Fractional Hadamard powers of symmetric positive-definite matrices

P. Fischer J.D. Stegeman*

1 Hadamard powers

Let $A = (a_{ij})$ and $B = (b_{ij})$ be matrices of the same size. Then their *Hadamard product* (also called *Schur product*) $A \diamond B$ is defined by entrywise multiplication: $A \diamond B = (a_{ij} b_{ij})$. The *Hadamard unit matrix* is the matrix U all of whose entries are 1 (the size of U being understood). A matrix A is *Hadamard invertible* if all its entries are non-zero, and $A^{\diamond(-1)} = (a_{ij}^{-1})$ is then called the *Hadamard inverse* of A. If B is Hadamard invertible, then the *Hadamard quotient* $A \diamond / B$ of A and B is $(a_{ij} b_{ij}^{-1})$. The k-fold Hadamard product $A^{\diamond k}$ of A with itself ($k \geq 0$) is called the k-th *Hadamard power* of A; thus $(a_{ij})^{\diamond k} = (a_{ij}^k)$. In particular, $A^{\diamond 0} = U$ (conventionally we set $0^0 = 1$). If A is Hadamard invertible, then $A^{\diamond k}$ can be defined for negative integers as well, in an obvious manner. For more information on the Hadamard product, see [7, Chapter 5] and [5].

In this paper we restrict our attention to *real* matrices. If all entries of A are non-negative, then we can also consider *fractional* Hadamard powers of A: if $A = (a_{ij})$ with $a_{ij} \geq 0$ for all i, j, and $\alpha \in \mathbb{R}$, $\alpha \geq 0$, then $A^{\diamond\alpha} = (a_{ij}^{\alpha})$. If $a_{ij} > 0$ for all i, j, then $A^{\diamond\alpha}$ can be defined for all $\alpha \in \mathbb{R}$. Note that $(A \diamond B)^{\diamond\alpha} = A^{\diamond\alpha} \diamond B^{\diamond\alpha}$ whenever all these fractional Hadamard powers are defined.

Matrices of size $n \times 1$ ('column vectors') will be identified with elements of $\mathbb{R}^n$. In this way the Hadamard product $v \diamond w$, Hadamard powers $v^{\diamond k}$ ($k \geq 0$), and fractional Hadamard powers $v^{\diamond\alpha}$ (for suitable $\alpha \in \mathbb{R}$) are defined for $v, w \in \mathbb{R}^n$. The Hadamard unit vector is the vector u all of whose entries are 1. The transpose of a vector v will be denoted by v^*.

There is an interesting difference between matrices of rank one and matrices of higher rank: if a matrix has rank one, then the same holds for all its existing fractional Hadamard powers; but if the rank is at least two and there are no evident obstructions (such as a row of zeros or two equal rows), then almost all fractional Hadamard powers will have maximal rank.

2 Positive-definite matrices

A (real or complex) $n \times n$ matrix A will be called *positive-definite* (many authors use, more accurately, positive *semi*-definite) if

$$v^* A v \geq 0 \quad \text{for all } v \in \mathbb{C}^n\,, \tag{1}$$

where v^* is the transpose of the complex-conjugate of v. A *real* matrix A is positive-definite (in the above sense) if and only if A is *symmetric* and satisfies

$$v^* A v \geq 0 \quad \text{for all } v \in \mathbb{R}^n\,. \tag{2}$$

It is customary to call such real matrices *symmetric* positive-definite.

As in [2], we shall write $\mathcal{S}_n$ for the set of all (real) symmetric positive-definite $n \times n$ matrices, and $\mathcal{S}_n^+$ for those matrices in $\mathcal{S}_n$ for which all entries are non-negative. The following proposition is a fundamental result of Schur (1911) (cf. [6, 7.5.3]).

*The research of the second author was supported by a travel grant of the Netherlands Organization for Scientific Research (NWO).

Proposition 2.1 *Suppose that $A, B \in \mathcal{S}_n$. Then also $A \diamond B \in \mathcal{S}_n$. In particular, if $A \in \mathcal{S}_n$, then $A^{\diamond k} \in \mathcal{S}_n$ for all non-negative integers k.* □

Let $A = (a_{ij})$ be an $m \times n$ matrix. For $\emptyset \neq \lambda \subset \{1, \ldots, m\}$ and $\emptyset \neq \mu \subset \{1, \ldots, n\}$, the submatrix $A_{\lambda\mu}$ is defined as $A_{\lambda\mu} = (a_{ij})_{i \in \lambda,\, j \in \mu}$. If $m = n$ and $\emptyset \neq \lambda \subset \{1, \ldots, n\}$, then $A_{\lambda\lambda}$ is called a *principal* submatrix of A. The determinant of a square submatrix is called a *minor*, of a principal submatrix a *principal minor* of A (cf. [6, section 0.7.1]). In the next proposition another fundamental fact from the theory of symmetric positive-definite matrices is given (cf. [6, 7.1.2 and 7.2.5]).

Proposition 2.2 *A symmetric real $n \times n$ matrix A is positive-definite if and only if all its principal minors are non-negative.* □

Consider the following problem: for a matrix $A \in \mathcal{S}_n^+$ and a non-negative real number α, when will $A^{\diamond\alpha}$ again belong to $\mathcal{S}_n^+$? For integer values of α this is always the case, by Proposition 2.1. We can also make the following observation.

Proposition 2.3 *Let A be an $n \times n$ matrix with non-negative entries. If there exist arbitrarily small $\alpha > 0$ such that $A^{\diamond\alpha} \in \mathcal{S}_n$, then $A^{\diamond\alpha} \in \mathcal{S}_n$ for all $\alpha \geq 0$.*

Proof Let X be the set of all $\alpha \geq 0$ for which $A^{\diamond\alpha} \in \mathcal{S}_n^+$. From Proposition 2.2 it follows that X is closed, while Proposition 2.1 implies that X is dense in $[0, \infty)$. Thus $X = [0, \infty)$. □

Remark 2.4 Proposition 2.3 is of some interest in connection with the following criterion, mentioned in [5, p. 144]. There a matrix $A \in \mathcal{S}_n^+$ such that $A^{\diamond\alpha} \in \mathcal{S}_n^+$ for all $\alpha > 0$ is called *infinitely divisible*. Further, an $n \times n$ matrix A is called *conditionally* positive-definite if relation (2) holds for all $v \in \mathbb{C}^n$ for which $u^* v = 0$ (i.e. all v whose entries sum to zero). It is then shown that if $A \in \mathcal{S}_n^+$ has strictly positive entries, then A is infinitely divisible if and only if $\log^\diamond(A)$ (the entrywise logarithm of A) is conditionally positive-definite.

It is known that $A^{\diamond\alpha} \in \mathcal{S}_n^+$ will hold for *all* $A \in \mathcal{S}_n^+$ if and only if α is an integer or $\alpha > n - 2$; cf. [5, p. 144]. The necessity of the condition follows from a more general result proved by Horn in 1969 [4, Theorem 1.2 and Corollary 1.3]. The sufficiency was proved in 1977, by an inductive argument, in [3, Theorem 2.2]. In that same paper (p. 636) the necessity is proved again: an explicit example is constructed, for $\alpha < n - 2$ (and α not an integer, cf. Proposition 2.1) of a matrix $A \in \mathcal{S}_n^+$, of rank 2 and very close to the Hadamard unit matrix U, for which $A^{\diamond\alpha} \notin \mathcal{S}_n$. This example inspired us to do the research of the present paper; cf. Example 8.4.

In fact, we study the following question: if a matrix $T \in \mathcal{S}_n^+$ of rank 1 and with strictly positive entries is approximated by matrices of the form $T + \varepsilon V$, with $V \in \mathcal{S}_n$, when will $(T + \varepsilon V)^{\diamond\alpha}$ have a negative determinant for sufficiently small $\varepsilon > 0$? In other words, we approach T along straight lines in $T + \mathcal{S}_n$, the *positive-definite cone* at T, and study the behaviour of the determinant of the α-th power. Some restriction in the choice of the matrices V will be necessary, though, for the following reason. As was pointed out in Section 1, if $A \in \mathcal{S}_n^+$ has rank at least 2, then in general the rank of $A^{\diamond\alpha}$ will be maximal for most values of $\alpha > 0$. However, for matrices A whose elements are "not in general position" it will happen that no fractional power is of maximal rank. For instance, if A has a row of zeros, or if two of its rows are equal, then the same will be the case for $A^{\diamond\alpha}$. To avoid such exceptional matrices we shall introduce the notion of *Hadamard independent* matrices (Definition 5.12; see also Example 5.13). In this connection we introduce the notion of a *cloud* to represent a symmetric positive-definite matrix.

Theorems 7.3 and 7.6 provide an answer to the above question. Of course, the answer depends on the value of α, but in fact, as we shall see, only on $[\alpha]$, i.e. on the interval between two consecutive integers to which α belongs. It further turns out that the sign of the determinant (for $\varepsilon > 0$ small enough) depends only on the value p of the rank of V in relation to the size n of the matrix. For instance, if the approximation is done with a matrix V of rank p, then for $0 < \alpha < 1$ the determinant of the α-th power of $T + \varepsilon V$ will be positive for sufficiently small $\varepsilon > 0$ when $p = n$ or $p = n - 1$, but negative if $p = n - 2$.

3 Clouds

Consider an element $v \in \mathbb{R}^n$. The matrix product vv^* is an element of $\mathcal{S}_n$ (this follows from (2)) of rank 1. Precisely: if $v = (x_i)_i$, then $vv^* = (x_i x_j)_{i,j}$. In particular, one has $uu^* = U$.

Conversely, each $A \in \mathcal{S}_n$ of rank 1 can be written as $A = vv^*$. To see this, one can argue as follows. Let $A = (a_{ij})$ have rank 1. Then $a_{ii} \geq 0$ for all i (Proposition 2.2) and $a_{ij} = a_{ji} = \mathrm{sgn}(a_{ij})\sqrt{a_{ii}a_{jj}}$ (because of the rank). Assume, without loss of generality, that $a_{11} \neq 0$ (not all a_{ii} can be zero, again because of the rank). For arbitrary i, j one has $a_{ij} = a_{i1}a_{1j}a_{11}^{-1}$ (because of the rank), hence $\mathrm{sgn}(a_{ij}) = \mathrm{sgn}(a_{i1}) \cdot \mathrm{sgn}(a_{j1})$. Now take $v = \pm(\mathrm{sgn}(a_{i1})\sqrt{a_{ii}})_i$ (two possible choices). Then the i,j-th element of vv^* is $\mathrm{sgn}(a_{i1})\mathrm{sgn}(a_{j1})\sqrt{a_{ii}a_{jj}} = \mathrm{sgn}(a_{ij})\sqrt{a_{ii}a_{jj}} = a_{ij}$, as desired. This reasoning corrects an inaccuracy in [2, Section 2.4].

Now let $v_1, \ldots, v_p$ be linearly independent elements of $\mathbb{R}^n$. Then

$$A = v_1 v_1^* + \cdots + v_p v_p^* \tag{3}$$

is a symmetric positive-definite matrix of rank p (cf. [6, Theorem 7.5.2]). Conversely, any matrix $A \in \mathcal{S}_n$ of rank p can be written as in (3). For instance, for $A = (a_{ij})$ one can take $v_1 = \left(\dfrac{a_{i1}}{\sqrt{a_{11}}}\right)_i$ (assuming that $a_{11} \neq 0$; otherwise start the procedure with another column of A). Then $v_1 v_1^* = \left(a_{i1}a_{1j}a_{11}^{-1}\right)_{ij}$ (recall that $a_{1j} = a_{j1}$ and that $a_{11} > 0$), and hence the first row and the first column of $A - v_1 v_1^*$ are zero. One can therefore continue by choosing v_2 with its first coordinate equal to 0, etc. Of course, this procedure is not unique, and the question arises when two representations as in (3) give the same matrix. The answer is given in the following theorem.

Theorem 3.1 *Let $v_1, \ldots, v_p$ and $w_1, \ldots, w_p$ be two sets of linearly independent vectors in $\mathbb{R}^n$. Then the matrices $V = v_1 v_1^* + \cdots + v_p v_p^*$ and $W = w_1 w_1^* + \cdots + w_p w_p^*$ will satisfy $V = W$ if and only if there exists an orthogonal $p \times p$ matrix $S = (\sigma_{ij})$ such that*

$$w_i = \sum_{j=1}^{p} \sigma_{ij} v_j \qquad (1 \leq i \leq p)\,.$$

Proof Write $v_k = (x_{ik})_{1 \leq i \leq n}$ $(1 \leq k \leq p)$. Then $V = (y_{ij})$ with $y_{ij} = \sum_{k=1}^{p} x_{ik}x_{jk}$. Now define vectors $\widetilde{v}_i = (x_{ik})_{1 \leq k \leq p} \in \mathbb{R}^p$ $(1 \leq i \leq n)$. Then $y_{ij} = \widetilde{v}_i \cdot \widetilde{v}_j$, the dot denoting the usual inner product. Moreover, the vectors $\widetilde{v}_i$ span $\mathbb{R}^p$, because the row rank of the matrix (x_{ik}) equals its column rank, which is p. Now an orthogonal transformation of $\mathbb{R}^p$ transforms the system of the n vectors $\widetilde{v}_i$ into a similar system $\widetilde{w}_i$ with the same values for the inner products $\widetilde{w}_i \cdot \widetilde{w}_j$, so that the matrix W generated by the w_i is equal to V.

