On convexity of polynomial paths and generalized majorizations*

Marija Dodig†
Centro de Estruturas Lineares e Combinatórias, CELC, Universidade de Lisboa,
Av. Prof. Gama Pinto 2, 1649-003 Lisboa, Portugal
dodig@cii.fc.ul.pt

Marko Stošić
Instituto de Sistemas e Robótica and CAMGSD, Instituto Superior Técnico,
Av. Rovisco Pais 1, 1049-001 Lisbon, Portugal
mstosic@isr.ist.utl.pt

Submitted: Nov 15, 2009; Accepted: Apr 5, 2010; Published: Apr 19, 2010
Mathematics Subject Classification: 05A17, 15A21

Abstract

In this paper we give some useful combinatorial properties of polynomial paths. We also introduce generalized majorization between three sequences of integers and explore its combinatorics. In addition, we give a new, simple, purely polynomial proof of the convexity lemma of E. M. de Sá and R. C. Thompson. All these results have applications in matrix completion theory.

1 Introduction and notation

In this paper we prove some useful properties of polynomial paths and generalized majorization between three sequences of integers. All proofs are purely combinatorial, and the presented results are used in matrix completion problems, see e.g. [2, 4, 7, 10, 11].

We study chains of monic polynomials and polynomial paths between them. Polynomial paths are combinatorial objects that are used in matrix completion problems, see [7, 9, 11]. There is a certain convexity property of polynomial paths appeared for the first time in [5]. In Lemma 2 we give a simple, direct polynomial proof of that result. We also show that no additional divisibility relations are needed.

*This work was done within the activities of CELC and was partially supported by FCT, project ISFL-1-1431, and by the Ministry of Science of Serbia, projects no. 144014 (M. D.) and 144032 (M. S.).
†Corresponding author.

Also, we explore generalized majorization between three sequences of integers. It presents a natural generalization of a classical majorization in Hardy-Littlewood-Pólya sense [6], and it appears frequently in matrix completion problems when both prescribed and the whole matrix are rectangular (see e.g [1, 4, 11]).

We give some basic properties of generalized majorization, and we prove that there exists a certain path of sequences, such that every two consecutive sequences of the path are related by an *elementary generalized majorization*.

E. Marques de Sá [7] and independently R. C. Thompson [10], gave a complete solution for the problem of completing a principal submatrix to a square one with a prescribed similarity class. The proof of this famous classical result is based on induction on the number of added rows and columns, and one of the crucial steps is the convexity lemma. The original proofs of the convexity lemma, which are completely independent one from the another one, both in [7] and [10] are rather long and involved. Later on, new combinatorial proof of this lemma has appeared in [8]. In Theorem 1, we give simple and the first purely polynomial proof of this result.

1.1 Notation

All polynomials are considered to be monic.

Let $\mathbb{F}$ be a field. Throughout the paper, $\mathbb{F}[\lambda]$ denotes the ring of polynomials over the field $\mathbb{F}$ with variable λ. By $f|g$, where $f, g \in \mathbb{F}[\lambda]$ we mean that g is divisible by f.

If $\psi_1|\cdots|\psi_r$ is a polynomial chain, then we make a convention that $\psi_i = 1$, for any $i \leqslant 0$, and $\psi_i = 0$, for any $i \geqslant r+1$.

Also, for any sequence of integers satisfying $c_1 \geqslant \cdots \geqslant c_m$, we assume $c_i = +\infty$, for $i \leqslant 0$, and $c_i = -\infty$, for $i \geqslant m+1$.

2 Convexity and polynomial paths

Let $\alpha_1|\cdots|\alpha_n$ and $\gamma_1|\cdots|\gamma_{n+m}$ be two chains of monic polynomials. Let

$$\pi_j := \prod_{i=1}^{n+j} \operatorname{lcm}(\alpha_{i-j}, \gamma_i), \quad j = 0, \ldots, m. \tag{1}$$

We have the following divisibility:

Lemma 1 *$\pi_j \mid \pi_{j+1}$, $j = 0, \ldots, m-1$ (i.e. $\pi_0|\pi_1|\cdots|\pi_m$).*

Proof: By the definition of π_j, $j = 0, \ldots, m$, the statement of Lemma 1 is equivalent to

$$\prod_{i=1}^{n+j} \operatorname{lcm}(\alpha_{i-j}, \gamma_i) \mid \prod_{i=1}^{n+j+1} \operatorname{lcm}(\alpha_{i-j-1}, \gamma_i), \quad j = 0, \ldots, m-1,$$

i.e.,

$$\prod_{i=1}^{n} \operatorname{lcm}(\alpha_i, \gamma_{i+j}) \mid \gamma_{j+1} \prod_{i=1}^{n} \operatorname{lcm}(\alpha_i, \gamma_{i+j+1}), \quad j = 0, \ldots, m-1, \tag{2}$$

which is trivially satisfied. ■

By Lemma 1 we can define the following polynomials

$$\sigma_j := \frac{\pi_j}{\pi_{j-1}}, \quad j = 1, \ldots, m. \tag{3}$$

Then, we have the following convexity property of π_i's:

Lemma 2 *$\sigma_j \mid \sigma_{j+1}$, $j = 1, \ldots, m-1$ (i.e. $\sigma_1|\sigma_2|\cdots|\sigma_m$).*

Proof: By the definition of σ_j, $j = 1, \ldots, m$, the statement of Lemma 2 is equivalent to

$$\frac{\prod_{i=1}^{n+j} \operatorname{lcm}(\alpha_{i-j}, \gamma_i)}{\prod_{i=1}^{n+j-1} \operatorname{lcm}(\alpha_{i-j+1}, \gamma_i)} \mid \frac{\prod_{i=1}^{n+j+1} \operatorname{lcm}(\alpha_{i-j-1}, \gamma_i)}{\prod_{i=1}^{n+j} \operatorname{lcm}(\alpha_{i-j}, \gamma_i)}, \quad j = 1, \ldots, m-1,$$

i.e. for all $j = 1, \ldots, m-1$, we have to show that

$$\frac{\gamma_j \operatorname{lcm}(\alpha_1, \gamma_{j+1}) \operatorname{lcm}(\alpha_2, \gamma_{j+2}) \cdots \operatorname{lcm}(\alpha_n, \gamma_{j+n})}{\operatorname{lcm}(\alpha_1, \gamma_j) \operatorname{lcm}(\alpha_2, \gamma_{j+1}) \cdots \operatorname{lcm}(\alpha_n, \gamma_{j+n-1})}$$

$$\mid \frac{\gamma_{j+1} \operatorname{lcm}(\alpha_1, \gamma_{j+2}) \operatorname{lcm}(\alpha_2, \gamma_{j+3}) \cdots \operatorname{lcm}(\alpha_n, \gamma_{j+n+1})}{\operatorname{lcm}(\alpha_1, \gamma_{j+1}) \operatorname{lcm}(\alpha_2, \gamma_{j+2}) \cdots \operatorname{lcm}(\alpha_n, \gamma_{j+n})}. \tag{4}$$

