (a) Assume that there is a solution $y$ to $M_{r_0}(y) \geq 0$. Then, proceeding as in the proof of Theorem 3.5 using Theorem 1.6 in Curto and Fialkow [6], it follows that $y$ is the vector of moments of a rank($M_{r_0}(y)$)-atomic probability measure $\mu$, this time on $\mathbb{R}^n$, and with support on the real zeros of $G$. Therefore, the rank of $M_{r_0}(y)$ is not larger than the number $s_0$ of distinct real zeros of $G$.

Next, let $\{x(i)\}_{i=1}^{s_0}$ be the real distinct zeros of $G$ and let $\mu := s_0^{-1} \sum_{i=1}^{s_0} \delta_{x(i)}$ (with $\delta_x$ the Dirac measure at the point $x \in \mathbb{R}^n$). Let $y$ be the infinite sequence
of all the moments of $\mu$. It follows easily that the moment matrices $M_p(y)$ are exactly as in Definition 3.11 and moreover, $M_{r_0}(y) \succeq 0$ holds, as it is a necessary condition for $y$ to be a moment sequence. As (from its definition) $\mu$ is an $s_0$-atomic probability measure with support on the distinct real zeros of $G$, we conclude from what precedes that rank$(M_{r_0}(y)) = s_0$.

(b) Let $f \in I(V_G(I))$ and let $y$ be as in Proposition 3.12(b). From (a) it follows that $y$ is the sequence of moments (up to order $2r_0$) of a probability measure $\mu_y$ with support on the $s_0$ distinct real zeros of $G$ (that is, with support on all the points of $V_G(I)$). Hence, from (2.2)

$$0 = \int f^2 \, d\mu_y = \langle M_{r_0}(y) f, f \rangle \Rightarrow M_{r_0}(y) f = 0 \quad \text{because} \quad M_{r_0}(y) \succeq 0.$$ 

Conversely, let $f$ be such that $M_{r_0}(y) f = 0$. Then

$$\langle M_{r_0}(y) f, f \rangle = 0 \Rightarrow \int f^2 \, d\mu_y = 0 \Rightarrow f \equiv 0 \quad \mu_y\text{-almost everywhere,}$$

which implies that $f(x) = 0$ for all $x \in V_G(I)$, or, equivalently, $f \in I(V_G(I))$. \hfill \Box

Remark 3.13. (i) The assumption on the degree of $f$ in Proposition 3.12(b) is not restrictive, for one may first reduce $f$ w.r.t. the Gröbner basis $H$ of $I$ to end up with a polynomial of degree less than $r_0$.

(ii) Proposition 3.12(a) should not be misleading. Finding a vector $y$ such that $M_{r_0}(y) \succeq 0$ has maximal rank is not necessarily easy. (However, note that SDP solvers that use interior point methods usually find solutions with highest rank.) Proposition 3.12(a)-(b) should be viewed as an alternative characterization of the number of real zeros of $G$ and of the ideal $I(V_G(I))$ in terms of moment matrices.

Note that in contrast to counting techniques via multivariate Hermite’s quadratic form, knowledge of a basis of $\mathbb{R}[x_1, \ldots, x_n]/I$ is not needed in Proposition 3.12.

4. Conclusion

In this paper we have considered a system $G$ of polynomial equations in triangular form and show that several characterizations of the zeros of $G$ may be obtained from positive semidefinite (numerical) conditions on appropriate moment and localizing matrices. In particular, the triangular form of $G$ permits us to define the analogue for the multivariate case of Newton sums of a univariate polynomial. As in the univariate case, these multivariate Newton sums permit us to give explicit conditions on the coefficients of the polynomials $g_i$’s for $G$ to have only real zeros, and for those zeros to be in a given semi-algebraic set of $\mathbb{R}^n$.

5. Proofs

5.1. Proof of Proposition 2.7.2

Proof. The proof is by induction. In view of the triangular form (2.6)-(2.7), the zero set of $G$ in $\mathbb{C}^n$ (or, equivalently, the variety $V_G(I)$ associated with $I$) consists of $s := \prod_{j=1}^s r_j$ zeros that we label $z(i)$, $i = 1, \ldots, s$, counting their multiplicity.