Conversely, if the matrices V and W are equal, then the vectors $\widetilde{v}_i$ and $\widetilde{w}_i$ $(1 \leq i \leq n)$, defined as above, are sets of vectors in $\mathbb{R}^p$ satisfying $\widetilde{v}_i \cdot \widetilde{v}_j = \widetilde{w}_i \cdot \widetilde{w}_j$ for all i, j, and both sets span all of $\mathbb{R}^p$. By bilinearity it follows that for arbitrary scalars $a_1, \ldots, a_n$ the vectors $a_1\widetilde{v}_1 + \cdots + a_n\widetilde{v}_n$ and $a_1\widetilde{w}_1 + \cdots + a_n\widetilde{w}_n$ have the same length. In particular, if a linear combination of $\widetilde{v}_1, \ldots, \widetilde{v}_n$ equals zero, then so does the corresponding linear combination of $\widetilde{w}_1, \ldots, \widetilde{w}_n$. Now choose $i_1, \ldots, i_p$ such that the vectors $\widetilde{v}_{i_1}, \ldots, \widetilde{v}_{i_p}$ form a basis of $\mathbb{R}^p$. It follows from the observation above that the corresponding vectors $\widetilde{w}_{i_1}, \ldots, \widetilde{w}_{i_p}$ form a basis of $\mathbb{R}^p$ as well. It further follows that if a vector $\widetilde{v}_k$ $(1 \leq k \leq n)$ is written as a linear combination of the basis vectors $\widetilde{v}_{i_j}$ $(1 \leq j \leq p)$, then $\widetilde{w}_k$ is equal to the same linear combination of the vectors $\widetilde{w}_{i_j}$. The linear transformation that maps $\widetilde{v}_{i_j}$ to $\widetilde{w}_{i_j}$ $(1 \leq j \leq p)$ therefore maps each $\widetilde{v}_k$ to the corresponding $\widetilde{w}_k$ $(1 \leq k \leq n)$. Finally, it follows in the same way that this mapping preserves orthogonality and length, and hence is an orthogonal transformation. $\square$

If S is a permutation matrix, then the w_i are just a permutation of the v_i. Furthermore, the case $p = 1$ implies that if $vv^* = ww^*$, then $v = \pm w$, in accordance with the first paragraph of this section. For a more intricate case see Example 5.13.

Theorem 3.1 motivates the following definition.

Definition 3.2 *Let $n \geq 1$ be fixed. Consider ordered n-tuples $(\widetilde{v}_1, \ldots, \widetilde{v}_n)$ of elements in $\mathbb{R}^n$. Two n-tuples $(\widetilde{v}_1, \ldots, \widetilde{v}_n)$ and $(\widetilde{w}_1, \ldots, \widetilde{w}_n)$ will be called* equivalent *if there is an orthogonal transformation S of $\mathbb{R}^n$ such that $S\widetilde{v}_i = \widetilde{w}_i$ $(1 \leq i \leq n)$. The equivalence class to which $(\widetilde{v}_1, \ldots, \widetilde{v}_n)$ belongs will be denoted by $[\widetilde{v}_1, \ldots, \widetilde{v}_n]$. A class $[\widetilde{v}_1, \ldots, \widetilde{v}_n]$ will be called a* cloud of size n *(or, simply, a* cloud, *when n is understood). Each representing element $(\widetilde{v}_1, \ldots, \widetilde{v}_n)$ of a cloud is called a* positioning *of the cloud. The dimension of the linear subspace spanned by the vectors of any positioning of a cloud will be called the* dimension *of the cloud.*

Note that the dimension of a cloud is well-defined. If a cloud of size n has dimension p (with $0 \leq p \leq n$), then one can in particular consider those positionings $(\widetilde{v}_1, \ldots, \widetilde{v}_n)$ of the cloud for which all $\widetilde{v}_i$ $(1 \leq i \leq n)$ have their last $n-p$ coordinates equal to zero. These positionings may then be considered, in an evident way, as n-tuples of vectors in $\mathbb{R}^p$. And then, clearly, two n-tuples in $\mathbb{R}^p$ are equivalent (in the above sense) if and only if there is an orthogonal transformation S of $\mathbb{R}^p$ that transforms one of the n-tuples into the other. We shall occasionally call a cloud of size n and dimension at most p a cloud of size n in $\mathbb{R}^p$.

Definition 3.3 *Let C be a cloud of size n, say $C = [\widetilde{v}_1, \ldots, \widetilde{v}_n]$. The matrix $A_C \in \mathcal{S}_n$, defined by $A_C = (\widetilde{v}_i \cdot \widetilde{v}_j)_{i,j}$, will be called the (symmetric positive-definite) matrix* determined by C.

The matrix A_C is well-defined because it is independent of the chosen positioning of the cloud. In fact, if C has size n and dimension p, and $(\widetilde{v}_1, \ldots, \widetilde{v}_n)$ is a positioning of C in $\mathbb{R}^p$, then $A_C = v_1 v_1^* + \cdots + v_p v_p^*$ with $v_k = (x_{ik})_{1 \leq i \leq n}$ $(1 \leq k \leq p)$ if $\widetilde{v}_i = (x_{ik})_{1 \leq k \leq p} \in \mathbb{R}^p$ $(1 \leq i \leq n)$. The v_k are the column vectors, the $\widetilde{v}_i$ the row vectors of the $n \times p$ matrix $X = (x_{ik})$; cf. the proof of Theorem 3.1 above.

Theorem 3.1 together with the computation preceding it can now be reformulated in the following way.

Theorem 3.4 *The mapping $C \mapsto A_C$ establishes a bijective correspondence between the clouds of size n and the elements of $\mathcal{S}_n$. Moreover, the dimension of C is equal to the rank of A_C.* □

As an extension of Definition 3.3 we can now define, for a matrix $A \in \mathcal{S}_n$, the *cloud* of A to be the cloud C such that $A_C = A$.

Example 3.5 The zero matrix corresponds to the cloud $[0, \ldots, 0]$, and $(0, \ldots, 0)$ is the only positioning of this cloud. If $A \in \mathcal{S}_n$ has rank 1, then its cloud in $\mathbb{R}$ is given by an n-tuple of real numbers; such a cloud has only two positionings in $\mathbb{R}$. The cloud $[1, 1, \ldots, 1]$ (with n identical points in $\mathbb{R}$) corresponds to the Hadamard unit matrix U (of size n). The cloud of size n whose positionings are the orthonormal bases in $\mathbb{R}^n$ corresponds to the classical $n \times n$ unit matrix. A matrix $A \in \mathcal{S}_n$ will be close (in some natural sense to be specified) to a matrix in $\mathcal{S}_n$ of rank p if the elements of any positioning of its cloud are close to a p-dimensional subspace of $\mathbb{R}^n$.

In (3) the vectors $v_1, \ldots, v_p$ were required to be linearly independent, and this asssumption was used in Theorem 3.1 and for the definition of clouds. For the representation as such, however, linear independence is not necessary, and one can even take infinitely many vectors, providing that the sum (as in (3)) is convergent. Generalizing a result in [2], we prove a theorem about the determinant of such matrices. We need a definition first.

Definition 3.6 *Consider an arbitrary index set K. Let v_k, $k \in K$, be vectors in $\mathbb{R}^n$, and let λ be a subset of K with n elements (notation: $|\lambda| = n$). Then the volume of the parallelepiped spanned by the vectors v_k $(k \in \lambda)$ will be denoted by S_λ. Explicitly: if V_λ is a matrix whose columns are the v_k with $k \in \lambda$, in any order, then $S_\lambda = |\det V_\lambda|$.*

Furthermore, let also scalars c_k, $k \in K$, be given. Then c_λ is defined by $c_\lambda = \prod_{k \in \lambda} c_k$. (In fact, λ might even be infinite, provided the infinite product is convergent.)

Theorem 3.7 *Let K be any index set and let for each $k \in K$ a column vector $v_k \in \mathbb{R}^n$ and a scalar c_k be given such that $\sum_{k\in K} c_k v_k v_k^*$ is entrywise absolutely convergent. Then:*

$$\det\left(\sum_{k\in K} c_k v_k v_k^*\right) = \sum_{\lambda\subset K, |\lambda|=n} c_\lambda S_\lambda^2 .$$

Proof For finite K this is Proposition 1 in [2]. The general case follows by continuity. □

We end this section with a notion that will be needed in Sections 7 and 8. The *multiplicative trace* $\mathrm{mtr}(A)$ of a square matrix $A = (a_{ij})$ is defined as follows:

$$\mathrm{mtr}(A) = \prod_i a_{ii} . \tag{4}$$

If $A \in \mathcal{S}_n$, then $\mathrm{mtr}(A) \geq 0$, and $\mathrm{mtr}(A) = 0$ occurs only if A has a complete row (and corresponding column) of zeros. One also has, for $A \in \mathcal{S}_n^+$ and $\alpha \geq 0$:

$$\mathrm{mtr}(A^{\diamond\alpha}) = (\mathrm{mtr}(A))^\alpha. \tag{5}$$

Lemma 3.8 *Suppose that $A \in \mathcal{S}_n$ has rank 1. Let B be any $n \times n$ matrix. Then*

$$\det(A \diamond B) = \mathrm{mtr}(A) \cdot \det(B) .$$

Proof Write $A = (a_{ij})$ and $B = (b_{ij})$. There exists a vector $a = (a_i)$ such that $aa^* = A$, thus $a_{ij} = a_i a_j$. Then $A \diamond B = (a_i a_j b_{ij})$, and hence, by a standard property of the determinant: $\det(A \diamond B) = \prod_i a_i \prod_j a_j \det(b_{ij}) = \prod_i a_{ii} \cdot \det(B) = \mathrm{mtr}(A)\det(B)$. □

4 Polynomial coefficients

In this section we collect some notation, definitions and formulae that will be needed in the sequel. Let p be a positive integer. For $\mathbf{m} = (m_1, \ldots, m_p) \in \mathbb{N}^p$ we define:

$$|\mathbf{m}| = \sum_{i=1}^p m_i, \qquad \mathbf{m}! = \prod_{i=1}^p m_i! . \tag{6}$$

For $\lambda \subset \mathbb{N}^p$ we define:

$$|\lambda| = \mathrm{Card}\,(\lambda), \qquad ||\lambda|| = \sum_{\mathbf{m}\in\lambda} |\mathbf{m}|. \tag{7}$$

We shall denote the usual binomial coefficients $\binom{m}{k}$ by $(m\,|\,k)$; thus:

$$(m\,|\,k) = \frac{1}{k!}\prod_{i=0}^{k-1}(m-i) \qquad (0 \leq k \leq m) \tag{8}$$

(by convention, an empty product (which occurs when $k = 0$) equals 1). Actually, it will be convenient to use a more symmetric way to denote binomial coefficients, by defining

$$(m\,||\,k) = (m+k\,|\,k) , \tag{9}$$

or, explicitly:

$$(m\,||\,k) = \frac{1}{k!}\prod_{i=1}^{k}(m+i) \qquad (m, k \geq 0) . \tag{10}$$

Note that $(m\,||\,k) = (k\,||\,m)$. The well-known formula $(m+1\,|\,k) = (m\,|\,k) + (m\,|\,k-1)$ is transformed into the equality

$$(m+1\,||\,k) = (m\,||\,k) + (m+1\,||\,k-1) . \tag{11}$$

The following equalities are easily seen to hold for all integers $p \geq 1$ and $k \geq 0$.

$$\operatorname{Card}\{\mathbf{m} \in \mathbb{N}^p \mid |\mathbf{m}| = k\} = (p-1 \,||\, k)\,, \tag{12}$$

$$\operatorname{Card}\{\mathbf{m} \in \mathbb{N}^p \mid |\mathbf{m}| \leq k\} = (p \,||\, k)\,. \tag{13}$$

Let p and n be positive integers. We define:

$$\Lambda(p,n) = \{\lambda \mid \lambda \subset \mathbb{N}^p,\ |\lambda| = n\}\,; \tag{14}$$

$$\Lambda(p,n,l) = \{\lambda \mid \lambda \in \Lambda(p,n),\ ||\lambda|| = l\} \qquad (l \geq 0)\,; \tag{15}$$

$$L(p,n) = \min\{||\lambda|| \mid \lambda \in \Lambda(p,n)\}\,. \tag{16}$$

Example 4.1 For $n = 1$ we have $\Lambda(p,1) = \mathbb{N}^p$, if $\{\mathbf{m}\}$ and $\mathbf{m}$ are identified. Then $\Lambda(p,1,l) = \{\mathbf{m} \in \mathbb{N}^p \mid |\mathbf{m}| = l\}$, in particular $\Lambda(p,1,0) = \{(0,\ldots,0)\}$ and $L(p,1) = 0$. For $2 \leq n \leq p+1$ we have $L(p,n) = n-1$.

For brevity we shall often write L instead of $L(p,n)$. For instance, the set $\Lambda(p,n,L(p,n))$ of all $\lambda \in \Lambda(p,n)$ for which $||\lambda||$ is minimal will be denoted by $\Lambda(p,n,L)$. We clearly have:

$$\Lambda(p,n) = \bigcup_{l=L}^{\infty} \Lambda(p,n,l)\,. \tag{17}$$

For $i,j \in \mathbb{Z}$ the set $\{n \in \mathbb{Z} \mid i \leq n \leq j\}$ will be denoted by $[i,j]$ and will be called a (finite) interval in $\mathbb{Z}$. We consider finite intervals in $\mathbb{N}$ delimited by binomial coefficients $(m \,||\, k)$; they will play an important part later on. Explicitly, we define:

$$D_m(k) = [\,(m \,||\, k), (m \,||\, k+1)] \qquad (m \geq 1,\ k \geq 0)\,. \tag{18}$$

For each fixed $m \geq 1$ the $D_m(k)$ $(k \geq 0)$ are subsequent intervals, two adjacent intervals having one point in common. For instance, for $m = 3$ these intervals are: $[1,4]$, $[4,10]$, $[10,20]$, ... Note that for $m = 0$ one would get $D_0(k) = \{1\}$ for all k.