Before proceeding, note that for every two polynomials ψ and ϕ we have

$$\operatorname{lcm}(\psi, \phi) = \frac{\psi\phi}{\gcd(\psi, \phi)} \tag{5}$$

Thus, for every i and j, we have

$$\operatorname{lcm}(\alpha_i, \gamma_{i+j}) = \operatorname{lcm}(\operatorname{lcm}(\alpha_i, \gamma_{i+j-1}), \gamma_{i+j}) = \frac{\gamma_{i+j} \operatorname{lcm}(\alpha_i, \gamma_{i+j-1})}{\gcd(\operatorname{lcm}(\alpha_i, \gamma_{i+j-1}), \gamma_{i+j})}. \tag{6}$$

By applying (6), equation (4) becomes equivalent to

$$\gamma_j \prod_{i=1}^{n} \gcd(\operatorname{lcm}(\alpha_i, \gamma_{i+j}), \gamma_{i+j+1}) \mid \gamma_{n+j+1} \prod_{i=1}^{n} \gcd(\operatorname{lcm}(\alpha_i, \gamma_{i+j-1}), \gamma_{i+j}). \tag{7}$$

By shifting indices, the right hand side of (7) becomes

$$\gamma_{n+j+1} \gcd(\operatorname{lcm}(\alpha_1, \gamma_j), \gamma_{j+1}) \prod_{i=1}^{n-1} \gcd(\operatorname{lcm}(\alpha_{i+1}, \gamma_{i+j}), \gamma_{i+j+1}).$$

This, together with obvious divisibilities $\gamma_j | \gcd(\operatorname{lcm}(\alpha_1, \gamma_j), \gamma_{j+1})$ and $\gcd(\operatorname{lcm}(\alpha_n, \gamma_{n+j}), \gamma_{n+j+1}) | \gamma_{n+j+1}$, proves (7), as wanted. ■

Frequently when dealing with polynomial paths we have the following additional assumptions

$$\gamma_i \mid \alpha_i, \quad i = 1, \ldots, n \tag{8}$$

and

$$\alpha_i \mid \gamma_{i+m}, \quad i = 1, \ldots, n. \tag{9}$$

Then the following lemma follows trivially from the definition of π_i's, for $i = 0$ and $i = m$:

Lemma 3 *$\pi_0 = \prod_{i=1}^{n} \alpha_i$ and $\pi_m = \prod_{i=1}^{n+m} \gamma_i$.*

2.1 Polynomial paths

Let $\alpha = (\alpha_1, \ldots, \alpha_n)$ and $\gamma = (\gamma_1, \ldots, \gamma_{n+m})$ be two systems of nonzero monic polynomials such that $\alpha_1 | \cdots | \alpha_n$ and $\gamma_1 | \cdots | \gamma_{n+m}$. A polynomial path between α and γ has been defined in a following way in [7, 9], see also [11]:

Definition 1 *Let $\epsilon^j = (\epsilon_1^j, \ldots, \epsilon_{n+j}^j)$, $j = 0, \ldots, m$, be a system of nonzero monic polynomials. Let $\epsilon^0 := \alpha$ and $\epsilon^m := \gamma$. The sequence*

$$\epsilon = (\epsilon^0, \epsilon^1, \ldots, \epsilon^m)$$

is a path from α to γ if the following is valid:

$$\epsilon_i^j | \epsilon_{i+1}^j, \quad i = 1, \ldots, n+j-1, \quad j = 0, \ldots, m, \tag{10}$$

$$\epsilon_i^j | \epsilon_i^{j-1} | \epsilon_{i+1}^j, \quad i = 1, \ldots, n+j-1, \quad j = 1, \ldots, m. \tag{11}$$

Consider the polynomials $\beta_i^j := \operatorname{lcm}(\alpha_{i-j}, \gamma_i)$, $i = 1, \ldots, n+j$, $j = 0, \ldots, m$ from (1). Let $\beta^j = (\beta_1^j, \ldots, \beta_{n+j}^j)$, $j = 0, \ldots, m$. Then the following proposition is valid (see Proposition 3.1 in [11] and Section 4 in [7]):

Proposition 1 *There exists a path from α to γ, if and only if*

$$\gamma_i | \alpha_i | \gamma_{i+m}, \quad i = 1, \ldots, n. \tag{12}$$

Moreover, if (12) is valid, then $\beta = (\beta^0, \ldots, \beta^m)$ is a polynomial path between α and γ, and for every path ϵ between α and γ hold

$$\beta_i^j \mid \epsilon_i^j, \quad i = 1, \ldots, n+j, \quad j = 0, \ldots, m.$$

Hence, β is a *minimal path* from α to γ.

The polynomials π_j from (1) are defined as $\pi_j = \prod_{i=1}^{n+j} \beta_i^j$. The polynomials σ_i were used by Sá [7, 9] and by Zaballa [11], but the convexity of π_j's, i.e. the result of Lemma 2, was obtained later by Gohberg, Kaashoek and van Schagen [5]. We gave a direct polynomial proof of this result and we have shown that it holds even without the divisibility relations (12).

3 Generalized majorization

Let $d_1 \geqslant \cdots \geqslant d_\rho$, $f_1 \geqslant \cdots \geqslant f_{\rho+l}$ and $a_1 \geqslant \cdots \geqslant a_l$, be nonincreasing sequences of integers.

