Matrix Analysis and Positivity Preservers

Apoorva Khare
Indian Institute of Science
1. Introduction

This text arose out of the course notes for Math 341: Matrix Analysis and Positivity, a one-semester course offered in Spring 2018 (and again in Fall 2019) at the Indian Institute of Science. The goal in this course is to briefly describe some notions of positivity in matrix theory, followed by our main focus: a detailed study of the operations that preserve these notions (and in the process, to understand some aspects of real functions). There are several different notions of positivity in analysis, studied for classical and modern reasons that are also touched upon in these notes. These include:

- Positive semidefinite (psd) and positive definite matrices.
- Entrywise positive matrices.
- A common strengthening of both of these notions, which involves totally positive (TP) and totally non-negative (TN) matrices.
- Settings somewhat outside matrix theory. For instance, consider discrete data associated to positive measures on locally compact abelian groups $G$. E.g. for $G = \mathbb{R}$, one obtains moment sequences, which are intimately related to positive semidefinite Hankel matrices. For $G = S^1$, the circle group, one obtains Fourier–Stieltjes sequences, which are connected to positive semidefinite Toeplitz matrices. (Works of Carathéodory, Hamburger, Hausdorff, Herglotz, and Stieltjes, among others.)
- More classically, positive definite functions and kernels have long been studied in analysis, on locally compact groups or metric spaces. (Bochner, Schoenberg, von Neumann, Pólya, to name a few.)

These notes begin by discussing the above notions, focussing on their properties and some results in matrix theory. The remainder of the notes then studies in detail the preservers of several of these notions of positivity. Just as it is useful to study objects in a suitable category, such as sets or metric spaces, but more enriching to study morphisms between them, such as (continuous) functions; similarly, the study of sets of positive matrices is enriched by studying operations between these sets. For instance, several classes of ‘positive’ matrices repeatedly get highlighted by way of studying positivity preservers – these include generalized Vandermonde matrices and Hankel moment matrices.

Specifically, in these notes we will study entrywise preservers of (various notions of) positivity. The question of why entrywise calculus – as compared to the usual holomorphic functional calculus – has a rich and classical history in the analysis literature, beginning with the work of Schoenberg, Rudin, Loewner, and Horn (these results are proved in Part 3 of these notes); but also drawing upon earlier works of Menger, Schur, Bochner, and others. Interestingly, this entrywise calculus also arises in modern-day applications from high-dimensional covariance estimation. (We elaborate on this in Section 13.1 and briefly also in Section 14.) Furthermore, this evergreen area of mathematics continues to be studied in the literature, and draws techniques from – and also contributes to – symmetric function theory, statistics and graphical models, combinatorics, and linear algebra (in addition to analysis).

In a sense, the course and these notes arose out of research carried out in significant measure by mathematicians at Stanford University (and their students) over the years. This includes Loewner and Karlin, and their students: FitzGerald, Horn, Micchelli, and Pinkus. Slightly less directly, there was also Katznelson, who had previously worked with Helson, Kahane, and Rudin, leading to Rudin’s strengthening of Schoenberg’s theorem. (Coincidentally, Pólya and Szegő, who made the original observation on entrywise preservers of positivity using the Schur product theorem, were again colleagues at Stanford.) On a personal note, the author’s contributions to this area also have their origins in his time spent at Stanford University,
collaborating with Alexander Belton, Dominique Guillot, Mihai Putinar, Bala Rajaratnam, and Terence Tao (though the collaboration with the last-named colleague was carried out almost entirely at the Indian Institute of Science).

We now describe the course, these notes, and their mathematical contents in detail. The notes were scribed by the students taking the course in Spring 2018 at IISc (followed by ‘homogenization’ by the author). Each section was originally intended to cover the notes of roughly one 90-minute lecture (or occasionally two); that said, some material has subsequently been moved around for logical, mathematical, and expositional reasons. The notes, and the course itself, require an understanding of basic linear algebra and calculus/analysis, perhaps with a bit of measure theory as well. Beyond these basic topics, we have tried to keep these notes as self-contained (with full proofs) as possible. To that end, we have included proofs of ‘preliminary’ results, including:

(i) early results of Schoenberg (following Menger, Fréchet, Bochner, von Neumann…) connecting metric geometry and positive definite functions to matrix positivity;
(ii) the Boas–Widder and Bernstein theorems on functions with positive iterated divided differences;
(iii) an extension to normed linear spaces, of (a special case of) a classical result of Ostrand on mid-convexity and local boundedness implying continuity;
(iv) Whitney’s density theorem for totally positive matrices inside totally non-negative matrices;
(v) Fekete’s result on totally positive matrices;
(vi) the discreteness of zeros of real analytic functions (and a sketch of the continuity of roots of complex polynomials);
(vii) Perron’s theorem (the precursor to Perron–Frobenius);
(viii) compound matrices and Kronecker’s theorem on their spectra;
(ix) Sylvester’s criterion and the Schur product theorem on positive (semi)definiteness (also, Jacobi’s formula);
(x) matrix identities by Andrieu and Cauchy–Binet, as well as the ‘continuous’ generalization of the latter;
(xi) Descartes’ rule of signs;
(xii) and the equivalence of Cauchy’s and Littlewood’s definitions of Schur polynomials (and of the Jacobi–Trudi and von Någelsbach–Kostka identities) via Lindström–Gessel–Viennot bijections, among other results.

Owing to considerations of time, we had to leave out some proofs, including of: theorems by Hamburger/Hausdorff/Stieltjes, Cauchy, Montel, and Morera; a Schur positivity phenomenon for ratios of Schur polynomials; the dominated convergence theorem; as well as the closure of real analytic functions under composition. Most of these can be found in multiple standard textbooks in mathematics. Nevertheless, as the previous and current paragraphs indicate, these notes cover many classical results by past experts, and acquaint the reader with a variety of tools in analysis (especially the study of real functions) and in matrix theory – many of these tools are not found in more ‘traditional’ courses on these subjects.