The number of elements in $D_m(k)$ (denoted as $|D_m(k)|$) is 1 more than the difference between its last and its first element. Applying (11) (with k and m interchanged) we therefore obtain:

$$|D_m(k)| = 1 + (m-1 \,||\, k+1)\,. \tag{19}$$

Lemma 4.2 *Let p and n be positive integers. Let $k \geq 0$ be such that $n \in D_p(k)$. Consider an element $\lambda \in \Lambda(p,n)$. Then $\lambda \in \Lambda(p,n,L)$ if and only if*

$$\{\mathbf{m} \in \mathbb{N}^p \mid |\mathbf{m}| \leq k\} \subset \lambda \subset \{\mathbf{m} \in \mathbb{N}^p \mid |\mathbf{m}| \leq k+1\}\ . \tag{20}$$

Proof Denote the set $\{\mathbf{m} \mid |\mathbf{m}| \leq k\}$ by M_k. From (13), (14) and the definition of k it follows that

$$\operatorname{Card} M_k \leq \operatorname{Card} \lambda \leq \operatorname{Card} M_{k+1}\,. \tag{21}$$

If $\lambda \in \Lambda(p,n)$ and $M_k \not\subset \lambda$, then also $\lambda \not\subset M_k$, by (21). Therefore, one can replace an element of $\lambda \backslash M_k$ by an element of $M_k \backslash \lambda$ and obtain an element λ' with $||\lambda'|| < ||\lambda||$. Thus $||\lambda||$ is not minimal in $\Lambda(p,n)$. A similar reasoning shows that $||\lambda||$ is not minimal in $\Lambda(p,n)$ if $\lambda \not\subset M_{k+1}$. So (20) is a necessary condition for minimality of $||\lambda||$. But it is also sufficient, because for all λ satisfying (20) $||\lambda||$ takes the same value. $\square$

Remark 4.3 In the special case where $n = (p \,||\, k+1)$ for some $k \geq 0$ one has $n \in D_p(k) \cap D_p(k+1)$. In this case, however, $\Lambda(p,n,L)$ contains only one element: $\lambda = M_{k+1}$, so that indeed (20) holds for both k and $k+1$.

Remark 4.4 In general, to obtain a $\lambda \subset \mathbb{N}^p$ with $|\lambda| = n$ and $\|\lambda\|$ minimal, one has to choose $n - (p \,\|\, k)$ elements (with k as in Lemma 4.2) from a set of $(p-1 \,\|\, k+1)$ elements (cf. (12)). Hence the number of elements λ for which $\|\lambda\|$ is minimal is:

$$\mathrm{Card}\,(\Lambda(p,n,L)) = ((p-1 \,\|\, k+1) \,|\, n - (p \,\|\, k)) \,. \tag{22}$$

This agrees with the previous remark for the special values of n considered there.

Now some generalizations of the 'classical' binomial coefficients will be defined; see (6) for the notation. For $\mathbf{m} \in \mathbb{N}^p$ $(p \geq 1)$ we define the *p-nomial coefficients*

$$(|\mathbf{m}| \,|\, \mathbf{m}) = \frac{|\mathbf{m}|!}{\mathbf{m}!} \,. \tag{23}$$

The quantity $(|\mathbf{m}| \,|\, \mathbf{m})$ equals the number of ways in which $|\mathbf{m}|$ objects can be divided into p numbered classes in such a way that the i-th class contains m_i objects $(1 \leq i \leq p)$. For $p = 2$ these are the usual binomial coefficients, with a slightly different notation.

The following formula is well known:

$$(a_1 + \cdots + a_p)^k = \sum_{\mathbf{m} \in \mathbb{N}^p, |\mathbf{m}| = k} (k \,|\, \mathbf{m})\, \mathbf{a}^{\mathbf{m}} \,, \tag{24}$$

where for $\mathbf{a} = (a_1, \ldots, a_p)$, $\mathbf{m} = (m_1, \ldots, m_p)$ we have used the notation:

$$\mathbf{a}^{\mathbf{m}} = \prod_{i=1}^{p} a_i^{m_i} \,. \tag{25}$$

For $\alpha \in \mathbb{R}$ and $m \in \mathbb{N}$ one has the generalized binomial coefficients:

$$(\alpha \,|\, m) = \frac{\alpha(\alpha-1) \cdot \cdots \cdot (\alpha - m + 1)}{m!} \,. \tag{26}$$

They occur in the Taylor formula

$$(1+x)^\alpha = \sum_{m=0}^{\infty} (\alpha \,|\, m)\, x^m \qquad (\alpha \in \mathbb{R},\ -1 < x < 1) \,. \tag{27}$$

We now define for $\mathbf{m} = (m_1, \ldots, m_p) \in \mathbb{N}^p$ $(p \geq 1)$ the *generalized p-nomial coefficients*:

$$(\alpha \,|\, \mathbf{m}) = \frac{\alpha(\alpha-1) \cdot \cdots \cdot (\alpha - |\mathbf{m}| + 1)}{\mathbf{m}!} \,. \tag{28}$$

For $\alpha = |\mathbf{m}|$ this definition coincides with the earlier definition (23). For $p = 1$ it coincides with (26), if $(\alpha \,|\, \{m\})$ and $(\alpha \,|\, m)$ are identified.

Remark 4.5 The apparent discrepancy with our earlier observation that (23) coincides with the usual coefficients, not for $p = 1$ but for $p = 2$, is understood on observing that for integers k and m with $k \geq m$ one has $(k \,|\, \{m\}) = (k \,|\, \{m, k - m\})$.

The following equality follows trivially from (26), (23), and (28):

$$(\alpha \,|\, |\mathbf{m}|)\, (|\mathbf{m}| \,|\, \mathbf{m}) = (\alpha \,|\, \mathbf{m}) \quad (\alpha \in \mathbb{R},\ \mathbf{m} \in \mathbb{N}^p\ (p \geq 1)) \,. \tag{29}$$

We finally define,for $\alpha \in \mathbb{R}$, $p \geq 1$, and $\lambda \subset \mathbb{N}^p$, λ finite:

$$(\alpha \,|\, \lambda) = \prod_{\mathbf{m} \in \lambda} (\alpha \,|\, \mathbf{m}) \,. \tag{30}$$

Lemma 4.6 *Let α be a real number and let p and n be positive integers. Let the integer k be such that $n \in D_p(k)$. If $\lambda \in \Lambda(p,n,L)$, then*

$$(\alpha \,|\, \lambda) = \prod_{j=0}^{k} (\alpha - j)^{n - (j \,\|\, p)} \Big/ \prod_{\mathbf{m} \in \lambda} \mathbf{m}! \,. \tag{31}$$

Proof Consider an element $\mathbf{m} \in \mathbb{N}^p$. The numerator of $(\alpha \,|\, \mathbf{m})$ as given in (28) contains a factor $\alpha - j$ if and only if $|\mathbf{m}| > j$. For $\lambda \in \Lambda(p,n,L)$ there are $|\lambda| - \mathrm{Card}\,\{\mathbf{m} \mid |\mathbf{m}| \le j\} = n - (j \,\|\, p)$ such elements $\mathbf{m}$ in λ (cf. Remark 4.4)). Therefore, for such a λ the numerator of $(\alpha \,|\, \lambda)$ is equal to $\prod_{j=0}^{k} (\alpha - j)^{n - (j \,\|\, p)}$. Finally, the denominator of $(\alpha \,|\, \lambda)$, as given in (31), is obtained by combining (30) and (28). □

Corollary 4.7 *Let α, p and n be as in* Lemma 4.6. *Then for $\lambda \in \Lambda(p,n,L)$ the coefficients $(\alpha \,|\, \lambda)$ all have the same sign.*

Proof The numerator in (31) is the same for all $\lambda \in \Lambda(p,n,L)$. □

5 The Hadamard span

Let E be any set. For $\mathbf{m} = (m_v)_{v \in E} \in \mathbb{N}^E$ we define, analogous to (6):

$$|\mathbf{m}| = \sum_{v \in E} m_v \,. \tag{32}$$

If E is infinite, then $|\mathbf{m}|$ can be $+\infty$. We shall denote by $\mathbb{N}_0^E$ the set of all $\mathbf{m} \in \mathbb{N}^E$ for which $|\mathbf{m}| < \infty$ (i.e. the set of all functions from E to $\mathbb{N}$ with finite support).

Now let E be a subset of $\mathbb{R}^n$. Each $\mathbf{m} \in \mathbb{N}_0^E$ defines a Hadamard product $v_{\mathbf{m}}$ of elements of E, in the following way:

$$v_{\mathbf{m}} = \prod_{v \in E}^{\diamond} v^{\diamond m_v} \qquad (\mathbf{m} = (m_v) \in \mathbb{N}_0^E) \tag{33}$$

(the diamond attached to the product sign indicates that the product is taken in the Hadamard sense). If E has p elements, say $E = \{v_1, \ldots, v_p\}$, then $\mathbb{N}_0^E = \mathbb{N}^E \cong \mathbb{N}^p$ and (33) becomes

$$v_{\mathbf{m}} = \prod_{i=1}^{p}{}^{\diamond}\, v_i^{\diamond m_i} = v_1^{\diamond m_1} \diamond \cdots \diamond v_p^{\diamond m_p} \qquad (\mathbf{m} \in \mathbb{N}^p)\,. \tag{34}$$

The formulae (33) and (34) define mappings $\mathbf{m} \mapsto v_{\mathbf{m}}$ from $\mathbb{N}_0^E$ and $\mathbb{N}^p$, respectively, into $\mathbb{R}^n$.

Definition 5.1 *Let E be a subset of $\mathbb{R}^n$. Then the set*

$$H(E) = \{ v_{\mathbf{m}} \mid \mathbf{m} \in \mathbb{N}_0^E \}$$

(the range of the above mapping) will be called the Hadamard span *of E.*

Note that, conventionally, $H(\emptyset) = \{u\}$.

The following terminology was introduced in [2].

Definition 5.2 *A subset E of a finite-dimensional vector space X is called* quasi linearly independent *if for all linear subspaces Y of X with $Y \neq X$ the set $Y \cap E$ contains at most* $\dim(Y)$ *elements. Otherwise E is called* quasi linearly dependent.

If, in the above definition, $Y = X$ is not excluded, then the definition reduces to ordinary linear (in)dependence. Some further easy observations are collected in the following proposition.

Proposition 5.3 *Let E be a subset of a finite-dimensional vector space X.*

(i) *E is quasi linearly independent if and only if every subset of E with at most* $\dim X$ *elements is linearly independent.*

(ii) *If E has at least* $\dim X$ *elements, then E is quasi linearly independent if and only if every subset of E with* $\dim X$ *elements is a basis for X.*

(iii) *Write $E = \{v_\iota\}_\iota$. If E is quasi linearly independent and $C = \{c_\iota\}_\iota$ is a set of non-zero real numbers, then also $E_C = \{c_\iota v_\iota\}_\iota$ is quasi linearly independent.* □

Example 5.4 If $E \subset V$ is quasi linearly independent and $\dim V > 0$, then $0 \notin E$. In $\mathbb{R}$ the set $\mathbb{R}\backslash\{0\}$ is quasi linearly independent. In $\mathbb{R}^2$ any curve not containing the origin and intersecting each line through the origin at most once, is quasi linearly independent.

The following definition is a combination of the two previous ones; yet another one in this chain of definitions will follow at the end of this section.

Definition 5.5 *A subset E of $\mathbb{R}^n$ will be called* Hadamard quasi linearly independent *if*

(i) *the mapping $\mathbf{m} \mapsto v_{\mathbf{m}}$ from $\mathbb{N}_0^E$ into $\mathbb{R}^n$ is injective;*

(ii) *the Hadamard span $H(E)$ of E is quasi linearly independent.*

If a set E is Hadamard quasi linearly independent, then so are all its subsets (including the empty set). It is therefore of interest to examine the case of a single vector $v \in \mathbb{R}^n$ more closely. When will the singleton $\{v\}$ be Hadamard quasi linearly independent? Condition (i) of Definition 5.5 is satisfied unless all coordinates of v are 0 or 1 or -1, in which case all even powers of v are equal, likewise all odd powers, and even *all* powers if all non-zero coordinates of v have the same sign. Condition (ii) is satisfied if the vectors $u, v, v^{\diamond 2}, v^{\diamond 3} \ldots$ form a quasi linearly independent set. For this latter condition we obtain a characterization via the following straightforward lemma.