Definition 2 *We say that*

$$f \prec' (d,a),$$

i.e., we have a generalized majorization between the partitions $d = (d_1,\dots,d_\rho)$, $a = (a_1,\dots,a_l)$ *and* $f = (f_1,\dots,f_{\rho+l})$, *if and only if*

$$d_i \geqslant f_{i+l}, \quad i = 1,\dots,\rho, \tag{13}$$

$$\textstyle\sum_{i=1}^{\rho+l} f_i = \sum_{i=1}^{\rho} d_i + \sum_{i=1}^{l} a_i, \tag{14}$$

$$\textstyle\sum_{i=1}^{h_q} f_i - \sum_{i=1}^{h_q-q} d_i \leqslant \sum_{i=1}^{q} a_i, \quad q = 1,\dots,l, \tag{15}$$

$$\text{where} \quad h_q = \min\{i | d_{i-q+1} < f_i\}, \quad q = 1,\dots,l.$$

Remark 1 *Recall that in Section 1.1 we have made a convention that* $f_i = +\infty$ *and* $d_i = +\infty$, *for* $i \leqslant 0$, *and that* $f_i = -\infty$, *for* $i > \rho + l$, *and* $d_i = -\infty$, *for* $i > \rho$. *Thus,* h_q*'s are well-defined. In particular, for every* $q = 1,\dots,l$, *we have* $q \leqslant h_q \leqslant q + l$, *and* $h_1 < h_2 < \ldots < h_l$.

Note that if $\rho = 0$, then the generalized majorization reduces to a classical majorization (in Hardy-Littlewood-Pólya sense [6]) between the partitions f and a ($f \prec a$).

If $l = 1$, (13)–(15) are equivalent to

$$d_i \geqslant f_{i+1}, \quad i = 1,\dots,\rho, \tag{16}$$

$$\textstyle\sum_{i=1}^{\rho+1} f_i = \sum_{i=1}^{\rho} d_i + a_1, \tag{17}$$

$$d_i = f_{i+1}, \quad i \geqslant h_1. \tag{18}$$

Indeed, for $l = 1$, (15) becomes

$$\sum_{i=1}^{h_1} f_i \leqslant \sum_{i=1}^{h_1-1} d_i + a_1.$$

The last inequality together with (14), gives

$$\sum_{i=h_1+1}^{\rho+1} f_i \geqslant \sum_{i=h_1}^{\rho} d_i. \tag{19}$$

Finally, from (13), we obtain that (19) is equivalent to (18), as wanted.

Generalized majorization for the case $l = 1$ will be called *elementary generalized majorization*, and will be denoted by

$$f \prec_1' (d,a).$$

In particular, if $l = 1$, and f, d and a satisfy $d_i \geqslant f_i$, $i = 1, \dots, \rho$ and (17), then $h_1 = \rho + 1$, and so $f \prec'_1 (d, a)$.

Note that if $f \prec' (d, a)$, then in the same way as in the proof of the equivalence of (15) and (18), we have

$$d_i = f_{i+l}, \quad i \geqslant h_l - l + 1. \tag{20}$$

The aim of this section is to show that there is a generalized majorization between the partitions d, a and f if and only if there are elementary majorizations between them, i.e. if and only if there exist intermediate sequences that satisfy (16)–(18). In certain sense, we show that there exists a path of sequences between d and f such that every neighbouring two satisfy the elementary generalized majorization (see Theorems 5 and 7 below).

More precisely, we shall show that

$$f \prec' (d, a)$$

if and only if there exist sequences $g^i = (g^i_1, \dots, g^i_{\rho+i})$, $i = 1, \dots, l-1$, with $g^i_1 \geqslant \cdots \geqslant g^i_{\rho+i}$, and with the convention $g^0 := d$ and $g^l := f$, such that

$$g^i \prec'_1 (g^{i-1}, a_i), \quad i = 1, \dots, l.$$

Lemma 4 *Let f, d and a be the sequences from Definition 1. If*

$$f \prec' (d, a),$$

then there exist integers $g_1 \geqslant \cdots \geqslant g_{\rho+l-1}$, such that

(i) $g_i \geqslant f_{i+1}, \quad i = 1, \dots, \rho + l - 1,$

(ii) $d_i \geqslant g_{i+l-1}, \quad i = 1, \dots, \rho,$

(iii) $g_i = f_{i+1}, \quad i \geqslant h,$ *where* $h := \min\{i | g_i < f_i\}$,

(iv) $\displaystyle\sum_{i=1}^{\tilde{h}_q} g_i - \sum_{i=1}^{\tilde{h}_q - q} d_i \leqslant \sum_{i=1}^{q} a_i, \quad q = 1, \dots, l-1,$ *where* $\tilde{h}_q = \min\{i | d_{i-q+1} < g_i\}$,

(v) $\displaystyle\sum_{i=1}^{\rho+l} f_i = \sum_{i=1}^{\rho+l-1} g_i + a_l.$

Proof: Let $H_1, \dots, H_{l-1}$ be integers defined as

$$H_q := \sum_{i=1}^{q} a_i - \sum_{i=1}^{h_q} f_i + \sum_{i=1}^{h_q - q} d_i, \quad q = 1, \dots, l-1,$$

and

$$H_0 := 0.$$

Note that from (15), we have that $H_q \geqslant 0$, $q = 1, \dots, l-1$.

Let

$$S_q := \sum_{i=h_{q-1}-q+2}^{h_q-q} d_i - \sum_{i=h_{q-1}+1}^{h_q} f_i, \quad q = 1, \dots, l-1.$$

Thus

$$H_q - H_{q-1} = S_q + a_q, \quad q = 1, \dots, l-1.$$

Since $a_1 \geqslant \cdots \geqslant a_{l-1}$, we have

$$H_1 - S_1 \geqslant H_2 - H_1 - S_2 \geqslant \cdots \geqslant H_{l-1} - H_{l-2} - S_{l-1}. \tag{21}$$

Now, define the numbers

$$H'_i := \min(H_i, H_{i+1}, \dots, H_{l-1}), \quad i = 0, \dots, l-1. \tag{22}$$

Thus, we have

$$H'_1 \leqslant \cdots \leqslant H'_{l-1}, \tag{23}$$

$$H'_{l-1} = H_{l-1} \quad \text{and} \quad H'_i \leqslant H_i, \quad i = 1, \dots, l-2. \tag{24}$$

We are going to define certain integers $g'_1, \dots, g'_{\rho+l-1}$. The wanted $g_1 \geqslant \cdots \geqslant g_{\rho+l-1}$ will be defined as the nonincreasing ordering of $g'_1, \dots, g'_{\rho+l-1}$.