In addition, still in view of (2.6)-(2.7), any particular zero $z(i) \in \mathbb{C}^n$ of $G$ can be written

$$z(i) = [z_1(i_1), z_2(i_1, i_2), \ldots, z_n(i_1, \ldots, i_n)].$$
for some multi-index \( i_1 \leq r_1, \ldots, i_n \leq r_n \), and where each \( z_k(i_1, \ldots, i_k) \in \mathbb{C} \) is a zero of the univariate polynomial \( x \mapsto g_k(z_1(i_1), \ldots, z_{k-1}(i_1, \ldots, i_{k-1}), x) \) (where multiplicity is taken into account).

Therefore, for every \( \alpha \in \mathbb{N}^n \), the \( \alpha \)-Newton sum \( y_\alpha \) defined in (2.8) can be written

\[
(5.1) \quad s y_\alpha := \sum_{i_1=1}^n z(i)^\alpha = \sum_{i_1 \leq r_1, \ldots, i_n \leq r_n} z_1(i_1)^{\alpha_1} z_2(i_1, i_2)^{\alpha_2} \cdots z_n(i_1, \ldots, i_n)^{\alpha_n}.
\]

Let us make the following induction hypothesis.

\[ H_k. \] For every \( p, q \in \mathbb{R}[x_1, \ldots, x_k] \)

\[
S_k(p, q) := \sum_{i_1, \ldots, i_k} \frac{p(z_1(i_1), \ldots, z_k(i_k))}{q(z_1(i_1), \ldots, z_k(i_k))}
\]

\[
(5.2) \quad = \sum_{i_1, \ldots, i_k} \frac{\sum_\alpha p_\alpha z_1(i_1)^{\alpha_1} \cdots z_k(i_k)^{\alpha_k}}{\sum_\alpha q_\alpha z_1(i_1)^{\alpha_1} \cdots z_k(i_k)^{\alpha_k}}
\]

is a rational fraction of coefficients of the polynomials \( g_i \)'s, \( i = 1, \ldots, k \).

Observe that (5.1) is a particular case of (5.2) in \( H_n \).

We first prove that \( H_1 \) is true. Let \( p, q \in \mathbb{R}[x_1] \) and

\[
S(p, q) = \sum_{j=1}^r \frac{\sum_{k} p_k z_1(j)^k}{\sum_{k} q_k z_1(j)^k},
\]

with \( \{z_1(j)\} \) being the zeros of \( x_1 \mapsto g_1(x_1) \), counting their multiplicity.

Reducing to a common denominator, \( S(p, q) \) reads

\[
S(p, q) = \frac{P(z_1(1), \ldots, z_1(r_1))}{Q(z_1(1), \ldots, z_1(r_1))},
\]

for some symmetric polynomials \( P, Q \) of the variables \( \{z_1(j)\} \) and whose coefficients are polynomials of coefficients of \( p, q \).

Therefore, by the fundamental theorem of symmetric functions, both numerator \( P(\cdot) \) and denominator \( Q(\cdot) \) are rational fractions of coefficients of \( g_1 \) (polynomials if \( g_1 \) is monic). Thus, \( H_1 \) is true, and we can write \( S_1(p, q) = u_{pq}(g_1)/v_{pq}(g_1) \) for some polynomials \( u_{pq}, v_{pq} \) of coefficients of \( g_1 \). The coefficients of \( u_{pq}, v_{pq} \) are themselves polynomials of coefficients of the polynomials \( p, q \).

Next, assume that \( H_j \) is true for all \( 1 \leq j \leq k \), that is, for all \( j = 1, \ldots, k \) and \( p, q \in \mathbb{R}[x_1, \ldots, x_j] \),

\[
(5.3) \quad S_j(p, q) = u_{pq}(g_1, \ldots, g_j)/v_{pq}(g_1, \ldots, g_j)
\]

for some polynomials \( u_{pq}, v_{pq} \) of coefficients of the polynomials \( g_1, \ldots, g_j \).

We are going to show that \( H_{k+1} \) is true. Let \( p, q \in \mathbb{R}[x_1, \ldots, x_{k+1}] \) and let

\[
S_{k+1}(p, q) = \sum_{i_1, \ldots, i_{k+1}} \frac{\sum_\alpha p_\alpha z_1(i_1)^{\alpha_1} \cdots z_{k+1}(i_1, \ldots, i_{k+1})^{\alpha_{k+1}}}{\sum_\alpha q_\alpha z_1(i_1)^{\alpha_1} \cdots z_{k+1}(i_1, \ldots, i_{k+1})^{\alpha_{k+1}}}.
\]