Lemma 5.6 *Let $v \in \mathbb{R}^n$ be given. Let r be a positive integer and let $m_1, \ldots, m_r$ be integers satisfying $0 \le m_1 < m_2 < \ldots < m_r$. Then the vectors $v^{\diamond m_1}, \ldots, v^{\diamond m_r}$ are linearly dependent if and only if there is a polynomial P of the form $P(x) = c_1 x^{m_1} + \ldots + c_r x^{m_r}$ with not all c_i equal to 0, such that all n coordinates of v are roots of P.*

Proof This is obvious because $c_1 v^{\diamond m_1} + \ldots + c_r v^{\diamond m_r} = 0$ is valid if and only if $c_1 x_i^{m_1} + \ldots + c_r x_i^{m_r} = 0$ for all coordinates x_i of v. □

Proposition 5.7 *Let $v \in \mathbb{R}^n$ be given. Then the vectors $u, v, v^{\diamond 2}, v^{\diamond 3} \ldots$ are quasi linearly dependent if and only if there exists a non-zero polynomial P with at most n terms such that all coordinates of v are roots of P.*

Proof By (i) of Proposition 5.3 these vectors are quasi linearly dependent if and only if there are integers m_i $(1 \le i \le n)$ with $0 \le m_1 < m_2 < \ldots < m_n$ such that the vectors $v^{\diamond m_1}, \ldots, v^{\diamond m_n}$ are linearly dependent. An application of Lemma 5.6 now gives the result. □

For the next proposition we need some notation. Let $x_1, \ldots, x_n$ be the coordinates of some $v \in \mathbb{R}^n$. Consider the polynomial Q_v be defined by $Q_v(x) = \prod_{i=1}^{n}(x - x_i)$. Set

$$e_i = \sum_{\lambda \subset \{1,\ldots,n\},\, |\lambda| = i} \; \prod_{j \in \lambda} x_j \qquad (1 \le i \le n)\,,$$

the elementary symmetric sums of the x_i. Then

$$Q_v(x) = x^n + \sum_{i=1}^{n} (-1)^i \, e_i \, x^{n-i}\,.$$

Proposition 5.8 *Let $v \in \mathbb{R}^n$ be given, with coordinates $x_1, \ldots, x_n$ and elementary symmetric sums $e_1, \ldots, e_n$. Then the following properties hold.*

(i) *The vectors $u, v, v^{\diamond 2}, \ldots, v^{\diamond (n-1)}$ are linearly dependent if and only if not all coordinates of v are distinct.*

(ii) *Suppose that all coordinates of* v *are distinct. Then the vectors* $u, v, v^{\diamond 2}, \ldots, v^{\diamond n}$ *are quasi linearly dependent if and only if* $e_i = 0$ *for some* i $(1 \leq i \leq n)$. *More specifically: if* $e_i = 0$, *then the* $v^{\diamond m}$ *with* $0 \leq m \leq n$ *and* $m \neq n - i$ *are linearly dependent.*

(iii) *If* v *has at least one coordinate equal to* 0, *then the vectors* $v, v^{\diamond 2}, \ldots, v^{\diamond n}$ *are linearly dependent.*

Proof The determinant of the matrix with columns $u, v, \ldots, v^{\diamond(n-1)}$ is a Vandermonde determinant; its value is $\prod_{1 \leq i < j \leq n}(x_i - x_j)$. This proves (i).

Concerning (ii), it follows from Lemma 5.6 that the said vectors are quasi linearly dependent if and only if there is a polynomial P of degree at most n and with at most n terms such that all coordinates of v are roots of P. On the other hand, such a polynomial must be divisible by Q_v (notation as before; here we use that all coordinates of v are distinct). Because Q_v has degree n, the only possibility is that P is a (non-zero) multiple of Q_v. By Proposition 5.7 quasi linear dependence occurs if and only if Q_v has at most n terms, thus if and only if at least one of the e_i equals 0. This proves the first part of (ii). The specific case follows from Lemma 5.6.

Finally, if v has at least one coordinate equal to 0, then $e_n = 0$, thus (iii) follows from (ii). Of course, (iii) is also evident without (ii): even *all* $v^{\diamond m}$ $(m \geq 1)$ belong to the $(n-1)$-dimensional subspace of all vectors with that specific coordinate equal to 0. □

It is instructive to verify the correctness of the above proposition, and its proof, when $n = 1$.

Theorem 5.9 *Let a vector* $v \in \mathbb{R}^n$ *be given.*

(i) *If* v *has a coordinate* 0, *or if two of its coordinates are equal, then the singleton* $\{v\}$ *is not Hadamard quasi linearly independent.*

(ii) *Suppose that the coordinates of* v *are non-zero, pairwise different, and of the same sign. Then* $\{v\}$ *is Hadamard quasi linearly independent.*

Proof In case (i) it follows from (iii) and (i) of Proposition 5.8, respectively, that the vectors $u, v, v^{\diamond 2}, \ldots, v^{\diamond n}$ are quasi linearly dependent, and hence that $\{v\}$ is not Hadamard quasi linearly independent.

Now suppose that v satisfies the conditions given in (ii). Then either all its coordinates are strictly positive, or they are strictly negative. Assume the former. Consider any polynomial P with at most n terms. By Descartes' rule of signs – the number of positive roots of a real polynomial is at most equal to the number of changes of sign in its sequence of non-zero coefficients; if it is less, then by an even number (cf. [8, Part V, Chapter 1]) – such a polynomial can have at most $n - 1$ positive roots. The desired result now follows from Proposition 5.7. If all coordinates of v are negative, the result follows by considering $-v$ (cf. (iii) of Proposition 5.3) or by applying Descartes' rule to $P(-x)$. □

Remark 5.10 In [2] we prove the stronger result that for a vector v as in (ii) above even the set $\{v^{\diamond \alpha} \mid \alpha \in \mathbb{R}\}$ is quasi linearly independent; see Lemma 1 and the remark which follows it in [2]. The above proof of the more restricted result is simpler and more direct.

Remark 5.11 For a singleton $\{v\}$ to be Hadamard quasi linearly independent it is not enough to require only that its entries are non-zero and mutually distinct *in absolute value*. For instance, if $v^* = (1, 3, -4)$, then, by (ii) of Proposition 5.8 (or by direct verification), the vectors $u, v, v^{\diamond 3}$ are linearly dependent (because $e_1 = 0$); and if $v^* = (1, 3, -3/4)$, then $u, v^{\diamond 2}, v^{\diamond 3}$ are linearly dependent (because $e_2 = 0$). As another example, consider the polynomial $x^5 - 2x + 1$. It has three real roots, say x_1, x_2, x_3, with $x_1 < -1$ and $1/2 < x_2 < x_3 = 1$. Let v be the vector with coordinates x_1, x_2, x_3. One verifies that the corresponding numbers e_1, e_2, e_3 are non-zero: e_1 is positive, while e_2 and e_3 are negative. Thus u, v, $v^{\diamond 2}$, $v^{\diamond 3}$ are quasi linearly independent, by (ii) of Proposition 5.8. But the vectors u, v, $v^{\diamond 5}$ are linearly dependent (explicitly: $v^{\diamond 5} - 2v + u = 0$), thus $\{v\}$ is not Hadamard quasi linearly independent. This example shows that the last condition in (ii) of Theorem 5.9 cannot be dropped.

Now consider the case of an arbitrary finite subset E of $\mathbb{R}^n$ with p elements, say $E = \{v_1, \ldots, v_p\}$. When will E be Hadamard quasi linearly independent? As we have seen, a necessary condition is that each singleton $\{v_k\}$ $(1 \le k \le p)$ is Hadamard quasi linearly independent. The elements of $H(E)$ are the vectors $v_{\mathbf{m}}$ with $\mathbf{m} \in \mathbb{N}^p$. Let us write $v_k^* = (x_{1k}, \ldots, x_{nk})$ $(1 \le k \le p)$. If $\mathbf{m} = (m_1, \ldots, m_p)$, then the coordinates of $v_{\mathbf{m}}$ are $\prod_{k=1}^{p} x_{ik}^{m_k}$ $(1 \le i \le n)$. Now $H(E)$ is Hadamard quasi linearly independent unless there are n different elements $\mathbf{m}_j \in \mathbb{N}^p$ $(1 \le j \le n)$ such that the vectors $v_{\mathbf{m}_j}$ are linearly dependent. Writing $\mathbf{m}_j = (m_{j1}, \ldots, m_{jp})$ $(1 \le j \le n)$, this is the case if and only if

$$\det \left(\prod_{k=1}^{p} x_{ik}^{m_{jk}} \right)_{ij} = 0\,. \tag{35}$$

The left hand side of (35) is a polynomial in the np variables x_{ik} $(1 \le i \le n,\ 1 \le k \le p)$, Moreover, this polynomial is homogeneous of degree $\sum_{k=1}^{p}\sum_{j=1}^{n} m_{jk} = \sum_{j=1}^{n} |\mathbf{m}_j|$. Equation (35) determines a conic manifold in $\mathbb{R}^{np}$, call it W_μ, where $\mu = (\mathbf{m}_1, \ldots, \mathbf{m}_n) = (m_{jk}) \in (\mathbb{N}^p)^n \cong \mathbb{N}^{np}$. We observe that W_μ has Lebesgue measure zero, and that hence so has the set $W = \bigcup_{\mu \in \mathbb{N}^{np}} W_\mu$. Representing, in a natural way, the set E as an element of $\mathbb{R}^{np}$, we conclude that E is Hadamard quasi linearly independent unless $E \in W$, a negligible subset of $\mathbb{R}^{np}$.

Any matrix $A \in \mathcal{S}_n$ of rank p corresponds to a unique cloud of size n in $\mathbb{R}^p$ (cf. Definition 3.2 and Theorem 3.4). As described after Definition 3.2, each positioning $(\widetilde{v}_1, \ldots, \widetilde{v}_n)$ of this cloud in $\mathbb{R}^p$ determines a p-tuple $\{v_1, \ldots, v_p\}$ of vectors in $\mathbb{R}^n$ such that $A = v_1 v_1^* + \cdots + v_p v_p^*$. It may well happen that some of these p-tuples are Hadamard quasi linearly dependent, whereas most others are not. One can try to position the cloud of A in such a way that the corresponding p-tuple is Hadamard quasi linearly independent.

These considerations motivate the following definition.

Definition 5.12 *A matrix $A \in \mathcal{S}_n$ is called* Hadamard independent *if it has a representation of the form $A = v_1 v_1^* + \cdots + v_p v_p^*$ with a set $\{v_1, \ldots, v_p\}$ (with p elements) which is Hadamard quasi linearly independent.*

For a matrix $A \in \mathcal{S}_n$ to be Hadamard independent it is certainly necessary that the n points in its cloud are all distinct and all different from zero. Indeed, if $\widetilde{v}_i = \widetilde{v}_j$ (notation as in Definitions 3.2 and 3.3), then the i-th and j-th row of A will be equal (the columns as well), and the equality of the two vectors persists after an orthogonal transformation; likewise if one of the vectors in the cloud is the zero vector. Now consider a matrix $A \in \mathcal{S}_n$ of rank p whose cloud consists of n non-zero and distinct points. If A has rank 1, it can still happen that A is not Hadamard independent, due to the fact that in $\mathbb{R}$ its cloud has only two positionings. For instance, if $A = vv^*$ with $v^* = (1, 3, -4)$, as in Remark 5.11, then the only positionings of its cloud $[1, 3, -4]$ in $\mathbb{R}$ are $\pm(1, 3, -4)$, and in both cases we get Hadamard dependency.

However, if A (as above) has rank at least 2, it seems plausible that a positioning of its cloud can be found for which the corresponding p-tuple of vectors is Hadamard quasi linearly independent. The following example will illustrate this.

Example 5.13 We take $n = 2$. Consider the cloud (of size 2) $C = [(1, 1), (a, b)]$ with a and b positive and different from 1 (as explained after Definition 3.2, it makes sense to consider the vectors in a cloud as *row* vectors). The matrix A_C determined by C is the matrix $A = v_1 v_1^* + v_2 v_2^*$ with $v_1^* = (1, a)$ and $v_2^* = (1, b)$; thus $A = \begin{pmatrix} 2 & a+b \\ a+b & a^2+b^2 \end{pmatrix}$.

Set $E = \{v_1, v_2\} \subset \mathbb{R}^2$. The Hadamard span $H(E)$ consists of all vectors $v_1^{\diamond m} v_2^{\diamond k} = (1, a^m b^k)^*$ with $m, k \in \mathbb{N}$. In this particular case condition (ii) in Definition 5.5 is trivially satisfied because $H(E)$ is contained in the line $\{(1, y) \mid y \in \mathbb{R}\}$ (cf. the last part of Example 5.4). One finds that condition (i) is satisfied unless b is a rational power of a. More generally, if v_1 and v_2 each have positive and different coordinates, say $v_1 = (x_1, x_2)^*$, $v_2 = (y_1, y_2)^*$, then E is Hadamard quasi linearly independent unless x_1/x_2 and y_1/y_2 are rational powers of one another.

As a specific example, let us take $a = 2$ and $b = 4$. Then $A = \begin{pmatrix} 2 & 6 \\ 6 & 20 \end{pmatrix}$, its rank is 2, but its generating set $E = \{v_1, v_2\}$ is not Hadamard quasi linearly independent, by the criterion above (we have $b = a^2$). Now apply an orthogonal transformation, say a rotation over θ. Then a generating set $F = \{w_1, w_2\}$ is obtained with $w_1 = \cos\theta\, v_1 - \sin\theta\, v_2$, $w_2 = \sin\theta\, v_1 + \cos\theta\, v_2$, thus $w_1^* = (\cos\theta - \sin\theta\,,\, 2\cos\theta - 4\sin\theta)$, $w_2^* = (\cos\theta + \sin\theta\,,\, 4\cos\theta + 2\sin\theta)$ (note that indeed $w_1 w_1^* + w_2 w_2^* = A$). From the criterion above it follows that F is Hadamard quasi linearly independent, for instance, for values of θ close to 0 for which the quotient

$$\frac{\log(\cos\theta - \sin\theta) - \log(2\cos\theta - 4\sin\theta)}{\log(\cos\theta + \sin\theta) - \log(4\cos\theta + 2\sin\theta)}$$

is irrational.