Let

$$g'_i := d_{i-l+1}, \quad i > h_{l-1}. \tag{25}$$

We shall split the definition of $g'_1, \dots, g'_{h_{l-1}}$ into $l-1$ groups. For arbitrary $j = 1, \dots, l-1$, we define g'_i, $i = h_{j-1}+1, \dots, h_j$, (with convention $h_0 := 0$) in a following way:

If

$$f_{h_j} \geqslant H'_j - H'_{j-1} - S_j, \tag{26}$$

then we define $g'_{h_{j-1}+1} \geqslant \cdots \geqslant g'_{h_j-1}$ as a nonincreasing sequence of integers such that

$$d_{i-j+1} \geqslant g'_i \geqslant f_i$$

and

$$\sum_{i=h_{j-1}+1}^{h_j-1} g'_i - \sum_{i=h_{j-1}+1}^{h_j-1} f_i = H'_j - H'_{j-1}$$

(this is obviously possible because of (26)). Also, in this case, we define

$$g'_{h_j} := f_{h_j}.$$

If

$$f_{h_j} < H'_j - H'_{j-1} - S_j, \tag{27}$$

then we define

$$g'_i := d_{i-j+1}, \quad i = h_{j-1}+1, \ldots, h_j - 1,$$

and

$$g'_{h_j} := H'_j - H'_{j-1} - S_j.$$

Note that in both of the previous cases, (26) and (27), we have

$$\sum_{i=h_{j-1}+1}^{h_j} g'_i - \sum_{i=h_{j-1}+1}^{h_j} f_i = H'_j - H'_{j-1}, \quad j = 1, \ldots, l-1. \tag{28}$$

and

$$g'_{h_i} = \max(f_{h_i}, H'_i - H'_{i-1} - S_i), \quad i = 1, \ldots, l-1.$$

Now, let $i \in \{1, \ldots, l-2\}$.

If $g'_{h_{i+1}} = f_{h_{i+1}}$, then $g'_{h_{i+1}} \leqslant f_{h_i} \leqslant g'_{h_i}$.

If $g'_{h_{i+1}} = H'_{i+1} - H'_i - S_{i+1} > f_{h_{i+1}}$, then, from (28), we have that $H'_{i+1} > H'_i$, and so $H'_i = H_i$. However, this together with (21), gives

$$\begin{aligned} g'_{h_{i+1}} &= H'_{i+1} - H'_i - S_{i+1} \leqslant H_{i+1} - H'_i - S_{i+1} = H_{i+1} - H_i - S_{i+1} \\ &\leqslant H_i - H_{i-1} - S_i = H'_i - H_{i-1} - S_i \leqslant H'_i - H'_{i-1} - S_i \leqslant g'_{h_i}. \end{aligned}$$

Hence, we have

$$g'_{h_1} \geqslant g'_{h_2} \geqslant \cdots \geqslant g'_{h_{l-1}}. \tag{29}$$

Also, from the definition of h_i, $i = 1, \ldots, l-1$, the subsequence of g'_i's for $i \in \{1, \ldots, \rho + l - 1\} \setminus \{h_1, \ldots, h_{l-1}\}$ is in nonincreasing order, and satisfies:

$$d_{i-j+1} \geqslant g'_i \geqslant f_i, \quad h_{j-1} < i < h_j, \quad j = 1, \ldots, l. \tag{30}$$

For $i \geqslant h_l$, from (20), we have

$$d_{i-l+1} = g'_i = f_{i+1}, \quad i \geqslant h_l. \tag{31}$$

Now, since $g'_i \geqslant f_{i+1}$ for all $i = 1, \ldots, \rho + l - 1$, and since g_i's are the nonincreasing ordering of g'_i's, we have (i).

Moreover, since $g'_{h_{l-1}} \geqslant f_{h_{l-1}} > d_{h_{l-1}-l+2} = g'_{h_{l-1}+1}$, we have that $g_i = g'_i$, for $i > h_{l-1}$. Then, from (30), we have $g_i \geqslant f_i$, for $i < h_l$, which together with $g_{h_l} = g'_{h_l} = d_{h_l-l+1} < f_{h_l}$, implies $h = h_l$. Thus, (31) implies (iii).

If we denote by $\nu_1 \geqslant \cdots \geqslant \nu_\rho$ the subsequence of g'_i's for $i \in \{1, \ldots, \rho + l - 1\} \setminus \{h_1, \ldots, h_{l-1}\}$, then from (30) and (31) we have

$$d_i \geqslant \nu_i, \quad i = 1, \ldots, \rho, \tag{32}$$

which implies (ii).

Also, by summing all inequalities from (28), for $j = 1, \ldots, l-1$, we have

$$\sum_{i=1}^{h_{l-1}} g'_i - \sum_{i=1}^{h_{l-1}} f_i = H'_{l-1},$$

which together with (24) and the definition of H_{l-1}, gives

$$\sum_{i=1}^{h_{l-1}} g'_i - \sum_{i=1}^{h_{l-1}-l+1} d_i = \sum_{i=1}^{l-1} a_i.$$

The last equation, together with the definition of the remaining g'_i's (25), the fact that $\sum_{i=1}^{\rho+l-1} g_i = \sum_{i=1}^{\rho+l-1} g'_i$, and (14), gives (v).

Before going to the proof of (iv), we shall establish some relations between h_q's and $\tilde{h}_q$'s. So, let $q \in \{1, \ldots, l-1\}$. The sequence of g_i's is defined as the nonincreasing ordering of g'_i's. As we have shown, the sequence of g'_i's is the union of two nonincreasing sequences: $g'_{h_1} \geqslant g'_{h_2} \geqslant \ldots \geqslant g'_{h_{l-1}}$ and $\nu_1 \geqslant \nu_2 \geqslant \ldots \geqslant \nu_\rho$.