These notes are broadly divided into four parts. In part 1, the key objects of interest – namely, positive semidefinite / totally positive / totally non-negative matrices – are introduced, together with some basic results as well as some important classes of examples. In part 2, we begin the study of functions acting on such matrices entrywise, and preserving the relevant notion of positivity. Here we will restrict ourselves to studying power functions that act on various sets of matrices of a fixed size. This is a long-studied question, including
by Bhatia, Elsner, Fallat, FitzGerald, Hiai, Horn, Jain, Johnson, and Sokal; as well as by the author in collaboration with Guillot and Rajaratnam. We also obtain certain necessary conditions on general entrywise functions that preserve positivity, including multiplicative mid-convexity and continuity. We explain some of the modern motivations, and end with some problems that are the subject of current research.

**Part 3** deals with some of the foundational results on matrix positivity preservers. After mentioning some of the early history of the field – including work by Menger, Fréchet, Bochner, and Schoenberg – we present an essentially self-contained proof of the classification of entrywise functions that preserve positive semidefiniteness (= positivity) in all dimensions, or total non-negativity on Hankel matrices of all sizes. This is (a strengthening of) a celebrated result of Schoenberg – later strengthened by Rudin – which is a converse to the Schur product theorem. The proof given in these notes is different from the previous approaches of Schoenberg and Rudin, is essentially self-contained, and uses relatively less sophisticated machinery compared to *loc. cit.* Moreover, it goes through proving a variant by Vasudeva, for matrices with only positive entries; and it lends itself to a multivariate generalization (which will not be covered in these notes). The starting point of these proofs is a necessary condition for entrywise preservers in fixed dimension, proved by Loewner and Horn in the late 1960s. To this day, this result remains essentially the only known condition in fixed dimension \( n \geq 3 \), and a proof of a stronger version is also provided in these notes. In addition to techniques and ingredients introduced by the above authors, these notes also borrow from the author’s joint work with Belton, Guillot, and Putinar.

In the final **part 4**, we return to the study of entrywise functions preserving positivity in fixed dimension. This is a challenging problem – it is still open in general, even for \( 3 \times 3 \) matrices – and we restrict ourselves in this part to studying polynomial preservers. By the Schur product theorem, if the polynomial has all coefficients non-negative then it is easily seen to be a preserver; but interestingly, until 2016 not a single example was known of any other entrywise polynomial preserver of positivity in a fixed dimension \( n \geq 3 \). Very recently, this question has been answered to some degree of satisfaction (by the author in collaboration – first with Belton, Guillot, and Putinar; and subsequently with Tao). These notes end by covering some of this recent progress.

To conclude, thanks are owed to the scribes (listed below), as well as to Alexander Belton, Shabarish Chenakkod, Projesh Nath Choudhury, Julian R. D’Costa, Dominique Guillot, Prakhar Gupta, Sarvesh Ravichandran Iyer, Poomenud Kumar, Paramita Pramanick, Mihai Putinar, Shubham Rastogi, Kartik Singh, G.V. Krishna Teja, Raghavendra Tripathi, Prateek Kumar Vishwakarma, and Pranjal Warade for helpful suggestions that improved the text. I am of course deeply indebted to my collaborators for their support and all of their research efforts in positivity – but also for many stimulating discussions, which helped shape my thinking about the field as a whole and the structure of this text in particular. Finally, I am grateful to the University Grants Commission (UGC, Govt. of India); the Science and Engineering Research Board (SERB) and the Department of Science and Technology (DST) of the Govt. of India; the Infosys Foundation; and the Tata Trusts – for their support through a CAS-II grant; through a MATRICS grant and the Ramanujan and Swarnajayanti Fellowships; through a Young Investigator Award; and through their Travel Grants, respectively.

Apoorva Khare
Indian Institute of Science
& Analysis and Probability Research Group
Bangalore, India
List of scribes: Each section denotes the material covered in a lecture (more or less), in some cases tweaked by the author in order to better structure the notes – so the notes may be longer or shorter than 90 minutes’ worth. Moreover, Sections 4 and 8 in Part 1, as well as the Appendices, were not covered in the lectures in Spring 2018. The notes also include two substantial sets of ‘general remarks’ by the author; see Sections 13.1 and 14.

Part 1: Preliminaries
(1) January 03: (Much of the above text.)
(2) January 08: Sarvesh Iyer
(3) January 10: Prateek Kumar Vishwakarma
(4) Apoorva Khare
(5) January 17: Prakhar Gupta
(6) January 19: Pranjal Warade
(7) January 24: Swarnalipa Datta
(8) Apoorva Khare

Part 2: Entrywise powers preserving (total) positivity in fixed dimension
(9) January 31: Raghavendra Tripathi
(10) February 02: Pabitra Barman
(11) February 07: Kartick Ghosh
(12) February 09: K. Philip Thomas (and Feb 14: Pranab Sarkar)
(13) February 28: Ratul Biswas (with an introduction by Apoorva Khare)
(14) March 02: Prateek Kumar Vishwakarma (with remarks by Apoorva Khare)

Part 3: Entrywise functions preserving positivity in all dimensions
(15) February 14, 16: Pranab Sarkar, Pritam Ganguly, and Apoorva Khare
(16) March 07: Shubham Rastogi
(17) March 09: Poornendu Kumar, Sarvesh Iyer, and Apoorva Khare
(18) March 14: Poornendu Kumar and Shubham Rastogi
(19) March 16: Lakshmi Kanta Mahata and Kartick Ghosh
(20) March 19, 21: Kartick Ghosh, Swarnalipa Datta, and Apoorva Khare
(21) March 23: Sarvesh Iyer and Raghavendra Tripathi
(22) Appendix A: Poornendu Kumar, Shubham Rastogi, and Apoorva Khare

Part 4: Entrywise polynomials preserving positivity in fixed dimension
(23) March 26: Prakhar Gupta, Pranjal Warade, and Apoorva Khare
(24) March 28: Ratul Biswas and K. Philip Thomas
(25) April 02: Pritam Ganguly and Pranab Sarkar
(26) April 11: Pabitra Barman, Lakshmi Kanta Mahata, and Apoorva Khare
(27) Appendix B: Apoorva Khare
Contents