For larger values of n and p the computations become rather complicated. One can for instance embed the orthogonal group as a compact manifold O_p of dimension $p(p+1)/2$ in the p^2-dimensional vector space of all $p \times p$ matrices. Each choice $M = \{\mathbf{m}_1, \ldots, \mathbf{m}_n\}$ of n elements in $\mathbb{N}^p$ determines a polynomial, say P_M, in the p^2 variables of the vector space. It remains then to show that the manifold O_p is not contained in the union of the countably many manifolds $P_M = 0$. However, our computations are not conclusive so far.

6 A Taylor expansion

In this section we obtain an expansion of Taylor type for the determinant of fractional powers of matrices near the unit matrix U.

Lemma 6.1 *Let V be an $n \times n$ matrix all of whose entries are less than 1 in absolute value. Let α be a real number. Then the α-th fractional Hadamard power of $U + V$ is given by:*

$$(U+V)^{\diamond\alpha} = \sum_{k=0}^{\infty} (\alpha \,|\, k) V^{\diamond k}\,.$$

Proof Apply (27) entrywise. □

Lemma 6.2 *Let $v_1, \ldots, v_p$ be vectors in $\mathbb{R}^n$ and let the matrix V be given by*

$$V = v_1 v_1^* + \cdots + v_p v_p^*\,.$$

Then the integer Hadamard powers $V^{\diamond k}$ $(k \geq 0)$ are given by the formula (cf. (34) for the notation):

$$V^{\diamond k} = \sum_{\mathbf{m}\in\mathbb{N}^p,\, |\mathbf{m}|=k} (k \,|\, \mathbf{m})\, v_{\mathbf{m}} v_{\mathbf{m}}^*\,.$$

Proof Set $v_l = (x_{il})_i$ $(1 \leq l \leq p)$; then $V = \left(\sum_{l=1}^{p} x_{il} x_{jl}\right)_{i,j}$. Now apply (24) and (25) first, and then (34). □

Lemma 6.3 *Let $v_1, \ldots, v_p$ and V be as in* Lemma 6.2 *above. Let α be an arbitrary real number. Then for ε sufficiently close to zero the α-th fractional Hadamard power of the matrix $U + \varepsilon V$ is given by*

$$(U + \varepsilon V)^{\diamond\alpha} = \sum_{\mathbf{m}\in\mathbb{N}^p} (\alpha \,|\, \mathbf{m})\, \varepsilon^{|\mathbf{m}|} v_{\mathbf{m}} v_{\mathbf{m}}^*\,. \tag{36}$$

Proof Take ε so close to zero that all entries of εV are less than 1 in absolute value. Now combine the Lemmas 6.1 and 6.2 and apply (29). $\square$

The following key result can now be proved (cf. (34), (30), (15), (16) and Definition 3.6 for the notation).

Theorem 6.4 *Let $v_1, \ldots, v_p$ be vectors in $\mathbb{R}^n$ and let the matrix V be given by*

$$V = v_1 v_1^* + \cdots + v_p v_p^* .$$

Then for ε sufficiently close to zero and any $\alpha \in \mathbb{R}$ the determinant of the matrix $(U + \varepsilon V)^{\diamond\alpha}$ is given by the formula

$$\det((U + \varepsilon V)^{\diamond\alpha}) = \sum_{l=L}^{\infty} C_l \, \varepsilon^l , \tag{37}$$

where $L = L(p, n)$ and the C_l $(l \geq L)$ are given by

$$C_l = \sum_{\lambda \in \Lambda(p,n,l)} (\alpha \,|\, \lambda) S_\lambda^2 . \tag{38}$$

Proof Applying Theorem 3.7 to (36) in Lemma 6.3 we obtain:

$$\det((U + \varepsilon V)^{\diamond\alpha}) = \sum_{\lambda \subset \mathbb{N}^p,\, |\lambda| = n} \left(\prod_{\mathbf{m} \in \lambda} (\alpha \,|\, \mathbf{m}) \varepsilon^{|\mathbf{m}|} \right) S_\lambda^2 .$$

Substitution of (14), (30), and (7) into this equality entails

$$\det((U + \varepsilon V)^{\diamond\alpha}) = \sum_{\lambda \in \Lambda(p,n)} (\alpha \,|\, \lambda) \, \varepsilon^{\|\lambda\|} S_\lambda^2 . \tag{39}$$

Finally, rearranging this sum according to the partition given in (17), formula (37) follows. $\square$

Remark 6.5 From Theorem 6.4 it is clear why integer values for the exponent α are exceptional in the sense that integer powers of symmetric positive-definite matrices are again positive-definite (Schur's theorem (Proposition 2.1)). Indeed, in this case the factors $(\alpha \,|\, \mathbf{m})$ (cf. (28)) don't take negative values: they are positive for $|\mathbf{m}| \leq \alpha$ and zero for $|\mathbf{m}| \geq \alpha + 1$; consequently, the coefficients $(\alpha \,|\, \lambda)$ (cf. (30)) are 0 for λ sufficiently large, the series (37) is finite, and the sum is positive.

Example 6.6 Consider the case $n = 2$. First suppose that $p = 1$. Take $v = \begin{pmatrix} x \\ y \end{pmatrix}$; then $V = vv^* = \begin{pmatrix} x^2 & xy \\ xy & y^2 \end{pmatrix}$. We want to compute

$$\det(U + \varepsilon V)^\alpha = \begin{vmatrix} (1 + \varepsilon x^2)^\alpha & (1 + \varepsilon xy)^\alpha \\ (1 + \varepsilon xy)^\alpha & (1 + \varepsilon y^2)^\alpha \end{vmatrix} .$$

For $|\varepsilon|$ small enough we can write $(1 + \varepsilon x^2)^\alpha = \sum_{r=0}^{\infty} (\alpha \,|\, r) \varepsilon^r x^{2r}$, and similarly for the other terms. Substituting this into the determinant we obtain

$$\det(U + \varepsilon V)^\alpha = \sum_{r_1=0}^{\infty} \sum_{r_2=0}^{\infty} (\alpha \,|\, r_1)(\alpha \,|\, r_2) \, \varepsilon^{r_1 + r_2} \left(x^{2r_1} y^{2r_2} - x^{r_1+r_2} y^{r_1+r_2} \right) .$$

The terms with $r_1 = r_2$ vanish, and the remaining terms can be combined pairwise. This gives

$$\begin{aligned}\det(U+\varepsilon V)^\alpha &= \sum_{0\le r_1<r_2} (\alpha\,|\,r_1)(\alpha\,|\,r_2)\,\varepsilon^{r_1+r_2}\left(x^{2r_1}y^{2r_2} - 2x^{r_1+r_2}y^{r_1+r_2} + x^{2r_2}y^{2r_1}\right)\\ &= \sum_{\lambda\in\Lambda(1,2)} (\alpha\,|\,\lambda)\,\varepsilon^{\|\lambda\|}\left|\begin{array}{cc} x^{r_1} & x^{r_2}\\ y^{r_1} & y^{r_2}\end{array}\right|^2,\end{aligned}$$

which is equality (39).

Next take $p = 2$, the general case for a 2×2 matrix. We have then

$$V = v_1v_1^* + v_2v_2^* = \left(\begin{array}{cc} x_1^2+x_2^2 & x_1y_1+x_2y_2\\ x_1y_1+x_2y_2 & y_1^2+y_2^2\end{array}\right),$$

and we find, as before, that $\det(U+\varepsilon V)^\alpha$ is equal to

$$\sum_{r_1=0}^{\infty}\sum_{r_2=0}^{\infty}(\alpha\,|\,r_1)(\alpha\,|\,r_2)\,\varepsilon^{r_1+r_2}\left((x_1^2+x_2^2)^{r_1}(y_1^2+y_2^2)^{r_2} - (x_1y_1+x_2y_2)^{r_1+r_2}\right).$$

For a fixed pair (r_1, r_2) the long expression in parentheses is equal to

$$\sum_{s_1=0}^{r_1}\sum_{s_2=0}^{r_2}(r_1\,|\,s_1)(r_2\,|\,s_2)\left(x_1^{2s_1}x_2^{2t_1}y_1^{2s_2}y_2^{2t_2} - x_1^{s_1+s_2}y_1^{s_1+s_2}x_2^{t_1+t_2}y_2^{t_1+t_2}\right),$$

where we have written $t_i = r_i - s_i$ $(i = 1, 2)$. Combining (r_1, r_2) and (r_2, r_1) we find that $\det(U+\varepsilon V)^\alpha$ is equal to

$$\sum_{0\le r_1\le r_2}(\alpha\,|\,r_1)(\alpha\,|\,r_2)\,\varepsilon^{r_1+r_2}\sum_{s_1=0}^{r_1}\sum_{s_2=0}^{r_2}(r_1\,|\,s_1)(r_2\,|\,s_2)\left(x_1^{s_1}x_2^{t_1}y_1^{s_2}y_2^{t_2} - x_1^{s_2}x_2^{t_2}y_1^{s_1}y_2^{t_1}\right)^2.$$

The terms with $r_1 = r_2$ and $s_1 = s_2$ vanish. Now write $\mathbb{N}^2 = \{(s,t)\mid s\ge 0,\, t\ge 0\}$ and

$$\Lambda(2,2) = \{\lambda = \{(s_1,t_1),(s_2,t_2)\}\mid s_1,s_2,t_1,t_2\ge 0,\ s_1+t_1\le s_2+t_2,\ s_1=s_2\Rightarrow t_1<t_2\}.$$

Then $(\alpha\,|\,r_1)(r_1\,|\,s_1) = (\alpha\,|\,(s_1,t_1))$ (cf. (29) and Remark 4.5). Likewise for the index 2, and hence (cf. (30)) $(\alpha\,|\,r_1)(\alpha\,|\,r_2)(r_1\,|\,s_1)(r_1\,|\,s_2) = (\alpha\,|\,(s_1,t_1))(\alpha\,|\,(s_2,t_2)) = (\alpha\,|\,\lambda)$.
Further, $r_1 + r_2 = s_1 + t_1 + s_2 + t_2 = \|\lambda\|$. Finally, we have

$$S_\lambda^2 = \left|\begin{array}{cc} x_1^{s_1}x_2^{t_1} & x_1^{s_2}x_2^{t_2}\\ y_1^{s_1}y_2^{t_1} & y_1^{s_2}y_2^{t_2}\end{array}\right|^2 = \left(x_1^{s_1}x_2^{t_1}y_1^{s_2}y_2^{t_2} - x_1^{s_2}x_2^{t_2}y_1^{s_1}y_2^{t_1}\right)^2.$$

Making all these substitutions we again obtain (39).

7 The main theorem

To state our main theorem, we have to define a certain pattern of plus and minus signs first; cf.(18) for the notation.

Definition 7.1 *For integers $p\ge 1$ and $a\ge 0$ the function $T_{p,a} : [p,+\infty)\to\{1,-1\}$ is defined according to the following rules:*

(i) *if $p\le n\le (p\,\|\,a+1)$ (i.e. if $n\in\{p\}\cup D_p(1)\cup\ldots\cup D_p(a)$), then $T_{p,a}(n) = 1$;*

(ii) *for $t\ge 0$ the function $T_{p,a}$ is* constant *on $D_p(a+t)$ when t is even, and* alternating *when t is odd.*

The functions $T_{p,a}$ are well-defined because for each $t\ge 0$ the subsequent sets $D_p(a+t)$ and $D_p(a+t+1)$ have precisely one element in common, so that $n\mapsto T_{p,a}(n)$ $(n\ge p)$ is defined successively on the sets $D_p(a+t)$ $(t\ge 0)$.

Example 7.2 The smallest n for which $T_{p,a}(n)$ is negative is the second element of $D_p(a+1)$. Let us denote this element by $N_{p,a}$. We have then:

$$N_{p,a} = 1 + (a+1 \,\|\, p) = 1 + \frac{(a+2)(a+3)\dots(a+p+1)}{p!}\,. \tag{40}$$

For instance, $N_{1,a} = a+3$, and $N_{2,a} = (a^2+5a+8)/2$. Similarly, $N_{p,0} = p+2$ and $N_{p,1} = (p^2+3p+4)/2$. We also note that $N_{p,a} = N_{a+1,p-1}$.

The following theorem is the main result of the paper.

Theorem 7.3 *Let p and n be integers with $1 \le p \le n$. Let V be a Hadamard independent symmetric positive-definite $n \times n$ matrix of rank p. Let α be a positive non-integer real number. Then*

$$\lim_{\varepsilon \downarrow 0}(\mathrm{sgn}(\det((U+\varepsilon V)^{\circ\alpha}))) = T_{p,[\alpha]}(n)\,. \tag{41}$$

Proof By Definition 5.12, there is a Hadamard quasi linearly independent set of vectors $v_1,\dots,v_p$ in $\mathbb{R}^n$ such that $V = v_1v_1^* + \dots + v_pv_p^*$. Write $\det((U+\varepsilon V)^{\circ\alpha})$ as in (37) and (38) in Theorem 6.4. The Hadamard span $\{v_{\mathbf{m}} \mid \mathbf{m} \in \mathbb{N}^p\}$ is quasi linearly independent, in view of Definition 5.5. Hence it follows from (ii) of Proposition 5.3 that $S_\lambda \neq 0$ for all $\lambda \in \Lambda(p,n)$. Moreover, for $\lambda \in \Lambda(p,n,L)$ all coefficients $(\alpha \,|\, \lambda)$ have the same sign (Corollary 4.7). This implies that $C_L \neq 0$.