\( S_{k+1}(p, q) \) can be rewritten as

\[
(5.4) \quad S_{k+1}(p, q) = \sum_{i_1, \ldots, i_k} \left[ \sum_{j=1}^{r_{k+1}} \frac{\sum_\alpha p_\alpha z_1(i_1)^{\alpha_1} \cdots z_k(i_1, \ldots, i_k)^{\alpha_k} z_{k+1}(i_1, \ldots, i_k, j)^{\alpha_{k+1}}}{\sum_\alpha q_\alpha z_1(i_1)^{\alpha_1} \cdots z_k(i_1, \ldots, i_k)^{\alpha_k} z_{k+1}(i_1, \ldots, i_k, j)^{\alpha_{k+1}}} \right].
\]
In (5.4), the term
\[
A := \left[ \sum_{j=1}^{r_{k+1}} a(z_k(i_1, \ldots, i_k))^\alpha_k \right] 
\]
can in turn be written as
\[
(5.5) \quad A = \sum_{j=1}^{r_{k+1}} \frac{\tilde{p}(z_k+1(i_1, \ldots, i_k, j))}{\tilde{q}(z_k+1(i_1, \ldots, i_k, j))},
\]
for some univariate polynomials \(\tilde{p}, \tilde{q} \in \mathbb{R}[x]\) of the variable \(z_k+1(i_1, \ldots, i_k, j)\) (which is a zero of the univariate polynomial \(x \mapsto g_k+1(z_1(i_1), \ldots, z_k(i_1, \ldots, i_k), x)\)) and whose coefficients are polynomials in the variables \(z_1(i_1), z_2(i_1, i_2), \ldots, z_k(i_1, \ldots, i_k)\).

Proof. (i) Let \(U, V \in \mathbb{R}[x_1, \ldots, x_k]\) whose coefficients are polynomials of coefficients of \(g_k+1\).

The coefficients of the polynomials \(u_{\tilde{p}q}, v_{\tilde{p}q}\) are themselves polynomials of coefficients of \(p, q\) and of \(z_1(i_1), \ldots, z_k(i_1, \ldots, i_k)\). Hence, substituting for \(A\) in (5.3) we obtain
\[
S_{k+1}(p, q) = \sum_{i_1, \ldots, i_k} \frac{\sum_{j=1}^{r_{k+1}} \tilde{p}(z_k+1(i_1, \ldots, i_k))}{\sum_{j=1}^{r_{k+1}} \tilde{q}(z_k+1(i_1, \ldots, i_k))}.
\]
for some polynomials \(U, V \in \mathbb{R}[x_1, \ldots, x_k]\) whose coefficients are polynomials of coefficients of \(g_k+1\).

We next use the induction hypothesis \(H_k\) by which \(S_k(U(g_k+1), V(g_k+1))\) is a rational fraction \(f_{UV}(g_1, \ldots, g_k)/h_{UV}(g_1, \ldots, g_k)\) of coefficients of the polynomials \(g_1, \ldots, g_k\). As the coefficients of \(f_{UV}, h_{UV}\) are themselves rational fractions of coefficients of \(g_{k+1}\) we finally obtain that
\[
S_{k+1}(p, q) = \frac{u_{\tilde{p}q}(g_1, \ldots, g_{k+1})}{v_{\tilde{p}q}(g_1, \ldots, g_{k+1})},
\]
that is, a rational fraction of coefficients of the polynomials \(g_1, \ldots, g_{k+1}\). Hence \(H_{k+1}\) is true, and therefore, the induction hypothesis is true.

Now Proposition 2.2 follows from \(H_k\) and the expression (5.1) for the \(\alpha\)-Newton sum \(y_0\). That the \(\{y_\alpha\}\) can be computed recursively is clear from the above proof of the induction hypothesis \(H_k\).