1. Introduction 3

Part 1: Preliminaries

2. The cone of positive semidefinite matrices. Examples. Schur complements. 11
4. Fekete's result. Hankel moment matrices are TN. Hankel psd and TN matrices. 23
5. Generalized Vandermonde matrices. Cauchy–Binet formula and its generalization. 27
6. Hankel moment matrices are TP. Andriev's identity. Density of TP in TN. 31
7. Density of symmetric TP matrices. (Non-)Symmetric TP completion problems. 35
8. Theorems of Perron and Kronecker. Spectra of TP and TN matrices. 39

Part 2: Entrywise powers preserving (total) positivity in fixed dimension

9. Entrywise powers preserving positivity in fixed dimension: I. 45
10. Entrywise powers preserving total positivity: I. 49
11. Entrywise powers preserving total positivity: II. 53
12. Entrywise functions preserving total positivity. Mid-convex implies continuous. 57
   The test set of Hankel TN matrices.
13. Entrywise powers (and functions) preserving positivity: II. Matrices with zero patterns. 63
14. Entrywise powers preserving positivity: III. Chordal and other graphs. Loewner monotonicity and super-additivity. 69

Part 3: Entrywise functions preserving positivity in all dimensions

15. History: Schoenberg's theorem. Rudin, Herz, Vasudeva. Metric geometry, positive definite functions, spheres, and correlation matrices. 77
17. The stronger Horn–Loewner theorem. Mollifiers. 91
18. The stronger Vasudeva and Schoenberg theorems. Bernstein’s theorem. Moment-sequence transforms. 95
20. Proof of stronger Schoenberg Theorem: II. Smoothness implies real analyticity. 107
22. Appendix A: The Boas–Widder theorem on functions with positive differences. 113

Part 4: Entrywise polynomials preserving positivity in fixed dimension

23. Entrywise polynomial preservers. Horn-type necessary conditions. Classification of sign patterns. 125
24. Polynomial preservers for generic rank-one matrices. Schur polynomials. 129
25. First-order approximation / leading term of Schur polynomials. From rank-one matrices to all matrices. 135
26. Exact quantitative bound: monotonicity of Schur ratios. Real powers and power series. 139
Part 1: Preliminaries
Part 1: Preliminaries


In these notes, we will assume familiarity with linear algebra and a first course in calculus/analysis. To set notation: a capital letter with a two integer subscript (such as $A_{m \times n}$) represents a matrix with $m$ rows and $n$ columns. If $m, n$ are clear from context or unimportant, then they will be omitted. Three examples of real matrices are $0_{m \times n}, 1_{m \times n}, \text{Id}_{n \times n}$, which are the (rectangular) matrix consisting of all zeros, all ones, and the identity matrix, respectively. The entries of a matrix $A$ will be denoted $a_{ij}, a_{jk}$, etc. Vectors are denoted by smaller case letters (occasionally in bold), and are columnar in nature. All matrices, unless specified otherwise, are real; and similarly, all functions, unless otherwise specified otherwise, are defined on $-\mathbb{R}^m$ for some $m$.

2.1. Preliminaries. We begin with several basic definitions.

Definition 2.1. A matrix $A_{n \times n}$ is said to be symmetric if $a_{jk} = a_{kj}$ for all $1 \leq j, k \leq n$. A symmetric matrix $A_{n \times n}$ is said to be positive semidefinite (psd) if the real number $x^T A x$ is non-negative for all $x \in \mathbb{R}^n$ — in other words, the quadratic form given by $A$ is positive semidefinite. If, furthermore, $x^T A x > 0$ for all $x \neq 0$ then $A$ is said to be positive definite.

The spectrum of $A$, denoted $\sigma(A) \subset \mathbb{C}$, is the multiset of all eigenvalues of $A$ (counted with multiplicities).

We state the spectral theorem for symmetric (i.e. self-adjoint) operators without proof.

Theorem 2.2 (Spectral theorem for symmetric matrices). For $A_{n \times n}$ a symmetric matrix, $A = U^T D U$ for some orthogonal matrix $U$ (that is, $U^T U = \text{Id}$) and (real) diagonal matrix $D$. $D$ contains all the eigenvalues of $A$ in some order (and counting multiplicities) along its diagonal.

As a consequence, $A = \sum_{j=1}^{n} \lambda_j v_j v_j^T$, where each $v_j$ is an eigenvector for $A$ with real eigenvalue $\lambda_j$, and the $v_j$ together form an orthonormal basis of $\mathbb{R}^n$. (Note that the $v_j$ are the columns of $U^T$.)

We also have the following consequence for two commuting matrices:

Theorem 2.3 (Spectral theorem for commuting symmetric matrices). Let $A_{n \times n}$ and $B_{n \times n}$ be two commuting symmetric matrices. Then $A$ and $B$ are simultaneously diagonalizable, i.e., for the same orthogonal $U$, $A = U^T D_1 U$ and $B = U^T D_2 U$ for $D_1$ and $D_2$ diagonal matrices (whose diagonal entries comprise the spectra $\sigma(A), \sigma(B)$ respectively).

2.2. Criteria for positive (semi)definiteness. We write down several equivalent criteria for positive (semi)definiteness. There are three initial criteria which are easy to prove, and a final criterion which requires separate treatment.

Theorem 2.4 (Criteria for positive (semi)definiteness). Given $A_{n \times n}$ a real symmetric matrix of rank $0 \leq r \leq n$, the following are equivalent:

1. $A$ is positive semidefinite (respectively, positive definite).
2. $\sigma(A) \subseteq [0, \infty)$. (Respectively, $\sigma(A) \subseteq (0, \infty)$.)
3. There exists a matrix $B \in \mathbb{R}^{r \times n}$ of rank $r$ such that $B^T B = A$. (In particular, if $A$ is positive definite then $B$ is square and non-singular.)