When $\varepsilon \downarrow 0$, the term $C_L\varepsilon^L$ in the series (37) becomes dominant. In other words, the limit in formula (41) will be the sign of C_L. By (38) this is equal to the sign of the $(\alpha \,|\, \lambda)$ for $\lambda \in \Lambda(p,n,L)$. It remains to show that this sign is equal to $T_{p,[\alpha]}(n)$ for all $n \ge p$. This will be done by induction.

Let us denote the smallest elements of $\mathbb{N}^p$ as follows: $\mathbf{m}_0 = (0,\dots,0)$, $\mathbf{m}_j = (\delta_{1j},\dots,\delta_{pj})$ $(1 \le j \le p)$, where δ_{ij} is the Kronecker symbol (δ_{ij} is 1 if $i=j$, and 0 if $i \neq j$). Then $|\mathbf{m}_0| = 0$ and $|\mathbf{m}_j| = 1$ $(1 \le j \le p)$. Set $\lambda_0 = \{\mathbf{m}_0, \mathbf{m}_1, \dots, \mathbf{m}_p\}$ and $\lambda_j = \lambda_0 \backslash \{\mathbf{m}_j\}$ $(1 \le j \le p)$. For instance, if $p=1$ then $\lambda_0 = \{(0),(1)\}$ and $\lambda_1 = \{(0)\}$, and if $p=2$, then $\lambda_0 = \{(0,0),(1,0),(0,1)\}$, $\lambda_1 = \{(0,0),(0,1)\}$ and $\lambda_2 = \{(0,0),(1,0)\}$. For $1 \le j \le p$ we have $|\lambda_j| = p$, thus $\lambda_j \in \Lambda(p,p)$, and $\|\lambda_j\| = p-1$. It is also clear (cf. Remark 4.3) that these λ_j are the only $\lambda \in \Lambda(p,p)$ for which $\|\lambda\|$ is minimal. Thus $L(p,p) = p-1$ and $\Lambda(p,p,p-1) = \{\lambda_1,\dots,\lambda_p\}$. Similarly, $|\lambda_0| = p+1$ and $\|\lambda_0\| = p$, and λ_0 is the only $\lambda \in \Lambda(p,p+1)$ for which $\|\lambda\|$ is minimal. Thus $L(p,p{+}1) = p$ and $\Lambda(p,p{+}1,p) = \{\lambda_0\}$. In the terminology of the proof of Lemma 4.2: $M_0 = \{\mathbf{m}_0\}$, $M_1 = \{\mathbf{m}_0,\mathbf{m}_1,\dots,\mathbf{m}_p\}$, and $M_0 \subset \lambda_j \subset \lambda_0 = M_1$ $(1 \le j \le p)$.

For any $\alpha \in \mathbb{R}$, we have (cf. (30) and (28)):

$$(\alpha \,|\, \lambda_0) = \prod_{k=0}^{p}(\alpha \,|\, \mathbf{m}_k) = \alpha^p, \quad (\alpha \,|\, \lambda_j) = (\alpha \,|\, \lambda_0)/(\alpha \,|\, \mathbf{m}_j) = \alpha^{p-1}\ (1 \le j \le p)\,.$$

For $\alpha > 0$ this implies that $C_L (= C_{p-1}) > 0$ when $n = p$ (and $C_L = C_p > 0$ when $n = p+1$). Thus (41) is proved for $n = p$ (and for $n = p+1$).

Now suppose that $n > p$ (or, if one prefers, $n > p+1$), and that (41) has been proved already for $n-1$ instead of n. Let $t \ge 0$ be the unique integer such that

$$([\alpha]+t \,\|\, p) \le n-1 < n \le ([\alpha]+t+1 \,\|\, p)\,.$$

Then both $n-1$ and n belong to $D_p([\alpha]+t)$. Now apply Lemma 4.6 with $k = [\alpha]+t$ (cf. (16)). The factor $\alpha - j$ is positive if and only if $j \le [\alpha]$. Therefore the number of negative factors in the product in the numerator of (31) is equal to

$$\sum_{j=[\alpha]+1}^{k} (n - (j \,\|\, p)) = (k-[\alpha])n - \sum_{j=[\alpha]+1}^{k} (j \,\|\, p)\,,$$

(recall that α is not an integer). This number is $k-[\alpha]$ more than the corresponding number for $n-1$. This shows that the limiting sign in (37) obeys rule (ii) in Definition 7.1. This completes the proof. $\square$

It is worth observing that, for a given size n of the matrices and a fixed value of α, the limiting sign of $\det((U+\varepsilon V)^{\circ\alpha})$ when $\varepsilon \downarrow 0$ (i.e. the left-hand side in (41)) depends only on the *rank* of V, not on its particular shape otherwise.

Example 7.4 The equality $N_{p,a} = N_{a+1,p-1}$ (cf. Example 7.2) implies the following fact. The smallest matrix size for which the limiting sign of $\det((U+\varepsilon V)^{\circ\alpha})$ is negative in the case where $\mathrm{rank}(V) = p$ and $a < \alpha < a+1$ is equal to the smallest size for which the limiting sign is negative in the case where $\mathrm{rank}(V) = a+1$ and $p-1 < \alpha < p$. To give an example, that smallest size when V has rank 10 and $5 < \alpha < 6$ is the same as when V has rank 6 and $9 < \alpha < 10$. This size is $1 + (10 \,\|\, 6) = 8009$ (the second element of $D_{10}(6)$, which is equal to the second element of $D_6(10)$). However, the final elements of $D_{10}(6)$ and $D_6(10)$ are not the same: for the former it is $(10 \,\|\, 7) = 19448$, whereas for the latter it is $(6 \,\|\, 11) = 12376$. This implies (cf. (ii) of Definition 7.1) that in the former case the limiting sign is negative for all odd sizes from 8009 up to and including 19447, whereas in the latter case the limiting sign is negative for all odd sizes from 8009 up to and including 12375 only. After these final values there is again a long stretch of size values n for which the limiting sign is positive: in the former case from 19448 up to $(10 \,\|\, 8) = \frac{18}{8} \cdot (10 \,\|\, 7) = 43758$, in the latter case from 12376 up to $(6 \,\|\, 12) = \frac{18}{12} \cdot (6 \,\|\, 11) = 18564$.

We shall say a little more about such 'sign patterns' in Section 9.

Remark 7.5 In Theorem 7.3 the requirement that V be Hadamard independent was made to guarantee that for all $\lambda \subset \mathbb{N}^p$ with $|\lambda| = n$ the vectors $v_{\mathbf{m}}$ $(\mathbf{m} \in \lambda)$ would be linearly independent, so that all S_λ would be non-zero and $C_L \neq 0$ would hold. But in each specific case a much weaker condition will already suffice. In fact, formula (41) holds as soon as $S_\lambda \neq 0$ holds for at least one $\lambda \in \Lambda(p,n,L)$. In particular, it does not matter if Hadamard dependency occurs for products of higher powers of the generating vectors v_j. To illustrate this, let us examine the cases $n = p$ and $n = p+1$, considered in the proof of Theorem 7.3. We saw there that $L(p,p) = p-1$ and $L(p,p+1) = p$. If $n = p$, then

$$C_L = \sum_{\lambda \in \Lambda(p,p,p-1)} (\alpha \,|\, \lambda)\, S_\lambda^2 = \alpha^{p-1} \sum_{j=1}^{p} S_{\lambda_j}^2 \,,$$

where S_{λ_j} is the volume of the parallelepiped spanned by the vectors $u, v_1, \ldots, v_p$ except v_j. If $n = p+1$, then

$$C_L = \sum_{\lambda \in \Lambda(p,p+1,p)} (\alpha \,|\, \lambda)\, S_\lambda^2 = \alpha^p S_{\lambda_0}^2 \,,$$

where S_{λ_0} is the volume of the parallelepiped spanned by the vectors $u, v_1, \ldots, v_p$. To conclude that $C_L \neq 0$ we don't need Hadamard quasi linear independence of the vectors $v_1 \ldots, v_p$. Indeed, from the way they are chosen it follows that these vectors are linearly independent. If $p = n$, then they form a basis of $\mathbb{R}^n$, and hence there is at least one index j, $1 \leq j \leq p$, such that when v_j is replaced by u we have again a basis, and hence $S_{\lambda_j} \neq 0$. If $p = n-1$, then the only requirement is that u does not belong to the space spanned by the vectors $v_1, \ldots, v_p$. Observe that this is implied by Hadamard quasi linear independence. Loosely speaking, the larger $n - p$ is, the more the Hadamard quasi linear independence of the v_j is needed.

Using Lemma 3.8 we can obtain a more general result, where U is replaced by an arbitrary matrix $T \in \mathcal{S}_n^+$ of rank 1 and with strictly positive entries. Let T be such a matrix, and let $V \in \mathcal{S}_n$ be a matrix of rank p for some $p \leq n$. Consider the matrix $A = T + V$ (an element of the positive-definite cone at T, cf. Section 2). Let us say that A is *sufficiently close* to T with respect to a real number α, if $\det(T+\varepsilon V)^{\circ\alpha}$ is of constant sign for $0 < \varepsilon \leq 1$.

Theorem 7.6 *Suppose that $T \in \mathcal{S}_n^+$ is a matrix of rank 1 with strictly positive entries. Let $V \in \mathcal{S}_n$ be a matrix of rank p for some $p \leq n$. Let α be a non-integer positive real number. If the*

Hadamard quotient $V \diamond/ T$ is Hadamard independent, and A is sufficiently close to T with respect to α, then the sign of $\det A^{\diamond\alpha}$ *is equal to $T_{p,[\alpha]}(n)$.*

Proof Take vectors $v_0, v_1, \ldots, v_p$ such that $T = v_0 v_0^*$ and $V = v_1 v_1^* + \cdots + v_p v_p^*$. All coordinates of v_0 are different from 0, therefore we can consider the vectors $w_j = v_j \diamond/ v_0$ $(1 \leq j \leq p)$. Set $W = w_1 w_1^* + \cdots + w_p w_p^*$. Then $V = T \diamond W$, thus $A = T \diamond (U + W)$. We have $A^{\diamond\alpha} = T^{\diamond\alpha} \diamond (U+W)^{\diamond\alpha}$, and hence by Lemma 3.8: $\det A^{\diamond\alpha} = (\mathrm{mtr}(T))^{\diamond\alpha} \cdot \det(U+W)^{\diamond\alpha}$. Now $(\mathrm{mtr}(T))^{\diamond\alpha} > 0$, thus $\det(A^{\diamond\alpha})$ has the same sign as $\det(U+W)^{\diamond\alpha}$, and the latter determinant is equal to $T_{p,[\alpha]}(n)$, by Theorem 7.3. □

8 The case of lowest rank

In this section we take $p = 1$, in other words, we examine the case that the approximation is done with a matrix V of rank 1, say $V = vv^*$ with $v \in \mathbb{R}^n$, $v \neq 0$. We identify $\mathbb{N}^p$ with $\mathbb{N}$ by identifying $\mathbf{m} = (m)$ with m. The set $\Lambda(1, n)$ consists of all subsets λ of $\mathbb{N}$ with n elements, and $\Lambda(1, n, l)$ consists of those $\lambda \in \Lambda(1, n)$ for which $||\lambda||$ (the sum of all elements of λ) is equal to l. Clearly $\{0, 1, 2, \ldots, n-1\}$ is the element λ of $\Lambda(1, n)$ for which $||\lambda||$ is minimal. Let us denote this element by λ_0 (for $n = 2$ this is the same λ_0 as in the proof of Theorem 7.3). Thus $L(1, n) = ||\lambda_0|| = n(n-1)/2$, and $\Lambda(1, n, L) = \{\lambda_0\}$. For the case $p = 1$ we now obtain from Theorem 7.3 the following result.

Theorem 8.1 *Let $n \geq 2$ be given and let α be a positive non-integer real number. Let $v \in \mathbb{R}^n$ have pairwise distinct coordinates. Set $V = vv^*$. Then*

$$\lim_{\varepsilon \downarrow 0}(\mathrm{sgn}(\det((U + \varepsilon V)^{\diamond\alpha})))$$

is positive *if $\alpha > n-2$ or $n-2 < \alpha + 4k < n$ for some integer $k \geq 1$, and* negative *if $n-4 < \alpha + 4k < n-2$ for some $k \geq 0$.*

Proof As just observed, we have $\Lambda(1, n, L) = \{\lambda_0\}$. As was pointed out in Remark 7.5, the conclusion of Theorem 7.3 (formula (41)) is valid if $S_\lambda \neq 0$ for at least one element of $\Lambda(1, n, L)$, thus if $S_{\lambda_0} \neq 0$. Now $S_{\lambda_0} = |\det(u, v, \ldots, v^{\diamond(n-1)})|$, and it follows from (i) of Proposition 5.8 that this is the case if and only if the coordinates of v are pairwise different.