Let r_q be the index such that

$$\nu_{r_q} \geqslant g'_{h_q} > \nu_{r_q+1}.$$

First of all, from the definition of g'_{h_q} and h_q, we have that $g'_{h_q} \geqslant f_{h_q} > d_{h_q-q+1} \geqslant \nu_{h_q-q+1}$, and so

$$r_q \leqslant h_q - q. \tag{33}$$

Furthermore, the subsequence $g_1 \geqslant g_2 \geqslant \ldots \geqslant g_{r_q+q}$ is the nonincreasing ordering of the union of sequences $g'_{h_1} \geqslant g'_{h_2} \geqslant \ldots \geqslant g'_{h_q}$ and $\nu_1 \geqslant \nu_2 \geqslant \ldots \geqslant \nu_{r_q}$, with g'_{h_q} being the smallest among them, i.e. $g_{r_q+q} = g'_{h_q}$. Thus, $\nu_i \geqslant g_{i+q-1}$, for $i = 1, \ldots, r_q$, and so from (32), for every $i \leqslant r_q$ we have that $d_i \geqslant \nu_i \geqslant g_{i+q-1}$, i.e.

$$\tilde{h}_q \geqslant r_q + q. \tag{34}$$

By (33), we have two possibilities for r_q:

If $r_q = h_q - q$, as proved above, we have $g_{h_q} = g'_{h_q}$, which then implies $g_{h_q} \geqslant f_{h_q} > d_{h_q-q+1} \geqslant \nu_{h_q-q+1}$, and so $\tilde{h}_q \leqslant h_q$, which together with (34) in this case gives $\tilde{h}_q = h_q = r_q + q$.

If $r_q < h_q - q$, then $g'_{h_q} > \nu_{h_q-q} \geqslant f_{h_q}$, and so from the definition of g'_i's, we have that $\nu_i = d_i$, for $i = r_q + 1, \ldots, h_q - q$. Thus $g_{r_q+q} = g'_{h_q} > \nu_{r_q+1} = d_{r_q+1}$, and so $\tilde{h}_q \leqslant r_q + q$, which together with (34) gives $\tilde{h}_q = r_q + q$.

Thus, altogether we have that $\tilde{h}_q \leqslant h_q$, and $g_1 \geqslant g_2 \geqslant \ldots \geqslant g_{\tilde{h}_q}$ is the nonincreasing ordering of the union of sequences $g'_{h_1} \geqslant g'_{h_2} \geqslant \ldots \geqslant g'_{h_q}$ and $\nu_1 \geqslant \nu_2 \geqslant \ldots \geqslant \nu_{\tilde{h}_q-q}$, with $g_{\tilde{h}_q} = g'_{h_q}$, and that $\tilde{h}_q < h_q$ implies $\nu_i = d_i$, for $i = \tilde{h}_q - q + 1, \ldots, h_q - q$.

Finally, we can pass to the proof of (iv). Let $q \in \{1, \dots, l-1\}$. We shall prove (iv) for this q in the following equivalent form

$$\sum_{i=1}^{\tilde{h}_q} \tilde{g}_i - \sum_{i=1}^{\tilde{h}_q - q} d_i \leqslant H_q + \sum_{i=1}^{h_q} f_i - \sum_{i=1}^{h_q - q} d_i. \tag{35}$$

If $\tilde{h}_q = h_q$, (35) is equivalent to

$$\sum_{i=1}^{h_q} (g'_i - f_i) \leqslant H_q, \tag{36}$$

which follows from (24) and (28).

If $\tilde{h}_q < h_q$, we have that $\nu_i = d_i$, for $i = \tilde{h}_q - q + 1, \dots, h_q - q$. Hence, the condition (35) is again equivalent to (36), which concludes our proof.

■

By iterating the previous result, we obtain the following

Theorem 5 *Let f, d and a be the sequences from Definition 1. If*

$$f \prec' (d, a),$$

then there exist sequences of integers $g^j = (g^j_1, \dots, g^j_{\rho+j})$, $j = 1, \dots, l-1$, with $g^j_1 \geqslant \dots \geqslant g^j_{\rho+j}$, such that

$$g^j \prec'_1 (g^{j-1}, a_j), \quad j = 1, \dots, l,$$

where $g^0 = d$ and $g^l = f$.

Proof: For $l = 1$, the claim of theorem follows trivially.

Let $l > 1$, and suppose that theorem holds for $l - 1$. By Lemma 4, there exists a sequence $g = (g_1, \dots, g_{\rho+l-1})$, such that $g_1 \geqslant \dots \geqslant g_{\rho+l-1}$ and such that they satisfy conditions $(i) - (v)$ from Lemma 4. Set $g^{l-1} := g$. From (i), (iii) and (v) we have

$$f \prec'_1 (g^{l-1}, a_l). \tag{37}$$

From (ii), (iv) and (v), we have

$$g^{l-1} \prec' (d, a'), \tag{38}$$

where $a' = (a_1, \dots, a_{l-1})$.

By induction hypothesis there exist sequences $g^1, \dots, g^{l-2}$, such that

$$g^j \prec'_1 (g^{j-1}, a_j), \quad j = 1, \dots, l-1.$$

This together with (37) finishes our proof. ■

The following two results give converse of Lemma 4 and Theorem 5:

Lemma 6 *Let* $d_1 \geqslant \cdots \geqslant d_\rho$, $f_1 \geqslant \cdots \geqslant f_{\rho+l}$ *and* $g_1 \geqslant \cdots \geqslant g_{\rho+1}$ *be integers. Let* a'_1 *and* $a'_2 \geqslant \cdots \geqslant a'_l$ *be integers. Let* $a_1 \geqslant \cdots \geqslant a_l$ *be integers such that*

$$(a'_1, a'_2, \ldots, a'_l) \prec (a_1, a_2, \ldots, a_l). \tag{39}$$

If

$$\begin{array}{rl} (i) & d_i \geqslant g_{i+1}, \quad i = 1, \ldots, \rho, \\ (ii) & g_i \geqslant f_{i+l-1}, \quad i = 1, \ldots, \rho+1, \\ (iii) & d_i = g_{i+1}, \quad i \geqslant \bar{h}_1, \quad \textit{where } \bar{h}_1 = \min\{i | d_i < g_i\}, \\ (iv) & \sum_{i=1}^{\tilde{h}_q} f_i - \sum_{i=1}^{\tilde{h}_q - q} g_i \leqslant \sum_{i=2}^{q+1} a'_i, \quad q = 1, \ldots, l-1, \quad \textit{where } \tilde{h}_q = \min\{i | g_{i-q+1} < f_i\}, \\ (v) & \sum_{i=1}^{\rho+l} f_i = \sum_{i=1}^{\rho+1} g_i + \sum_{i=2}^{l} a'_i = \sum_{i=1}^{\rho} d_i + \sum_{i=1}^{l} a'_i, \end{array}$$

then

$$\sum_{i=1}^{h_q} f_i - \sum_{i=1}^{h_q - q} d_i \leqslant \sum_{i=1}^{q} a_i, \quad q = 1, \ldots, l, \tag{40}$$

where $h_q = \min\{i | d_{i-q+1} < f_i\}$, $q = 1, \ldots, l$.