5.2. Proof of Theorem 3.3

Proof. (i) Let \(p > r_0\) be fixed, arbitrary, and write
\[
M_p(\mu^*, y) = \begin{bmatrix} M_{r_0}(\mu^*, y) & B \\ -B' & C \end{bmatrix}.
\]
Consider an arbitrary column \(B(\cdot, j)\), \(C(\cdot, j)\). By definition of the moment matrix, \(B(1, j)\) is a monomial \(z_1^{\gamma_1}z_2^{\gamma_2}\) for which \(\gamma_i > r'_i\) or \(\eta_k > r'_k\) for at least one index \(i\) or
k. By Proposition 3.1
\[(5.6)\quad z^{\alpha}z^{\beta} = \sum_{\alpha,\beta} u_{\alpha\beta}(\eta, \gamma) z^{\alpha} y^{\beta}, \quad \alpha_i, \beta_i < r_i', \forall i = 1, \ldots, n,\]
for some scalars \(\{u_{\alpha\beta}(\eta, \gamma)\}\), so that, from the construction of \(M_p(\mu^*, y)\), we have
\[
B(1, j) = y_{\gamma \gamma} = \sum_{\alpha,\beta} u_{\alpha\beta}(\eta, \gamma) y_{\alpha\beta} = \sum_{\alpha,\beta} u_{\alpha\beta}(\eta, \gamma) M_{\gamma \gamma}(\mu^*, y)(1, j_{\alpha\beta}),
\]
where \(j_{\alpha\beta}\) is the index of the column of \(M_{\gamma \gamma}(\mu^*, y)\) corresponding to the monomial \(z^{\alpha} y^{\beta}\). Next, consider an element \(B(k, j)\) of the column \(B(., j)\). The element \(k\) of \(M_p(\mu^*, y)(k, 1)\) is a monomial \(z^{\alpha}z^{\beta}\) and from the definition of \(M_p(\mu^*, y)\), we have \(B(k, j) = y_{q, \gamma + p}\). Now, from (5.6) we have
\[
z^{\alpha}z^{\beta} = \sum_{\alpha,\beta} u_{\alpha\beta}(\eta, \gamma) z^{\alpha} y^{\beta},
\]
which implies
\[
y_{q, \gamma + p} = \sum_{\alpha,\beta} u_{\alpha\beta}(\eta, \gamma) y_{\alpha + q, \beta + p},
\]
or, equivalently,
\[
B(k, j) = \sum_{\alpha,\beta} u_{\alpha\beta}(\eta, \gamma) M_{\gamma \gamma}(k, j_{\alpha\beta}).
\]
The same argument holds for \(C(., j)\). Hence,
\[
\begin{bmatrix} B \\ C \end{bmatrix}(j) = \sum_{\alpha,\beta} u_{\alpha\beta}(\eta, \gamma) \begin{bmatrix} M_{\gamma \gamma}(\mu^*, y) \\ B^* \end{bmatrix}(j) \quad \forall j,
\]
which, in view of \(M_p(\mu^*, y) \succeq 0\), implies that
\[
\text{rank}(M_p(\mu^*, y)) = \text{rank}(M_{\gamma \gamma}(\mu^*, y)).
\]
As \(p > r_0\) was arbitrary, and using the terminology of Curto and Fialkow [39], it follows that the matrices \(M_p(\mu^*, y)\) are flat positive extensions of \(M_{\gamma \gamma}(\mu^*, y)\) for all \(p > r_0\). This in turn implies that, indeed, the entries of \(M_{\gamma \gamma}(\mu^*, y)\) are moments of some rank\((M_{\gamma \gamma}(\mu^*, y))-atomic probability measure \(\mu\).

We next prove that \(\mu = \mu^*\), i.e., the condition \(M_{\gamma \gamma}(\mu^*, y) \succeq 0\) determines a unique vector \(y = y^*\) that corresponds to the vector of moments of \(\mu^*\), up to order \(2r_0\).

Given the Gröbner basis \(H = \{h_i\}_{i=1}^m\) of \(I = \langle g_1, \ldots, g_m \rangle\) (already considered in the proof of Proposition 3.1), let \(\overline{h_i} \in \mathbb{C}[z, \overline{z}]\) be the conjugate polynomial of \(h_i\), i.e., \(\overline{h_i}(z, \overline{z}) = h_i(\overline{z})\) for all \(i = 1, \ldots, m\).

We first prove that
\[
(5.7) \quad M_p(h_i y) = 0, \quad M_p(\overline{h_i} y) = 0, \quad i = 1, \ldots, m, \quad p \in \mathbb{N},
\]
where \(M_p(h_i y)\) (resp. \(M_p(\overline{h_i} y)\)) is the localizing matrix associated with the polynomials \(h_i\) (resp. \(\overline{h_i}\)).