Proof. We prove only the positive semidefinite statements. The corresponding positive definite statements will follow with minor changes.
If (1) holds and λ is any eigenvalue, with corresponding (nonzero) eigenvector \( x \), then 
\[ x^T A x = \lambda \|x\|^2 \geq 0. \]
Hence \( \lambda \geq 0 \), proving (2). Conversely, if (2) holds then by the spectral theorem, 
\[ A = \sum_j \lambda_j v_j v_j^T, \]
with all \( \lambda_j \geq 0 \). Now for any \( x \in \mathbb{R}^n \),
\[ x^T A x = \sum_j \lambda_j x^T v_j v_j^T x = \sum_j \lambda_j (x^T v_j)^2 \geq 0, \]
and so \( A \) is positive semidefinite.

Next if (1) holds then write \( A = U^T D U \) by the spectral theorem; note that \( D = U A U^T \) has
the same rank as \( A \). Since \( D \) has non-negative diagonal entries \( d_{jj} \), it has a square root \( \sqrt{D} \), which is a diagonal matrix with diagonal entries \( \sqrt{d_{jj}} \). Write 
\[ D = \begin{pmatrix} D'_{r \times r} & 0 \\ 0 & 0_{(n-r) \times (n-r)} \end{pmatrix}, \]
where \( D' \) is a diagonal matrix with positive diagonal entries. Correspondingly, write 
\[ U = \begin{pmatrix} P_{r \times r} & Q \\ R & S_{(n-r) \times (n-r)} \end{pmatrix}. \]
If we set 
\[ B := (\sqrt{D'} P \mid \sqrt{D'} Q)_{r \times n}, \]
then it is easily verified that
\[ B^T B = \begin{pmatrix} P^T D' P & P^T D' Q \\ Q^T D' P & Q^T D' Q \end{pmatrix} = U^T D U = A. \]

Hence (1) \( \implies (3) \). Conversely, if (3) holds then \( x^T A x = \|Bx\|^2 \geq 0 \) for all \( x \in \mathbb{R}^n \).
Hence \( A \) is positive semidefinite. Moreover, we claim that \( B \) and \( B^T B \) have the same null space and hence the same rank. Indeed, if \( Bx = 0 \) then \( B^T B x = 0 \), while
\[ B^T B x = 0 \implies x^T B^T B x = 0 \implies \|Bx\|^2 = 0 \implies Bx = 0. \]

**Corollary 2.5.** For any real symmetric matrix \( A_{n \times n} \), the matrix \( A - \lambda_{\text{min}} \text{Id}_{n \times n} \) is positive semidefinite, where \( \lambda_{\text{min}} \) denotes the smallest eigenvalue of \( A \).

We now state Sylvester’s criterion for positive (semi)definiteness. This requires some additional notation.

**Definition 2.6.** Given an integer \( n > 0 \), define \([n] := \{1, \ldots, n\} \). Now given a matrix \( A_{m \times n} \) and subsets \( J \subseteq [m] \), \( K \subseteq [n] \), define \( A_{J \times K} \) to be the submatrix of \( A \) with entries \( a_{jk} \) for \( j \in J \), \( k \in K \) (always considered to be arranged in increasing order in these notes). If \( J, K \) have the same size then \( \det A_{J \times K} \) is called a **minor** of \( A \). If \( A \) is square and \( J = K \) then \( A_{J \times K} \) is a **principal submatrix** of \( A \), and \( \det A_{J \times K} \) is a **principal minor**. The principal submatrix (and principal minor) are **leading** if \( J = K = \{1, \ldots, m\} \) for some \( 1 \leq m \leq n \).

**Theorem 2.7** (Sylvester’s criterion). A symmetric matrix is positive semidefinite (respectively, positive definite) if and only if all its principal minors are non-negative (respectively, positive).

We will prove this with the help of some lemmas.

**Lemma 2.8.** If \( A_{n \times n} \) is a positive semidefinite (respectively, positive definite) matrix, then so are all principal submatrices of \( A \).

**Proof.** Fix a subset \( J \subseteq [n] = \{1, \ldots, n\} \) (so \( B := A_{J \times J} \) is the corresponding principal submatrix of \( A \)), and let \( x \in \mathbb{R}^{|J|} \). Define \( x' \in \mathbb{R}^n \) to be the vector such that \( x'_j = x_j \) for all \( j \in J \) and 0 otherwise. It is easy to see that \( x^T B x = (x')^T A x' \). Hence \( B \) is positive (semi)definite if \( A \) is.
As a corollary, all the principal minors of a positive semidefinite (positive definite) matrix are non-negative (positive), since the corresponding principal submatrices have non-negative (positive) eigenvalues and hence non-zero (positive) determinants. So one direction of Sylvester’s criterion trivially holds.

**Lemma 2.9.** Sylvester’s criterion is true for positive definite matrices.

**Proof.** We induct on the dimension of the matrix $A$. Suppose $n = 1$. Then $A$ is just an ordinary real number, whence its only principal minor is $A$ itself, and so the result is trivial.

Now, suppose the result is true for matrices of dimension $\leq n - 1$. We claim that $A$ has at least $n - 1$ positive eigenvalues. To see this, let $\lambda_1, \lambda_2 \leq 0$ be eigenvalues of $A$. Let $W$ be the $n - 1$ dimensional subspace of $\mathbb{R}^n$ with last entry 0. If $v_j$ are orthogonal eigenvectors for $\lambda_j$, $j = 1, 2$, then the span of the $v_j$ must intersect $W$ non-trivially, since the sum of dimensions of these two subspaces of $\mathbb{R}^n$ exceeds $n$. Define $u := c_1 v_1 + c_2 v_2 \in W$; then $u^T A u = u^T A|_W u > 0$, since $A|_W$ is positive definite (by Lemma 2.8). However, $u^T A u = (c_1 v_1^T + c_2 v_2^T) A (c_1 v_1 + c_2 v_2) = c_1^2 \lambda_1 ||v_1||^2 + c_2^2 \lambda_2 ||v_2||^2 \leq 0$ giving a contradiction, and proving the claim.