Proof: From the definition of h_q, $\tilde{h}_q$ and $\bar{h}_1$, we obtain the following inequalities

$$h_q \geqslant \max(\tilde{h}_{q-1}, \min(\bar{h}_1 + q - 1, \tilde{h}_q)), \quad q = 1, \ldots, l-1, \ (\tilde{h}_0 = 0), \tag{41}$$

and

$$h_l \geqslant \max(\tilde{h}_{l-1}, \bar{h}_1 + l - 1). \tag{42}$$

This is true since for $q = 1, \ldots, l-1$, and $j < \min(\bar{h}_1 + q - 1, \tilde{h}_q)$, we have that

$$d_{j-q+1} \geqslant g_{j-q+1} \geqslant f_j.$$

Therefore, $h_q \geqslant \min(\bar{h}_1 + q - 1, \tilde{h}_q)$. Also, for every $q = 1, \ldots, l$, and $j < \tilde{h}_{q-1}$, we have $d_{j-q+1} \geqslant g_{j-q+2} \geqslant f_j$, which gives $h_q \geqslant \tilde{h}_{q-1}$. Furthermore, for every $j < \bar{h}_1 + l - 1$, we have $d_{j-l+1} \geqslant g_{j-l+1} \geqslant f_j$, and so $h_l \geqslant \bar{h}_1 + l - 1$. Altogether, we have (41) and (42).

Let $q \in \{1, \ldots, l-1\}$. From (41), we have the following three possibilities on h_q:

$$\begin{array}{ll} a) & h_q \geqslant \tilde{h}_q, \text{ in the case } \tilde{h}_q \leqslant \bar{h}_1 + q - 1, \\ b) & \tilde{h}_q > h_q \geqslant \max(\tilde{h}_{q-1}, \bar{h}_1 + q - 1) \text{ if } \tilde{h}_q > \bar{h}_1 + q - 1, \\ c) & h_q \geqslant \tilde{h}_q > \max(\tilde{h}_{q-1}, \bar{h}_1 + q - 1) \text{ if } \tilde{h}_q > \bar{h}_1 + q - 1. \end{array}$$

Observe these cases separately:

$a)$ Let $h_q \geqslant \tilde{h}_q$ ($\tilde{h}_q \leqslant \bar{h}_1 + q - 1$), then by (iv) we have

$$\sum_{i=1}^{h_q} f_i = \sum_{i=1}^{\tilde{h}_q} f_i + \sum_{\tilde{h}_q+1}^{h_q} f_i \leqslant \sum_{i=1}^{\tilde{h}_q-q} g_i + \sum_{\tilde{h}_q+1}^{h_q} f_i + \sum_{i=2}^{q+1} a'_i$$
$$\leqslant \sum_{i=1}^{\tilde{h}_q-q} d_i + \sum_{\tilde{h}_q-q+1}^{h_q-q} d_i + \sum_{i=2}^{q+1} a'_i = \sum_{i=1}^{h_q-q} d_i + \sum_{i=2}^{q+1} a'_i.$$

The second inequality is true since $\tilde{h}_q - q < \bar{h}_1$. So, we have $d_i \geqslant g_i$ for all $i \leqslant \tilde{h}_q - q$. Also, from $h_q < h_{q+1}$, we obtain $f_i \leqslant d_{i-q}$, for all $i \leqslant h_q < h_{q+1}$.

Finally, from (39), we have

$$\sum_{i=2}^{q+1} a'_i \leqslant \sum_{i=1}^{q} a_i,$$

and so

$$\sum_{i=1}^{h_q} f_i \leqslant \sum_{i=1}^{h_q-q} d_i + \sum_{i=1}^{q} a_i,$$

which proves (40), as wanted.

$b)$ Let $\tilde{h}_q > h_q \geqslant \max(\bar{h}_1 + q - 1, \tilde{h}_{q-1})$, then by (iv), we have

$$\sum_{i=1}^{h_q} f_i = \sum_{i=1}^{\tilde{h}_{q-1}} f_i + \sum_{\tilde{h}_{q-1}+1}^{h_q} f_i \leqslant \sum_{i=1}^{\tilde{h}_{q-1}-q+1} g_i + \sum_{\tilde{h}_{q-1}+1}^{h_q} f_i + \sum_{i=2}^{q} a'_i$$
$$\leqslant \sum_{i=1}^{\tilde{h}_{q-1}-q+1} g_i + \sum_{\tilde{h}_{q-1}-q+2}^{h_q-q+1} g_i + \sum_{i=2}^{q} a'_i.$$

The second inequality is true, since $h_q < \tilde{h}_q$, and so, $g_{i-q+1} \geqslant f_i$, for all $i \leqslant h_q$.

Moreover, since $h_q - q + 1 \geqslant \bar{h}_1$, by conditions (iii) and (v), we have

$$\sum_{i=1}^{h_q-q+1} g_i = \sum_{i=1}^{\rho+1} g_i - \sum_{i=h_q-q+2}^{\rho+1} g_i = \sum_{i=1}^{\rho} d_i + a'_1 - \sum_{i=h_q-q+1}^{\rho} d_i = \sum_{i=1}^{h_q-q} d_i + a'_1,$$

and so

$$\sum_{i=1}^{h_q-q+1} g_i + \sum_{i=2}^{q} a'_i = \sum_{i=1}^{h_q-q} d_i + \sum_{i=1}^{q} a'_i.$$

Last equality together with (39) gives

$$\sum_{i=1}^{h_q} f_i \leqslant \sum_{i=1}^{h_q-q} d_i + \sum_{i=1}^{q} a_i,$$

which proves (40), as wanted.

c) Let $h_q \geqslant \tilde{h}_q > \max(\bar{h}_1 + q - 1, \tilde{h}_{q-1})$, then by (iv), we have

$$\begin{aligned}
\sum_{i=1}^{h_q} f_i = \sum_{i=1}^{\tilde{h}_{q-1}} f_i + \sum_{\tilde{h}_{q-1}+1}^{h_q} f_i &\leqslant \sum_{i=1}^{\tilde{h}_{q-1}-q+1} g_i + \sum_{\tilde{h}_{q-1}+1}^{h_q} f_i + \sum_{i=2}^{q} a'_i \\
&= \sum_{i=1}^{\tilde{h}_{q-1}-q+1} g_i + \sum_{\tilde{h}_{q-1}+1}^{\tilde{h}_q-1} f_i + \sum_{\tilde{h}_q}^{h_q} f_i + \sum_{i=2}^{q} a'_i \\
&\leqslant \sum_{i=1}^{\tilde{h}_{q-1}-q+1} g_i + \sum_{\tilde{h}_{q-1}-q+2}^{\tilde{h}_q-q} g_i + \sum_{\tilde{h}_q-q}^{h_q-q} d_i + \sum_{i=2}^{q} a'_i \\
&= \sum_{i=1}^{h_q-q} d_i + a'_1 + \sum_{i=2}^{q} a'_i.
\end{aligned}$$

The second inequality follows from the definition of $\tilde{h}_q$ and the fact that $h_q < h_{q+1}$, while the last equality is true since $\tilde{h}_q - q \geqslant \bar{h}_1$. Now, we finish the proof as in the previous case.