By Proposition 3.1, recall that any entry \(y_{\alpha\beta}\) of \(M_p(\mu^*, y)\) is replaced by a linear combination of the \(y_{\alpha\beta}\)’s with \(\alpha_i, \beta_i < r_i'\) for all \(i = 1, \ldots, n\). This linear combination is coming from the reduction of the monomials \(\{z^\alpha\}_{\alpha \in \mathbb{N}^n}\) with respect to \(H\);
that is, let us call \( J \) the set of indices \( \beta \) corresponding to the irreducible monomials \( z^\beta \) w.r.t. \( H \). Then, the reduction of \( z^\alpha \) w.r.t. \( H \) yields

\[
z^\alpha = \sum_{i=1}^{m} q_i(z) h_i(z) + \sum_{\beta \in J} u_\beta(\alpha) z^\beta \quad \text{denoted } z^\alpha \to \sum_{\beta \in J} u_\beta(\alpha) z^\beta,
\]

and similarly,

\[
t^\alpha = \sum_{i=1}^{m} q_i(t) h_i(t) + \sum_{\beta \in J} u_\beta(\alpha) t^\beta \quad \text{denoted } t^\alpha \to \sum_{\beta \in J} u_\beta(\alpha) t^\beta.
\]

From this, we obtain (see the proof of Proposition 3.1)

\[
z^\gamma t^\gamma \to \left( \sum_{\beta \in J} u_\beta(\gamma) z^\beta \right) \left( \sum_{\beta \in J} u_\beta(\eta) t^\beta \right) \to \sum_{\alpha,\beta \in J} u_{\alpha \beta}(\eta, \gamma) t^\alpha z^\beta,
\]

for some scalars \( \{ u_{\alpha \beta}(\eta, \gamma) \} \), and thus the entry \( y_{\eta \gamma} \) of \( M_p(\mu^*, y) \) is replaced with \( \sum_{\alpha,\beta \in J} u_{\alpha \beta}(\eta, \gamma) y_{\alpha \beta} \), or, equivalently,

\[
y_{\eta \gamma} - \sum_{\alpha,\beta \in J} u_{\alpha \beta}(\eta, \gamma) y_{\alpha \beta} = 0.
\]

So let \( p \in \mathbb{N} \) be fixed, and consider the entry \( M_p(h_i y)(k, l) \) of the localizing matrix \( M_p(h_i y) \). Recall that \( M_p(y)(k, l) = y_{\phi \zeta} \) for some \( \phi, \zeta \in \mathbb{N}^n \), and so \( M_p(h_i y)(k, l) \) is just the expression \( t^\beta z^\gamma h_i(z) \), where each monomial \( t^\beta z^\gamma \) is replaced with \( y_{\alpha \beta} \); see (2.4). Next, by definition, \( t^\beta z^\gamma h_i \to 0 \) for all \( i = 1, \ldots, m \). Therefore, when \( y \) is as in Definition 3.2 (that is, when (5.9) holds), writing

\[
t^\beta z^\gamma h_i = \sum_{\eta,\gamma \in \mathbb{N}^n} v_{\eta \gamma} t^\alpha z^\beta \to 0,
\]

and using (5.8)–(5.9), yields

\[
M_p(h_i y)(k, l) = \sum_{\alpha,\beta \in J} v_{\eta \gamma} u_{\alpha \beta}(\eta, \gamma) y_{\alpha \beta} = 0.
\]

Recall that \( p \in \mathbb{N} \), and \( k, l \) were arbitrary. Therefore, when \( y \) is as in Definition 3.2 we have \( M_p(h_i y) = 0 \) (and similarly, \( M_p(h_i y) = 0 \), for all \( i = 1, \ldots, m \) and all \( p \in \mathbb{N} \). That is, (5.7) holds.

Hence, let \( \mu \) be the \( r \)-atomic probability measure encountered earlier (with \( r := \text{rank}(M_{r, \mu^*}(y)) \)), and let \( \{ z(k) \}_{k=1}^{r} \subset \mathbb{C}^n \) be the \( r \) distinct points of the support of \( \mu \), that is,

\[
\mu = \sum_{k=1}^{r} u_k \delta_{z(k)} \quad \text{with } u_k > 0, \quad k = 1, \ldots, r,
\]

where \( \delta_z \) denotes the Dirac measure at \( z \).

For every \( 1 \leq i \leq r \), let \( q_i \in \mathbb{C}[z, \overline{z}] \) be an interpolation polynomial that vanishes at all \( z(k) \), \( k \neq i \), and with \( q_i(z(i), \overline{z(i)}) \neq 0 \). Let \( p \geq \deg q_i \). Then for all \( j = 1, \ldots, m \), we have (also denoting \( q_i \) as the vector of coefficients of \( q_i \in \mathbb{C}[z, \overline{z}] \))

\[
0 = \langle q_i, M_p(h_j y) q_i \rangle = \int |q_i(z, \overline{z})|^2 h_j(z) \mu(dz) = u_i |q_i(z(i), \overline{z(i)})|^2 h_j(z(i)),
\]