Now since the determinant of $A$ is positive (it is the minor consisting of $A$ itself), it follows that all eigenvalues are positive, completing the proof. □

We will now prove Jacobi’s formula, an important result in its own right. A corollary of this result will be used, along with the previous result and the idea that positive semidefinite matrices can be expressed as entrywise limits of positive definite matrices, to prove Sylvester’s criteria for all positive semidefinite matrices.

**Theorem 2.10** (Jacobi’s formula). Let $A_t : \mathbb{R} \rightarrow \mathbb{R}^{n \times n}$ be a matrix-valued differentiable function. Then,

$$\frac{d}{dt} (\det A_t) = \text{tr} \left( \text{adj}(A_t) \frac{dA_t}{dt} \right),$$

(2.11)

where $\text{adj}(A_t)$ denotes the adjugate matrix of $A_t$.

**Proof.** The first step is to compute the differential of the determinant. We claim that for any $n \times n$ real matrices $A, B$,

$$d(\det(A))(B) = \text{tr}(\text{adj}(A)B).$$

As a special case, at $A = \text{Id}_{n \times n}$, the differential of the determinant is precisely the trace.

To show the claim, we need to compute the directional derivative

$$\lim_{\epsilon \rightarrow 0} \frac{\det(A + \epsilon B) - \det A}{\epsilon}.$$

The fraction is a polynomial in $\epsilon$ with vanishing constant term (e.g., set $\epsilon = 0$ to see this); and we need to compute the coefficient of the linear term. Expand $\det(A + \epsilon B)$ using the Laplace expansion as a sum over permutations $\sigma \in S_n$; now each individual summand $(-1)^\sigma \prod_{k=1}^n (a_{\sigma(k)} + \epsilon b_{\sigma(k)})$ splits as a sum of $2^n$ terms. (It may be illustrative to try and work out the $n = 3$ case by hand.) From these $2^n \cdot n!$ terms, choose the ones that are linear in $\epsilon$. For each $1 \leq i, j \leq n$, there are precisely $(n - 1)!$ terms corresponding to $b_{ij}$; and added together, they equal the $(i, j)$th cofactor $C_{ij}$ of $A$ – which equals $\text{adj}(A)_{ji}$. Thus, the coefficient of $\epsilon$ is:

$$d(\det(A))(B) = \sum_{i, j=1}^n C_{ij} b_{ij},$$

and this is precisely $\text{tr}(\text{adj}(A)B)$, as claimed.

More generally, the above argument shows that if $B(\epsilon)$ is any family of matrices, with limit $B(0)$ as $\epsilon \to 0$, then

$$\lim_{\epsilon \to 0} \frac{\det(A + \epsilon B(\epsilon)) - \det A}{\epsilon} = \text{tr}(\text{adj}(A)B(0)). \quad (2.12)$$

Returning to the proof of the theorem, for $\epsilon \in \mathbb{R}$ small and $t \in \mathbb{R}$ we write:

$$A_{t+\epsilon} = A_t + \epsilon B(\epsilon)$$

where $B(\epsilon) \to B(0) := dA_t/dt$ as $\epsilon \to 0$, by definition. Now compute using (2.12):

$$\frac{d}{dt} (\det A_t) = \lim_{\epsilon \to 0} \frac{\det(A_t + \epsilon B(\epsilon)) - \det A_t}{\epsilon} = \text{tr}(\text{adj}(A_t)dA_t),$$

as desired. \(\square\)

With these results in hand, we can finish the proof of Sylvester’s criterion for positive semidefinite matrices.

Proof of Theorem 2.7. For positive definite matrices, the result was proved in Lemma 2.9.

Now suppose $A_{n \times n}$ is positive semidefinite. We show the result by induction on $n$, with an easy argument for $n = 1$ similar to the positive definite case.

Now suppose the result holds for matrices of dimension $\leq n - 1$, and let $A_{n \times n}$ have all principal minors non-negative. Let $B$ be any principal minor of $A$, and define $f(t) := \det(B + t \text{Id}_{n \times n})$. Note that $f'(t) = \text{tr}(\text{adj}(B + t \text{Id}_{n \times n}))$ by Jacobi’s formula (2.11).

We claim that $f'(t) > 0$ for all $t > 0$. Indeed, all diagonal entries of $\text{adj}(B + t \text{Id}_{n \times n})$ are proper principal minors of $A + t \text{Id}_{n \times n}$, which is positive definite since $x^T(A + t \text{Id}_{n \times n})x = x^TAx + t\|x\|^2$ for $x \in \mathbb{R}^n$. The claim now follows using Lemma 2.8 and the induction hypothesis.

From the claim, we obtain: $f(t) > f(0) = \det B \geq 0$ for all $t > 0$. Therefore, all principal minors of $A + tI$ are positive, and by Sylvester’s criterion for positive definite matrices, $A + tI$ is positive definite for all $t > 0$. Now note that $x^T(A + t \text{Id}_{n \times n})x = x^TAx + t\|x\|^2$; therefore the non-negativity of the right-hand side implies that of the left-hand side for all $x \in \mathbb{R}^n$, completing the proof. \(\square\)

2.3. Examples of positive semidefinite matrices. We next discuss several examples of positive semidefinite matrices.

2.3.1. Gram matrices.

Definition 2.13. For any finite set of vectors $x_1, \ldots, x_n \in \mathbb{R}^m$, their Gram matrix is given by $\text{Gram}(\{x_j\}) := (\langle x_j, x_k \rangle)_{1 \leq j, k \leq n}$.

A correlation matrix is a positive semidefinite matrix with ones on the diagonal.

In fact we need not use $\mathbb{R}^m$ here; any inner product space/Hilbert space is sufficient.

Proposition 2.14. Given a real symmetric matrix $A_{n \times n}$, it is positive semidefinite if and only if there exist an integer $m > 0$ and vectors $x_1, \ldots, x_n \in \mathbb{R}^m$ such that $A = \text{Gram}(\{x_j\})$.