The only remaining case is $q = l$. Let $i > h_l$. Since $h_l \geqslant \max(\tilde{h}_{l-1}, \bar{h}_1 + l - 1)$, we have $i > \tilde{h}_{l-1}$. From (ii), (iv) and (v) we have that $f \prec' (g, a'')$, where $a'' = (a'_2, a'_3, \ldots, a'_l)$, and so (see (20)) we have $f_i = g_{i-l+1}$. Also, since $i > \bar{h}_1 + l - 1$, from (iii) we have $g_{i-l+1} = d_{i-l}$, and thus

$$f_i = d_{i-l}, \quad i > h_l. \tag{43}$$

Now, by (v), condition (40) for $q = l$ is equivalent to

$$\sum_{i=h_l+1}^{\rho+l} f_i \geqslant \sum_{i=h_l-l+1}^{\rho} d_i.$$

Finally, from (i), we have that $d_i \geqslant f_{i+l}$, $i = 1, \ldots, \rho$, and so condition (40) for $q = l$ is equivalent to (43), which concludes our proof.

■

By iterating the previous result, we obtain the following one:

Theorem 7 *Let $d_1 \geqslant \cdots \geqslant d_\rho$, $f_1 \geqslant \cdots \geqslant f_{\rho+l}$, $a_1 \geqslant \cdots \geqslant a_l$ and $a'_1, \ldots, a'_l$ be integers, such that*

$$(a'_1, \ldots, a'_l) \prec (a_1, \ldots, a_l).$$

Moreover, for every $j = 1, \ldots, l-1$, let $g^j = (g^j_1, \ldots, g^j_{\rho+j})$ be such that $g^j_1 \geqslant \cdots \geqslant g^j_{\rho+j}$. Also, let $g^0 := d$, and $g^l := f$.

If $g^j \prec'_1 (g^{j-1}, a'_j)$ for $j = 1, \ldots, l$, then $f \prec' (d, a)$.

Thus, Theorems 5 and 7 prove the existence of a path of sequences, as announced before Lemma 4. In particular, we have

Corollary 8 *Let $l \geqslant 2$, $d_1 \geqslant \cdots \geqslant d_\rho$, $f_1 \geqslant \cdots \geqslant f_{\rho+l}$, $a_1 \geqslant \cdots \geqslant a_l$ be integers. Then*

$$f \prec' (d, a)$$

if and only if there exists $g = (g_1, \dots, g_{\rho+s})$, for some $0 < s < l$, such that $g_1 \geqslant \cdots \geqslant g_{\rho+s}$ and

$$f \prec' (g, a')$$

$$g \prec' (d, a'')$$

where $a' = (a_1, \dots, a_{l-s})$ and $a'' = (a_{l-s+1}, \dots, a_l)$.

4 Convexity lemma

In this section we give a short polynomial proof of the convexity lemma, which is the crucial step in Sá-Thompson theorem [7, 10]. The original proofs of Sá and Thompson were long and complicated, and relied on very involved techniques. The proof in [7] (Proposition 4.1 and Lemma 4.2) uses nonelementary analytical tools, while the proof in [10] is elementary but very long and does not involve the concept of convexity. Later on shorter, combinatorial proof was given in [8].

Here we give the first purely polynomial proof of the convexity lemma.

Let $\alpha_1 | \cdots | \alpha_n$ and $\gamma_1 | \cdots | \gamma_{n+m}$ be two polynomial chains.

For every $j = 0, \dots, m$, let

$$\delta_i^j := \operatorname{lcm}(\alpha_{i-2j}, \gamma_i), \quad i = 1, \dots, n+j,$$

$$\delta^j := \prod_{i=1}^{n+j} \delta_i^j.$$

The difference between the convexity in this case and the result from Lemma 2 is in a different shift in the definition of δ^j comparing to π_j. This makes the problem much more difficult, and in particular here we do not have that $\delta^{j-1} | \delta^j$. However, the convexity of the degrees of δ^j holds:

Theorem 9 *(Convexity Lemma)*

$$d(\delta^j) - d(\delta^{j-1}) \leqslant d(\delta^{j+1}) - d(\delta^j), \text{ for } j = 1, \dots, m-1.$$

Before going to the proof we give one simple lemma:

Lemma 10 *Let ϕ_1, ϕ_2, ψ_1 and ψ_2 be polynomials such that $\phi_1|\phi_2$ and $\psi_1|\psi_2$. Then*

$$\operatorname{lcm}(\phi_1,\psi_1)\operatorname{lcm}(\phi_2,\psi_2)|\operatorname{lcm}(\phi_2,\psi_1)\operatorname{lcm}(\phi_1,\psi_2). \tag{44}$$

Proof: For $i = 1, 2$, we have

$$\operatorname{lcm}(\phi_i,\psi_2) = \operatorname{lcm}(\phi_i,\psi_1,\psi_2) = \operatorname{lcm}(\operatorname{lcm}(\phi_i,\psi_1),\psi_2) = \frac{\operatorname{lcm}(\phi_i,\psi_1)\psi_2}{\gcd(\operatorname{lcm}(\phi_i,\psi_1),\psi_2)}.$$

Now, by replacing this expression for $i = 1$ and $i = 2$ into (44), it becomes equivalent to the following obvious divisibility relation:

$$\gcd(\operatorname{lcm}(\phi_1,\psi_1),\psi_2) \mid \gcd(\operatorname{lcm}(\phi_2,\psi_1),\psi_2).$$