Furthermore, we have $D_1(k) = \{k+1, k+2\}$ $(k \geq 0)$ (cf. (18)). Thus $T_{1,a}(n)$ is positive if $1 \leq n \leq a+2$, while for $n = a+3$ it is negative (thus $N_{1,a} = a+3$, cf. Example 7.2). It is again negative for $n = a+4$, for the next two values of n it is positive, then negative for the next two values, and so on. The theorem follows. □

Remark 8.2 We can verify the above result in a more direct way. If $\lambda = \{m_1, \ldots, m_n\} \in \Lambda(1, n)$ and $\alpha \in \mathbb{R}\backslash\mathbb{Z}$, then $(\alpha \,|\, \lambda) = \prod_{m \in \lambda} (\alpha \,|\, m) = \prod_{i=1}^{n} \dfrac{\alpha(\alpha-1)\ldots(\alpha-m_i+1)}{m_i!}$. In particular:

$$(\alpha \,|\, \lambda_0) = \frac{\alpha^{n-1}(\alpha-1)^{n-2}\ldots(\alpha-n+3)^2(\alpha-n+2)}{1! \cdot 2! \cdots (n-1)!}.$$

It follows that $(\alpha \,|\, \lambda_0)$ is positive if $\alpha - n + 2 > 0$, thus for $n \leq [\alpha] + 2$. When $n = [\alpha] + 3$ we get a minus sign, and we get two more minus signs for the next n. Therefore, $(\alpha \,|\, \lambda_0)$ is negative for $n = [\alpha] + 3$ and also for $n = [\alpha] + 4$. For the next two values of n it is positive again because we get three and four extra minus signs, respectively. Continuing this argument, the same pattern as in the theorem is obtained.

In the next corollary the meaning of 'sufficiently small' is: so small that the sign of the determinant considered doesn't change anymore when the x_i are replaced by εx_i with $0 < \varepsilon < 1$; cf. the terminology in Theorem 7.6.

Corollary 8.3 *Let $n \geq 2$ be given, and let α be a positive non-integer real number. Let $x_1, \ldots, x_n$ be pairwise different real numbers such that either all x_i are non-negative or all x_i are at most one in absolute value. Let A be the $n \times n$ matrix whose general element is $\dfrac{1+x_ix_j}{\sqrt{(1+x_i^2)(1+x_j^2)}}$. Then if the x_i are sufficiently small,* $\det(A^{\diamond\alpha})$ *is* positive *if $\alpha > n-2$ or $n-2 < \alpha + 4k < n$ for some integer $k \geq 1$, and* negative *if $n-4 < \alpha + 4k < n-2$ for some $k \geq 0$.*

Proof Set $v^* = (x_1, \ldots, x_n)$ and $V = vv^*$. We compute:

$$U + V = (1 + x_ix_j)_{i,j} = \left(\sqrt{(1+x_i^2)(1+x_j^2)}\right)_{i,j} \diamond \left(\frac{1+x_ix_j}{\sqrt{(1+x_i^2)(1+x_j^2)}}\right)_{i,j} = T \diamond A\,,$$

and hence $(U+V)^{\diamond\alpha} = T^{\diamond\alpha} \diamond A^{\diamond\alpha}$ (the 'Hadamard calculus' is convenient for such computations!). Now T has rank 1, hence so has $T^{\diamond\alpha}$. We have $\mathrm{mtr}(T^{\diamond\alpha}) = \prod_i (1+x_i^2)^\alpha$ (cf. (4)), and therefore Lemma 3.8 implies that $\det A^{\diamond\alpha}$ has the same sign as $\det(U+V)^{\diamond\alpha}$. The result now follows from Theorem 8.1. □

Example 8.4 We prove Horn's result mentioned in Section 2. Let $n \geq 3$ be given and let α be a positive non-integer real number satisfying $\alpha < n-2$. We can take $m \in \mathbb{Z}$ with $m \leq n$ such that $m - 3 < \alpha < m - 2$. Take $A_\varepsilon = U + \varepsilon V$ as in Theorem 8.1, with ε sufficiently small. It then follows from that theorem that all principal minors of A_ε^α of size m are negative, and hence, by Proposition 2.2, $A_\varepsilon^\alpha \notin \mathcal{S}_n$, as desired. We note that also the principal minors of the sizes $m+1$, $m+4$, $m+5$, $m+8$, ... are negative, whereas those with size less than m or with size $m+2$, $m+3$, $m+6$, $m+7$, ... are positive.

The reader may find it instructive to compare the above with the concrete example in [3, p. 636]; there $V = vv^*$ with $v^* = (1, 2, \ldots, n)$ (all coordinates distinct!).

We note that it is essential that the matrices A_ε that produce negative principal minors (when taken to a fractional power) have rank 2, i.e. are of as low a rank as possible. And the larger α is, the more essential this low rank is. Indeed, if instead of $p = 1$ we take $p = 2$ (so that the A_ε have rank 3), then, as $N_{2,a} = (a^2 + 5a + 8)/2$, no principal minor as desired is obtained when $n \leq (a+2)(a+3)/2$. For instance, if $10 < \alpha < 11$, the size n has to be at least 79 to obtain (by our technique) negative principal minors of A_ε^α (and when $p = 3$ this minimal size is already 365). One might say: 'the larger the rank of a symmetric positive-definite matrix, the more stable its positive-definiteness is under taking fractional powers'.

Example 8.5 Consider the case $n = 2$. The positive-definite cone at U consists of all matrices A of the form

$$A = \begin{pmatrix} 1+x & 1+z \\ 1+z & 1+y \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} + \begin{pmatrix} x & z \\ z & y \end{pmatrix} = U + V\,,$$

with $x \geq 0$, $y \geq 0$ and $|z| \leq \sqrt{xy}$. The rank of A is 2 unless $(1+x)(1+y) = (1+z)^2$, and, using that $|z| \leq \sqrt{xy}$, one sees that this happens only if $x = y = z$.

The rank of V is 1 if $|z| = \sqrt{xy}$, and 2 if $|z| < \sqrt{xy}$. This example illustrates that (also for general n) the matrices V of rank $p = n-1$ give only a minor part of the matrices $U + V$ of rank n in the positive-definite cone at U. Therefore, it is essential that in Definition 7.1 and Theorem 7.3 also the case $p = n$ is considered. In the same vein, observe that the rank of $U + V$ is $p+1$ if $p < n$ (and V is Hadamard linearly independent), but it is only p if $p = n$.

But the case $n = 2$ is really trivial: if $A \in \mathcal{S}_2^+$ (hence $\det A \geq 0$), then $\det(A^{\diamond\alpha}) \geq 0$ for all $\alpha \geq 0$ because the relation $x_{11}x_{22} \geq x_{12}x_{21}$ is preserved when a positive power is taken (likewise if $\det A > 0$). This agrees with the fact that $(p\,\|\,a+1) \geq (p\,\|\,1) = p + 1 \geq 2$, hence $T_{1,a}(2) = +1$ (cf. Definition 7.1) for all $a \geq 0$ (and $p = 1$ or $p = 2$).

Example 8.6 The case $n = 3$ (and $p = 1$) is a lot more interesting already. Take $v^* = (x, y, z)$ with x, y, z pairwise different and either all non-negative or all at most 1 in absolute value. Explicit computation of the 3×3 determinant, and an application of Corollary 8.3 gives the following inequality, for $0 < \alpha < 1$ and x, y, z sufficiently small:

$$1 + 2\left(\frac{(1+yz)(1+xz)(1+xy)}{(1+x^2)(1+y^2)(1+z^2)}\right)^{\alpha}$$
$$< \left(\frac{(1+yz)^2}{(1+y^2)(1+z^2)}\right)^{\alpha} + \left(\frac{(1+xz)^2}{(1+x^2)(1+z^2)}\right)^{\alpha} + \left(\frac{(1+xy)^2}{(1+x^2)(1+y^2)}\right)^{\alpha}. \tag{42}$$

For $\alpha = 1$ we have equality, and for $\alpha > 1$ the opposite inequality holds, by the same corollary. We also remark that, when e.g. $x = y$, the relation reduces to an equality.

We shall now show, using a technique due to Thiemann [9], that for inequality (42) the unnatural restriction to sufficiently small values of x, y, z can be avoided (cf. Remark 8.13). We start with a definition.

Definition 8.7 *A triple $\{x, y, z\}$ of real numbers will be called* triangular *if $x \le y+z$ and $y \le x+z$ and $z \le x+y$, and* strictly triangular *if these inequalities are strict.*

Remark 8.8 Suppose, without loss of generality, that $x \le y \le z$. If $x < 0$, then $z > x + y$. Therefore, triangularity can occur only for non-negative real numbers. If $x = 0$, then triangularity holds if and only if $y = z$, and in this case two of the inequalities are equalities (or even all three, if $x = y = z = 0$). Finally, if $x > 0$, then triangularity holds if and only if $z \le x + y$, and the triangularity is strict unless $z = x + y$ (regardless of whether x and y are equal or not).

Lemma 8.9 *Let $p, q, r \in [-1, 1]$ and let $x, y, z \in [0, \pi]$ be such that $p = \cos x$, $q = \cos y$, $r = \cos z$. Then $1+2pqr-(p^2+q^2+r^2) \ge 0$ if and only if $\{x, y, z\}$ is triangular and $1+2pqr-(p^2+q^2+r^2) > 0$ if and only if $\{x, y, z\}$ is strictly triangular.*

Proof We compute: $p^2+q^2+r^2-2pqr-1 = (r-pq)^2-(1-p^2)(1-q^2) = (\cos z - \cos x \cos y)^2 - \sin^2 x \sin^2 y$. We shall write $\cos(x-y) = C_1$, $\cos(x+y) = C_2$, $\cos z = D$, for short. Then $p^2+q^2+r^2-2pqr-1 = \left(D - \frac{C_1+C_2}{2}\right)^2 - \left(\frac{C_1-C_2}{2}\right)^2 = D^2 - D(C_1+C_2) + C_1C_2 = (D-C_1)(D-C_2)$. Thus $1+2pqr-(p^2+q^2+r^2) \ge 0$ if and only if $C_1 \ge D \ge C_2$ (note that $C_1 \ge C_2$ because $C_1 - C_2 = 2\sin x \sin y$), thus if and only if $|x-y| \le z \le x+y$. But this is equivalent to $\{x, y, z\}$ being triangular. The other statement follows in the same way, using strict inequalities. □

Lemma 8.10 *Let f be defined, non-negative and non-decreasing on $[0, A]$, for some $A > 0$. Then the following holds.*

(i) *If $t \mapsto \dfrac{f(t)}{t}$ is non-increasing on $(0, A]$, then f preserves triangularity on $[0, A]$.*

(ii) *If $t \mapsto \dfrac{f(t)}{t}$ is strictly decreasing on $(0, A]$, then f transforms triangularity into strict triangularity on $(0, A]$.*

Similarly with $[0, \infty)$ instead of $[0, A]$.

Proof First of all, we can assume, without loss of generality, that f is defined on all of $[0, \infty)$. Indeed, if f is as in the lemma, then its domain can be extended to $[0, \infty)$ by setting $f(t) = f(A) + c\,(t - A)$ with $0 \le c \le A$ $(t > A)$.

Now let $x > 0$ and $y > 0$ be given. In case (i) it follows that $f(x+y) \le \dfrac{x+y}{x} f(x)$ and $f(x+y) \le \dfrac{x+y}{y} f(y)$, and hence $f(x) + f(y) \ge f(x+y)$; in case (ii) we even get $f(x) + f(y) > f(x+y)$. Using this and the fact that f is non-decreasing, we get for any z with $0 \le z \le x+y$ that $f(z) \le f(x+y) \le f(x) + f(y)$, and in case (ii) even $f(z) < f(x) + f(y)$. Therefore, both results follow. □

Lemma 8.11 *Let f be defined (and real-valued) on $[0, A]$, for some $A > 0$. If $f(0) \geq 0$ and f is* concave, *then $t \mapsto \dfrac{f(t)}{t}$ is non-increasing on $(0, A]$. Moreover, if there are x, y with $0 < x < y \leq A$ for which $\dfrac{f(x)}{x} = \dfrac{f(y)}{y}$, then $f(t) = ct$ for some $c \in \mathbb{R}$ and all $t \in [0, y]$.*

Proof If $0 < x < y \leq A$, then, by the concavity of f, we have $f(x) \geq \dfrac{y-x}{y} f(0) + \dfrac{x}{y} f(y)$, thus $\dfrac{f(x)}{x} \geq \dfrac{f(y)}{y} + (\dfrac{1}{x} - \dfrac{1}{y}) f(0) \geq \dfrac{f(y)}{y}$. Thus $\dfrac{f(t)}{t}$ is non-increasing on $(0, A]$. It also follows that $\dfrac{f(x)}{x} = \dfrac{f(y)}{y}$ implies that $f(0) = 0$ and that $\dfrac{f(t)}{t}$ is constant on $[x, y]$.

Secondly, if $0 < z < x < y \leq A$, then $f(x) \geq \dfrac{y-x}{y-z} f(z) + \dfrac{x-z}{y-z} f(y)$. Thus $f(z) \leq \dfrac{y-z}{y-x} f(x) - \dfrac{x-z}{y-x} f(y)$. If $\dfrac{f(x)}{x} = \dfrac{f(y)}{y}$, this gives $\dfrac{f(z)}{z} \leq \dfrac{x(y-z) - y(x-z)}{z(y-x)} \dfrac{f(y)}{y} = \dfrac{f(y)}{y}$. The opposite inequality also holds (as was proved above), thus $\dfrac{f(t)}{t}$ is constant on $(0, y]$. □

From Lemma 8.11 it follows that in Lemma 8.10 concavity of f is a sufficient condition to have $f(t)/t$ non-increasing; moreover, $f(t)/t$ is strictly decreasing, except possibly for an initial interval $[0, B]$ with $0 < B \leq A$ on which f is linear. On the other hand, it is easily seen that concavity is not a necessary condition. For instance, the function f defined by $f(x) = \sqrt{x}$ $(0 \leq x \leq 1)$, $f(x) = x$ $(x \geq 1)$ is not concave, but it satisfies the conditions of Lemma 8.10.