As a special case, correlation matrices precisely correspond to those Gram matrices for which the $x_j$ are unit vectors.
2. The cone of positive semidefinite matrices. Examples. Schur complements. 15

Proof. If $A$ is positive semidefinite, then by Theorem 2.4 we can write $A = B^T B$ for some matrix $B_{m \times n}$. It is now easy to check that $A$ is the Gram matrix of the columns of $B$.

Conversely, if $A = \text{Gram}(\{x_1, \ldots, x_n\})$ with all $x_j \in \mathbb{R}^m$, then to show that $A$ is positive semidefinite, we compute for any $u = (u_1, \ldots, u_n)^T \in \mathbb{R}^n$:

$$u^T A u = \sum_{j,k=1}^n u_j u_k \langle x_j, x_k \rangle = \left\| \sum_{j=1}^n u_j x_j \right\|^2 \geq 0,$$

using the bilinearity of the inner product. \hfill \Box

2.3.2. (Toeplitz) Cosine matrices.

Definition 2.15. A matrix $A = (a_{jk})$ is Toeplitz if $a_{jk}$ depends only on $j - k$.

Lemma 2.16. Let $\theta_1, \ldots, \theta_n \in [0, 2\pi]$. Then the matrix $C := (\cos(\theta_j - \theta_k))_{j,k=1}^n$ is positive semidefinite, with rank at most 2. In particular, $\alpha 1_{n \times n} + \beta C$ has rank at most 3 (for scalars $\alpha, \beta$), and it is positive semidefinite if $\alpha, \beta \geq 0$.

This family of Toeplitz matrices was used by Rudin in a 1959 paper on entrywise positivity preservers; we will discuss his result below.

Proof. Define the vectors $u, v \in \mathbb{R}^n$ via:

$$u^T = (\cos \theta_1, \ldots, \cos \theta_n), \quad v^T = (\sin \theta_1, \ldots, \sin \theta_n).$$

Then $C = uu^T + vv^T$ via the identity $\cos(a - b) = \cos a \cos b + \sin a \sin b$, and clearly the rank of $C$ is at most 2. (For instance, it can have rank 1 if the $\theta_j$ are equal.) As a consequence,

$$\alpha 1_{n \times n} + \beta C = \alpha 1_n 1_n^T + \beta uu^T + \beta vv^T$$

has rank at most 3; the last assertion is straightforward. \hfill \Box

As a special case, suppose $\theta_1, \ldots, \theta_n$ are in arithmetic progression, i.e., there exists $\theta$ such that $\theta_{j+1} - \theta_j = \theta$ for $j = 1, \ldots, n - 1$. The above construction now yields the positive semidefinite matrix

$$C = \begin{pmatrix}
1 & \cos \theta & \cos 2\theta & \cdots \\
\cos \theta & 1 & \cos \theta & \cdots \\
\cos 2\theta & \cos \theta & 1 & \cos \theta & \cdots \\
\vdots & \ddots & \ddots & \ddots & \ddots
\end{pmatrix},$$

which is Toeplitz.

2.3.3. Hankel matrices.

Definition 2.17. A matrix $A = (a_{jk})$ is Hankel if $a_{jk}$ depends only on $j + k$.

Example 2.18. $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ is Hankel but not positive semidefinite.

Example 2.19. For each $x \geq 0$, the matrix $\begin{pmatrix} 1 & x & x^2 \\
x & x^2 & x^3 \\
x^2 & x^3 & x^4 \end{pmatrix}$ is Hankel and positive semidefinite of rank 1.
A more general perspective is as follows. Define \( H_x := \begin{pmatrix} 1 & x & x^2 & \cdots \\ x & x^2 & x^3 & \cdots \\ x^2 & x^3 & x^4 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \), and let \( \delta_x \) be the Dirac measure at \( x \in \mathbb{R} \). The moments of this measure are given by \( s_k(\delta_x) = \int_\mathbb{R} y^k \, d\delta_x(y) = x^k \), \( k \geq 0 \). Thus \( H_x \) is the moment matrix of \( \delta_x \). More generally, given any non-negative measure \( \mu \) supported on \( \mathbb{R} \), with all moments finite, the corresponding Hankel matrix is the bi-infinite ‘matrix’ given by

\[
H_\mu = \begin{pmatrix} s_0 & s_1 & s_2 & \cdots \\ s_1 & s_2 & s_3 & \cdots \\ s_2 & s_3 & s_4 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}, \quad \text{where } s_k = s_k(\mu) := \int_\mathbb{R} y^k \, d\mu(y). 
\]  

(2.20)

Now we have:

**Lemma 2.21.** The matrix \( H_\mu \) is positive semidefinite. In other words, every finite principal submatrix is positive semidefinite.

Notice that this condition is equivalent to every leading principal submatrix being positive semidefinite.

**Proof.** Fix \( n \geq 1 \) and consider the finite principal (Hankel) submatrix \( H'_\delta \) with the first \( n \) rows and columns. Let \( H'_\delta \) be the Hankel matrix defined in a similar manner for the measure \( \delta_x \), \( x \in \mathbb{R} \). Now to show that \( H'_\mu \) is positive semidefinite, we compute for any vector \( u \in \mathbb{R}^n \):

\[
u^T H'_\mu u = \int_\mathbb{R} u^T H'_\delta u \, d\mu(x) = \int_\mathbb{R} ((1, x, \ldots, x^{n-1})u)^2 \, d\mu(x) \geq 0,
\]

where the final equality holds because \( H'_\delta \) has rank one, and factorizes as in Example 2.19.

(Note that the first equality holds because we are taking finite linear combinations of the integrals in the entries of \( H'_\mu \).) \( \square \)

**Remark 2.22.** Lemma 2.21 is (the easier) half of a famous classical result by Hamburger. The harder converse result says that if a semi-infinite Hankel matrix is positive semidefinite, with \((j,k)\)-entry \( s_{j+k} \) for \( j,k \geq 0 \), then there exists a non-negative Borel measure on the real line whose \( k \)th moment is \( s_k \) for all \( k \geq 0 \). This theorem will be useful later.