■

Proof of Theorem 9:
In order to prove the convexity, it is enough to prove that

$$\delta^j\delta^j \mid \delta^{j-1}\delta^{j+1}, \quad j = 1,\dots,m-1. \tag{45}$$

By definition, we have

$$\delta^j = \prod_{i=1}^{n+j}\operatorname{lcm}(\alpha_{i-2j},\gamma_i), \quad j = 0,\dots,m. \tag{46}$$

Since for all i and j we have

$$\operatorname{lcm}(\alpha_{i-2j},\gamma_i) = \operatorname{lcm}(\alpha_{i-2j},\operatorname{lcm}(\alpha_{i-2j-2},\gamma_i)) = \frac{\alpha_{i-2j}\operatorname{lcm}(\alpha_{i-2j-2},\gamma_i)}{\gcd(\alpha_{i-2j},\operatorname{lcm}(\alpha_{i-2j-2},\gamma_i))},$$

we can rewrite (46) as

$$\delta^j = \prod_{i=1}^{n+j}\frac{\alpha_{i-2j}\operatorname{lcm}(\alpha_{i-2j-2},\gamma_i)}{\gcd(\alpha_{i-2j},\operatorname{lcm}(\alpha_{i-2j-2},\gamma_i))} = \frac{\prod_{i=1}^{n-j}\alpha_i\prod_{i=1}^{n+j}\operatorname{lcm}(\alpha_{i-2j-2},\gamma_i)}{\prod_{i=1}^{n-j}\gcd(\alpha_i,\operatorname{lcm}(\alpha_{i-2},\gamma_{i+2j}))}. \tag{47}$$

We replace one δ^j on the left hand side and δ^{j+1} on the right hand side of (45) by the expression (46), while we replace the other δ^j and δ^{j-1} by the expression (47). Then (45) becomes equivalent to

$$\frac{\prod_{i=1}^{n+j}\operatorname{lcm}(\alpha_{i-2j},\gamma_i)\prod_{i=1}^{n-j}\alpha_i\prod_{i=1}^{n+j}\operatorname{lcm}(\alpha_{i-2j-2},\gamma_i)}{\prod_{i=1}^{n-j}\gcd(\alpha_i,\operatorname{lcm}(\alpha_{i-2},\gamma_{i+2j}))}$$
$$\mid \frac{\prod_{i=1}^{n+j+1}\operatorname{lcm}(\alpha_{i-2j-2},\gamma_i)\prod_{i=1}^{n-j+1}\alpha_i\prod_{i=1}^{n+j-1}\operatorname{lcm}(\alpha_{i-2j},\gamma_i)}{\prod_{i=1}^{n-j+1}\gcd(\alpha_i,\operatorname{lcm}(\alpha_{i-2},\gamma_{i+2j-2}))}.$$

After cancellations, the last divisibility becomes equivalent to

$$\mathrm{lcm}(\alpha_{n-j},\gamma_{n+j})\prod_{i=1}^{n-j+1}\gcd(\alpha_i,\mathrm{lcm}(\alpha_{i-2},\gamma_{i+2j-2}))$$

$$\mid \mathrm{lcm}(\alpha_{n-j-1},\gamma_{n+j+1})\alpha_{n-j+1}\prod_{i=1}^{n-j}\gcd(\alpha_i,\mathrm{lcm}(\alpha_{i-2},\gamma_{i+2j})).$$

By using the obvious divisibility relation

$$\gcd(\alpha_i,\mathrm{lcm}(\alpha_{i-2},\gamma_{i+2j-2}))\mid\gcd(\alpha_i,\mathrm{lcm}(\alpha_{i-2},\gamma_{i+2j})),$$

we are left with proving that

$$\mathrm{lcm}(\alpha_{n-j},\gamma_{n+j})\gcd(\alpha_{n-j+1},\mathrm{lcm}(\alpha_{n-j-1},\gamma_{n+j-1}))\mid\alpha_{n-j+1}\,\mathrm{lcm}(\alpha_{n-j-1},\gamma_{n+j+1}). \quad (48)$$

However, since

$$\gcd(\alpha_{n-j+1},\mathrm{lcm}(\alpha_{n-j-1},\gamma_{n+j-1}))=\frac{\alpha_{n-j+1}\,\mathrm{lcm}(\alpha_{n-j-1},\gamma_{n+j-1})}{\mathrm{lcm}(\alpha_{n-j+1},\gamma_{n+j-1})},$$

(48) becomes equivalent to the following

$$\mathrm{lcm}(\alpha_{n-j-1},\gamma_{n+j-1})\,\mathrm{lcm}(\alpha_{n-j},\gamma_{n+j})\mid\mathrm{lcm}(\alpha_{n-j-1},\gamma_{n+j+1})\,\mathrm{lcm}(\alpha_{n-j+1},\gamma_{n+j-1}),$$

which follows directly from Lemma 10. ■

References

[1] I. Baragaña, I. Zaballa, Column completion of a pair of matrices, Linear and Multilinear Algebra, 27 (1990) 243-273.

[2] M. Dodig, Matrix pencils completion problems, Linear Algebra Appl. 428 (2008), no. 1, 259-304.

[3] M.Dodig, M. Stošić, Similarity class of a matrix with prescribed submatrix, Linear and Multilinear Algebra, 57 (2009) 217-245.

[4] M. Dodig, Feedback invariants of matrices with prescribed rows, Linear Algebra Appl. 405 (2005) 121-154.

[5] I. Gohberg, M. A. Kaashoek, F. van Schagen, Eigenvalues of completions of submatrices, Linear and Multilinear Algebra, 25 (1989) 55-70.

[6] G. Hardy, J.E. Littlewood, G. Pólya, Inequalities, Cambridge University Press, 1991.

[7] E. M. Sá, Imbeding conditions for λ -matrices, Linear Algebra Appl. 24 (1979) 33-50.

[8] E. M. Sá, A convexity lemma on the interlacing inequalities for invariant factors, Linear Algebra Appl. 109 (1988) 107-113.

[9] E. M. Sá, Imersão de matrizes e entrelaçamento de factores invariantes, PhD Thesis, Univ. of Coimbra, 1979.

[10] R. C. Thompson, Interlacing inequalities for invariant factors, Linear Algebra Appl. 24 (1979) 1-31.

[11] I. Zaballa, Matrices with prescribed rows and invariant factors, Linear Algebra Appl. 87 (1987) 113-146.