and so, \( h_j(z_i) = 0 \) for all \( j = 1, \ldots, m \).
As this is true for all $1 \leq i \leq r$, it follows that
\[
h_j(z(i)) = 0, \quad i = 1, \ldots, r, \quad j = 1, \ldots, m,
\]
that is, $\mu$ has its support contained in $G$. Therefore, with $\{z(i)\}_{i=1}^{s_0}$ being the distinct zeros in $\mathbb{C}^n$ of $G$,
\[
\mu = \sum_{i=1}^{s_0} u_i \delta_{z(i)}, \quad \sum_{i=1}^{s_0} u_i = 1, \quad u_i \geq 0, \quad i = 1, \ldots, n,
\]
for some nonnegative scalars $\{u_i\}$, whereas $\mu^* = s^{-1} \sum_{i=1}^{s} \delta_{z(i)}$ (counting their multiplicity) or $\mu^* = \sum_{i=1}^{s_0} v_i \delta_{z(i)}$ for some nonnegative scalars $\{v_i\}$.

Remember that by definition of the matrices $M_{\alpha}(\mu^*)$ and $M_{\alpha}(\mu^*, y)$, their entries $\{s_\alpha\}$ (corresponding to the Newton sums) are the same. That is,
\[
s_\alpha = \int z^\alpha \, d\mu = \int z^\alpha \, d\mu^*, \quad \alpha_j \leq r_j - 1, \quad j = 1, \ldots, n.
\]

Now, we also know that $s_0$ is less than the number of independent monomials $\{z^{\beta(j)}\}$ (w.r.t. $H$) which form a basis of $\mathbb{R}[x_1, \ldots, x_n]/I$ (with equality if $I = \sqrt{T}$). Therefore, if $\mu \neq \mu^*$, we have
\[
\sum_{i=1}^{s_0} (u_i - v_i) z(i)^{\beta(j)} = 0, \quad j = 1, \ldots, s_0, \quad \text{with } u \neq v,
\]
which yields that the square matrix of the above linear system is singular. Hence some linear combination $\{\lambda_j\}$ of its rows vanishes, i.e.,
\[
\sum_{j=1}^{s_0} \lambda_j z(k)^{\beta(j)} \quad \forall k = 1, \ldots, s_0,
\]
in contradiction with the linear independence of the $\{z^{\beta(j)}\}$. Hence $u = v$, which in turn implies $\mu = \mu^*$. So it follows that $M_{\alpha}(\mu^*, y) \geq 0$ has only one solution, namely $y = y^*$, the (truncated) vector $y^*$ of moments up to order $2r_0$, of the probability measure $\mu^*$.

Finally, this implies that $s_0 = r = \text{rank}(M_{\alpha}(\mu^*, y)) = \text{rank}(M_{\alpha}(\mu^*))$ because by Curto and Fialkow [6, Theor. 1.6], the number of atoms of $\mu = \mu^*$ is precisely $\text{rank}(M_{\alpha}(\mu^*, y))$. This also proves that $M_{\alpha}(\mu^*, y) = M_{\alpha}(\mu^*)$ and thus, (i) and (ii).

To prove (iii), consider a polynomial $f \in \mathbb{C}[z, \overline{z}]$ of degree less than $2p$ with coefficient vector in the basis $\{\overline{z}^k\}$ still denoted $f$. It is clear that if $f(z(i)) = 0$ for all $i = 1, \ldots, s_0$, then
\[
0 = \int |f|^2 \, d\mu^* = \langle M_p(\mu^*)f, f \rangle,
\]
which in turn implies $M_p(\mu^*)f = 0$. Conversely,
\[
M_p(\mu^*)f = 0 \Rightarrow 0 = \langle M_p(\mu^*)f, f \rangle = \int |f|^2 \, d\mu^*,
\]
which in turn implies $f(z) = 0, \mu^*$-a.e.

Finally, let $f \in \mathbb{R}[x_1, \ldots, x_n]$. Recall that $\sqrt{T} = I(V_{\mathbb{C}}(I))$ where $V_{\mathbb{C}}(I) = \{z(i)\}_{i=1}^{s_0}$, that is, $f \in \sqrt{T}$ if and only if $f(z(i)) = 0$ for all $i = 1, \ldots, s_0$. In view of what precedes, $f \in \sqrt{T}$ if and only if $M_p(\mu^*)f = 0$. \qed
References


LAAS-CNRS and Institute of Mathematics, LAAS, 7 Avenue du Colonel Roche, 31077 Toulouse, Cédex, France.
E-mail address: lasserre@laas.fr