There is also a multivariate version of Lemma 2.21 which is no harder, modulo notation:

**Lemma 2.23.** Given a measure \( \mu \) on \( \mathbb{R}^d \) for some integer \( d \geq 1 \), we define its moments for tuples of non-negative integers \( n = (n_1, \ldots, n_d) \) via:

\[
s_n(\mu) := \int_{\mathbb{R}^d} x^n \, d\mu(x) = \int_{\mathbb{R}^d} \prod_{j=1}^d x_j^{n_j} \, d\mu,
\]

if these integrals converge. Now suppose \( \mu \geq 0 \) on \( \mathbb{R}^d \) and let \( \Psi : (\mathbb{Z}_{\geq 0})^d \to \mathbb{Z}_{\geq 0} \) be any bijection such that \( \Psi(0) = 0 \) (although this restriction is not really required). Define the semi-infinite matrix \( H_\mu := (a_{jk})_{j,k=0}^{\infty} \) via:

\[
a_{jk} := s_{\Psi^{-1}(j)+\Psi^{-1}(k)},
\]

where we assume that all moments of \( \mu \) exist. Then \( H_\mu \) is positive semidefinite.

Proof. Given a real vector \( u = (u_0, u_1, \ldots)^T \) with finitely many nonzero coordinates, we have:
\[
u^T H \mu u = \sum_{j,k \geq 0} \int_{\mathbb{R}^d} u_j u_k x^{\Psi^{-1}(j) + \Psi^{-1}(k)} \, d\mu(x) = \int_{\mathbb{R}^d} ((1, x^{\Psi^{-1}(1)}, x^{\Psi^{-1}(2)}, \ldots)u)^2 \, d\mu(x) \geq 0.
\]
\( \Box \)

2.3.4. Matrices with sparsity. Another family of positive semidefinite matrices involves matrices with a given zero pattern, i.e. structure of (non)zero entries. Such families are important in applications, as well as in combinatorial linear algebra, spectral graph theory, and graphical models / Markov random fields.

Definition 2.24. A graph \( G = (V, E) \) is simple if the sets of vertices/nodes \( V \) and edges \( E \) are finite, and \( E \) contains no self-loops \( (v, v) \) or multi-edges. In the sequel, all graphs will be finite and simple.

Given a finite simple graph \( G = (V, E) \), with node set \( V = [n] = \{1, \ldots, n\} \), define
\[
P_G := \{ A \in \mathbb{P}_n : a_{jk} = 0 \text{ if } j \neq k \text{ and } (j, k) \notin E \}, \tag{2.25}
\]
where \( \mathbb{P}_n \) comprises the (real symmetric) positive semidefinite matrices of dimension \( n \).

Remark 2.26. The set \( P_G \) is a natural mathematical generalization of the cone \( \mathbb{P}_n \) (and shares several of its properties). In fact, two ‘extreme’ special cases are: (i) \( G \) is the complete graph, in which case \( P_G \) is the full cone \( \mathbb{P}_n \) for \( n = |V| \); and (ii) \( G \) is the empty graph, in which case \( P_G \) is the cone of \( |V| \times |V| \) diagonal matrices with non-negative entries.

Akin to both of these cases, for all graphs \( G \), the set \( P_G \) is in fact a closed convex cone. Also notice by the Schur product theorem \[^{3.3}\] that \( P_G \) is closed under taking entrywise products.

Example 2.27. Let \( G = \{\{v_1, v_2, v_3\}, \{(v_1, v_3), (v_2, v_3)\}\} \). The adjacency matrix is given by
\[
A_G = \begin{pmatrix}
0 & 0 & 1 \\
0 & 0 & 1 \\
1 & 1 & 0
\end{pmatrix}.
\]
This matrix is not in \( P_G \) (but \( A_G + \lambda_{\min}(A_G) \text{Id}_{3 \times 3} \in P_G \), see Corollary \[^{2.5}\]).

Example 2.28. For any graph with node set \([n]\), let \( D_G \) be the diagonal matrix with \( (j, j) \) entry the degree of node \( j \), i.e. the number of edges adjacent to \( j \). Then the graph Laplacian, defined to be \( L_G := D_G - A_G \) (where \( A_G \) is the adjacency matrix), is in \( P_G \).

Example 2.29. An important class of examples of positive semidefinite matrices arises from the Hessian matrix of (suitably differentiable) functions. In particular, if the Hessian is positive definite at a point, then this is an isolated local minimum.

2.4. Schur complements. We mention some more preliminary results here; these may be skipped for now, but will get used in Lemma \[^{9.4}\] below.

Definition 2.30. Given a matrix \( M = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \), where \( A \) and \( D \) are square and \( D \) non-singular, the Schur complement of \( M \) with respect to \( D \) is given by
\[
M/D := A - BD^{-1}C.
\]

Schur complements arise naturally in theory and applications. As an important example, suppose \( X_1, \ldots, X_n \) and \( Y_1, \ldots, Y_m \) are random variables with covariance matrix \( \Sigma = \begin{pmatrix} A & B \\ B^T & C \end{pmatrix} \), with \( C \) non-singular. Then the conditional covariance matrix of \( X \) given \( Y \) is \( \text{Cov}(X|Y) := A - BC^{-1}B^T = \Sigma/C \). That such a matrix is also positive semidefinite is implied by the following result.
Theorem 2.31. Given a symmetric matrix $\Sigma = \begin{pmatrix} A & B \\ B^T & C \end{pmatrix}$, with $C$ positive definite, the matrix $\Sigma$ is positive (semi)definite if and only if the Schur complement $\Sigma / C$ is thus.