Theorem 8.12 *Let $p, q, r \in [0, 1]$ be such that $1 + 2\,pqr - (p^2 + q^2 + r^2) \geq 0$. Then the following holds.*

(i) $1 + 2\,(pqr)^\alpha - (p^{2\alpha} + q^{2\alpha} + r^{2\alpha}) \geq 0$ *for all* $\alpha \geq 1$.

(ii) *If $p, q, r < 1$, then $1 + 2\,(pqr)^\alpha - (p^{2\alpha} + q^{2\alpha} + r^{2\alpha}) > 0$ for all $\alpha > 1$.*

Proof The relation $p = \cos x$ establishes a one-to-one correspondence between the elements $p \in [0, 1]$ and $x \in [0, \pi/2]$ (cf. Lemma 8.9; this time we take $p \geq 0$ because we want to consider fractional powers). If p corresponds to x, then to p^α corresponds $\arccos((\cos x)^\alpha)$. Therefore we define, for each $\alpha > 0$, the function f_α by $f_\alpha(t) = \arccos((\cos t)^\alpha)$ $(0 \leq t \leq \pi/2)$.

It follows from Lemma 8.9 that for the proof of (i) it suffices to show that the f_α $(\alpha \geq 1)$ preserve triangularity for triples $\{x, y, z\}$ with $x, y, z \in [0, \pi/2]$, and for the proof of (ii) that the f_α $(\alpha > 1)$ transform triangularity into strict triangularity for triples $\{x, y, z\}$ with $x, y, z \in (0, \pi/2]$.

It is clear that the f_α are non-negative and non-decreasing on $[0, \pi/2]$. Therefore, by Lemma 8.10, it suffices to show that the f_α satisfy the assumptions (i) and (ii), respectively, of that lemma, for $\alpha \geq 1$ and $\alpha > 1$, respectively. By Lemma 8.11, to do this it suffices to show that on $[0, \pi/2]$ the f_α are concave for all $\alpha \geq 1$, and even strictly concave for all $\alpha > 1$. This, finally, can be proved by showing that for $0 < t < \pi/2$ one has $f''_\alpha(t) \leq 0$ for all $\alpha \geq 1$, and even $f''_\alpha(t) < 0$ for all $\alpha > 1$.

To determine f''_α, we rewrite the definition of f_α in the form

$$\cos(f_\alpha(t)) = (\cos t)^\alpha \,. \tag{43}$$

Logarithmic differentiation of (43) gives

$$f'_\alpha(t) = \alpha \, \frac{\tan t}{\tan f_\alpha(t)} \qquad (0 < t < \pi/2) \tag{44}$$

(note that indeed $f'_\alpha > 0$, even for all $\alpha > 0$). Differentiating (44) and substituting (43) in the result we obtain

$$f''_\alpha(t) = \frac{\alpha \tan^2 t}{\tan f_\alpha(t) \sin^2 f_\alpha(t)} \left(\frac{\sin^2 f_\alpha(t)}{\sin^2 t} - \alpha \right) \qquad (0 < t < \pi/2)\,.$$

Finally, substituting (43) once again and using a mean value argument, we find that

$$\frac{\sin^2 f_\alpha(t)}{\sin^2 t} - \alpha = \frac{1-(\cos^2 t)^\alpha}{1-\cos^2 t} - \alpha = \alpha\left(\xi^{\alpha-1}-1\right) ,$$

for some ξ satisfying $\cos t < \xi < 1$. It follows that, for all $t \in (0,\pi/2)$, $f''_\alpha(t)$ is negative for $\alpha > 1$, zero for $\alpha = 1$ and positive for $0 < \alpha < 1$. □

Remark 8.13 Part (i) of the above theorem is Thiemann's result; his proof is a combination of the proof of Lemma 8.9 and the above proof, starting from relation (43). Part (ii) is a slight generalization, needed to clarify the link between the requirements 'x, y, z distinct' and '$p, q, r < 1$' in Corollary 8.14 below. Definition 8.7 and the Lemmas 8.10 and 8.11 are added to obtain a wider perspective.

As a corollary we obtain the general validity of the inequality in Example 8.6.

Corollary 8.14 *Let x, y, z be pairwise distinct real numbers such that either $x, y, z \geq 0$ or $x, y, z \in [-1, 1]$. Then inequality (42) holds for all α with $0 < \alpha < 1$; for $\alpha = 1$ there is equality, and for $\alpha > 1$ the opposite inequality holds.*

Proof Denote the difference between the left-hand side and the right-hand side of (42) by $P_\alpha(x, y, z)$. We have to show that $P_\alpha(x, y, z)$ is negative when $0 < \alpha < 1$, zero when $\alpha = 1$, and positive when $\alpha > 1$. First of all, we note that $P_\alpha(x, y, z) = 0$ for all $\alpha \geq 0$ when x, y, z are not pairwise different. Secondly, an easy calculation shows that $P_1(x, y, z) = 0$ (or one may observe that $P_1(x, y, z)$ is the determinant of a 3×3 matrix of rank 2). Now set $p = \dfrac{1+yz}{\sqrt{(1+y^2)(1+z^2)}}$, and similarly for q and r. Then $P_\alpha(x, y, z) = 1 + 2(pqr)^\alpha - (p^{2\alpha} + q^{2\alpha} + r^{2\alpha})$. By the assumption on x, y, z we have $p, q, r \geq 0$. We also have $p, q, r < 1$; for instance, $p = 1$ occurs if and only if $y = z$.

As $P_1(x, y, z) = 0$, it now follows from (ii) of Theorem 8.12 that $P_\alpha(x, y, z) > 0$ if $\alpha > 1$. But it also follows that $P_\alpha(x, y, z) < 0$ if $0 < \alpha < 1$. Indeed, suppose that $P_\alpha(x, y, z) \geq 0$ for such an α; it would then follow, by applying (ii) of Theorem 8.12, not with p, q, r but with $p^{1/\alpha}, q^{1/\alpha}, r^{1/\alpha}$, that $P_1(x, y, z) > 0$. □

9 Sign patterns

In this final section we say a few words on the patterns of plus and minus signs determined by the functions $T_{p,a}$. From Definition 7.1 we know that for fixed p and a the function $n \mapsto T_{p,a}(n)$ is constant (and positive) for $p \leq n \leq (p \,\|\, a+1) = (a+2)(a+3)\ldots(a+p+1)/p!$, after which it is alternating for a while, then constant again, and so on. Let us call an interval $D_p(a+t)$ an 'interval of constancy' (relative to p and a) when t is even, and an 'interval of alternation' when t is odd. Recall that adjacent intervals have one point in common.

Whether the value of $T_{p,a}$ on two subsequent intervals of constancy is the same or opposite, depends on the parity of the length of the interval of alternation between these two intervals. This is specified in the next lemma.

Lemma 9.1 *Let $p \geq 1$ and $a \geq 0$ be integers and let $t \geq 1$ be an odd integer. The function $n \mapsto T_{p,a}(n)$ will have on $D_p(a+t+1)$ the* same *(constant) value as on $D_p(a+t-1)$ if $(a+t+1 \,\|\, p-1)$ is even, and the* opposite *value if it is odd.*

Proof If an interval of alternation has an odd number of elements, then the value of $T_{p,a}$ will be the same at its two end points, and hence on the two adjacent intervals of constancy; in the opposite case the value on the two adjacent intervals will be opposite. The result now follows from (19). □

The parity of the binomial coefficients in Lemma 9.1 can be visualized by a Pascal triangle that gives only the parity (the pattern of this triangle is that of the Sierpińsky gasket; cf. [1, pp. 10-11]). Binomial coefficients are more often even than odd; for instance, among the 4^N coefficients $(m \,\|\, k)$ with $m, k < 2^N$ there are only 3^N odd ones. Hence, from $N = 3$ on the even ones are in the majority. As a consequence, by Lemma 9.1, for matrices of not too small rank it will be more common to have subsequent intervals of constancy of equal sign than of opposite sign.

Can it happen that $T_{p,a}$ has the same (necessarily positive) sign on *all* intervals of constancy? It turns out that this occurs surprisingly often; in fact, in one quarter of all cases. To prove this, we need a lemma on the parity of binomial coefficients; its (elementary) proof is left to the reader.

Lemma 9.2 *Suppose that k and i are integers satisfying $0 \le i \le k$. Then the following properties hold.*

(i) *If k is even and i is odd, then $(k \mid i)$ is even.*

(ii) *For even i there are arbitrarily large even k such that $(k \mid i)$ is odd.*

(iii) *For arbitrary i there are arbitrarily large odd k such that $(k \mid i)$ is odd.* □

Let us call an integer n an *isolated* element of a set of integers E if $n \in E$ but $n - 1 \notin E$ and $n + 1 \notin E$. The following remarkable theorem can then be proved.

Theorem 9.3 *Let $p \ge 1$ and $a \ge 0$ be given. Then the set of integers $n \ge p$ for which $T_{p,a}(n) = -1$ has only isolated elements if and only if p is even and a is odd.*

Proof Let $a \ge 1$, $p \ge 0$, and $t \ge 1$ be fixed. If a is odd and p is even, then $p - 1$ is odd and $a + p + 1$ is even for all odd t, and hence, by Lemma 9.1 and (i) of Lemma 9.2, $T_{p,a}$ has the same value on all intervals of constancy. On $D_{p,a}(0)$, the first interval of constancy, $T_{p,a}$ is positive, hence $T_{p,a}$ is positive on all intervals of constancy.

It remains to show that $(a + t + 1 \mid p - 1)$ can be odd for arbitrarily large odd t, whenever a is even or p is odd. Now, if a is even and p is odd, this follows from (ii) of Lemma 9.2, while for a and p both even or both odd, it follows from (iii) of the same lemma. □

Example 9.4 Suppose that $p = 2^N$ for some positive N. Then $(p - 1)!$ contains relatively few factors 2, and hence $(a + t + 1 \,\|\, p - 1)$ 'has a good chance to be even'. Actually, it is easily seen that the smallest k for which $(k \,\|\, 2^N - 1)$ is odd is $k = 2^N - 1$. Now take $a = 0$. By Theorem 9.3, $T_{2^N,0}$ has intervals of constancy where it takes the value -1. The first such interval is $D_{2^N}(2^N)$, as follows from Lemma 9.1, together with the above-mentioned smallest k.

As an example, consider the case that $N = 4$, thus $p = 16$. For $16 \le n \le 601080389$ the value of $T_{16,0}(n)$ is negative for $n = 18, 20, \ldots, 152, 970, 972, \ldots, 4844, 20350, \ldots$, i.e. for all even n such that $(2s - 1 \,\|\, 16) < n < (2s \,\|\, 16)$ for an s with $1 \le s \le 8$, altogether (cf. (11)) for $\frac{1}{2}\sum_{s=1}^{8}(2s \,\|\, 15) = 199650082$ values of n (all even and all isolated), and positive for the other 401430292 values. But for the next 565722721 values of n the $T_{p,a}(n)$ are all negative. The number of negative values hence increases to 765372803, an increase from 33.2 to 65.6 percent. To state a very concrete case: imagine a (Hadamard independent) matrix of size 10^9, of rank 17, in the positive-definite cone of the unit matrix U and sufficiently close to U: a vast field of 10^{18} numbers, all very close to 1. Then the determinant of the Hadamard square root of that matrix is negative. The same is the case for sizes n satisfying $601080390 \le n \le 1166803110$; but not so for the previous, nor for the next n.

References

[1] G. A. Edgar, *Measure, topology, and fractal geometry.* Springer–Verlag, New York 1990.

[2] P. Fischer, J. D. Stegeman, Functions operating on positive-definite symmetric matrices. *Journal of Mathematical Analysis and Applications* 171, 461-470 (1992).

[3] C. H. FitzGerald, R. A. Horn, On fractional Hadamard powers of positive definite matrices. *Journal of Mathematical Analysis and Applications* 61, 633-642 (1977).

[4] R. A. Horn, The theory of infinitely divisible matrices and kernels. *Transactions of the American Mathematical Society* 136, 269-286 (1969).

[5] R. A. Horn, The Hadamard product. In: C. R. Johnson (editor), *Matrix theory and applications.* Proceedings of Symposia in Applied Mathematics 40, 87-168. AMS 1990.

[6] R. A. Horn, C. R. Johnson, *Matrix analysis.* Cambridge University Press 1985.

[7] R. A. Horn, C. R. Johnson, *Topics in matrix analysis.* Cambridge University Press 1991.

[8] G. Pólya, G. Szegö, *Problems and theorems in analysis,* Vol. 2. Springer–Verlag, Berlin 1976.

[9] J. G. F. Thiemann, *Personal communication* (1990).

Addresses of the authors:

P. Fischer
Department of Mathematics and Statistics
University of Guelph
Guelph, Ontario N1G 2W1
Canada
email: pfischer@uoguelph.ca

J. D. Stegeman
Department of Mathematics
Utrecht University
P.O. Box 80010
3508 TA Utrecht
The Netherlands
email: stegeman@math.uu.nl