Proof. We first write down a more general matrix identity: for a non-singular matrix $D$ and a square matrix $A$, one verifies:

$$
\begin{pmatrix} A & B \\ B' & D \end{pmatrix} = \begin{pmatrix} \text{Id} & BD^{-1} \\ 0 & \text{Id} \end{pmatrix} \begin{pmatrix} A - BD^{-1}B' & 0 \\ D & \text{Id} \end{pmatrix} \begin{pmatrix} \text{Id} & 0 \\ D^{-1}B' & \text{Id} \end{pmatrix}.
$$

(Note here that the identity matrices on the right-hand side are of two different sizes.) Specializing to $B' = B^T$ and $D = C$ yields: $\Sigma = X^T Y X$, where $X = \begin{pmatrix} \text{Id} & 0 \\ D^{-1}B' & \text{Id} \end{pmatrix}$ is non-singular, and $Y = \begin{pmatrix} A - BD^{-1}B^T & 0 \\ 0 & D \end{pmatrix}$ is block diagonal (and real symmetric). The result is not hard to show from here. \qed

Akin to Sylvester’s criterion, the above characterization has a variant for when $C$ is positive semidefinite; however, this is not as easy to prove. To discuss this variant, we first discuss the Moore–Penrose inverse of a general matrix.

Definition 2.33 (Moore–Penrose inverse). Given any real $m \times n$ matrix $A$, the pseudo-inverse or Moore–Penrose inverse of $A$ is an $n \times m$ matrix $A^\dagger$ satisfying:

$$
AA^\dagger A = A, \quad A^\dagger AA^\dagger = A^\dagger,
$$

and $(AA^\dagger)_{m\times m}$, $(A^\dagger A)_{n\times n}$ are symmetric.

The following fact can be found in standard texts, and is presented here without proof.

Lemma 2.34. For every $A_{m\times n}$, the matrix $A^\dagger$ exists and is unique.

Example 2.35. Here are some examples of the Moore–Penrose inverse of square matrices.

(a) If $D = \text{diag}(\lambda_1, \ldots, \lambda_r, 0, \ldots, 0)$, with all $\lambda_j \neq 0$, then $D^\dagger = \text{diag}\left(\frac{1}{\lambda_1}, \ldots, \frac{1}{\lambda_r}, 0, \ldots, 0\right)$.

(b) If $A$ is positive semidefinite, then $A = U^TDU$ where $D$ is a diagonal matrix. It is easy to verify that $A^\dagger = U^TD^\dagger U$.

(c) If $A$ is non-singular then $A^\dagger = A^{-1}$.

We now mention the connection between the positivity of a matrix and its Schur complement with respect to a singular submatrix. First note that the Schur complement is now defined in the expected way:

$$
M = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \implies M/D := A - BD^\dagger C,
$$

in the case that the (square) matrix $D$ is singular. Now the proof of the following result can be found in standard textbooks on matrix analysis.

Theorem 2.37. Given a symmetric matrix $\Sigma = \begin{pmatrix} A & B \\ B^T & C \end{pmatrix}$, with $C$ not necessarily invertible, the matrix $\Sigma$ is positive semidefinite if and only if the following conditions hold:

(a) $C$ is positive semidefinite,

(b) the Schur complement $\Sigma/C$ is positive semidefinite, and

(c) $(\text{Id} - CC^\dagger)B^T = 0$. 


3.1. The Schur product. We now make some straightforward observations about the set $\mathbb{P}_n$. The first is that $\mathbb{P}_n$ is topologically closed, convex, and closed under scaling by positive multiples (a ‘cone’):

**Lemma 3.1.** $\mathbb{P}_n$ is a closed convex cone in $\mathbb{R}^{n \times n}$.

*Proof.* All properties are easily verified using the definition of positive semidefiniteness. □

If $A$ and $B$ are positive semidefinite matrices, then we expect the product $AB$ to also be positive semidefinite. This is true if $AB$ is symmetric.

**Lemma 3.2.** For $A, B \in \mathbb{P}_n$, if $AB$ is symmetric then $AB \in \mathbb{P}_n$.

*Proof.* In fact $AB = (AB)^T = B^T A^T = BA$, whence $A$ and $B$ commute. Writing $A = U^T D_1 U$ and $B = U^T D_2 U$ as per the Spectral Theorem for commuting matrices, we have:

$$x^T (AB)x = x^T (U^T D_1 U \cdot U^T D_2 U)x = x^T U^T (D_1 D_2) U x = \| \sqrt{D_1 D_2} U x \|^2 \geq 0.$$  

Hence $AB \in \mathbb{P}_n$. □

Note however that $AB$ need not be symmetric even if $A$ and $B$ are symmetric. In this case, the matrix $AB$ certainly cannot be positive semidefinite; however, it still satisfies one of the equivalent conditions for positive semidefiniteness (shown above for symmetric matrices), namely, having a non-negative spectrum. We prove this with the help of another result, which shows the ‘tracial’ property of the spectrum:

**Lemma 3.3.** For $A_{n \times m}, B_{m \times n}$, we have $\sigma(AB) \cup \{0\} = \sigma(BA) \cup \{0\}$.

(Here, ‘tracial’ suggests that the expression for $AB$ equals that for $BA$, as does the trace.)

*Proof.* Assume without loss of generality that $1 \leq m \leq n$. The result will follow if we can show that $\det(\lambda I_{n \times n} - AB) = \lambda^{n-m} \det(\lambda I_{m \times m} - BA)$ for all $\lambda$. In turn, this follows from the equivalence of characteristic polynomials of $AB$ and $BA$ up to a power of $\lambda$, which is why we must take the union of both spectra with zero. (In particular, the sought-for equivalence would also imply that the non-zero eigenvalues of $AB$ and $BA$ are equal up to multiplicity).

The proof finishes by considering the two following block matrix identities:

$$
\begin{pmatrix}
I_{n \times n} & -A \\
0 & \lambda I_{m \times m}
\end{pmatrix} \begin{pmatrix}
\lambda I_{n \times n} & A \\
B & I_{m \times m}
\end{pmatrix} = \begin{pmatrix}
\lambda I_{n \times n} - AB & 0 \\
\lambda B & \lambda I_{m \times m}
\end{pmatrix}
$$

$$
\begin{pmatrix}
I_{n \times n} & 0 \\
-B & \lambda I_{m \times m}
\end{pmatrix} \begin{pmatrix}
\lambda I_{n \times n} & A \\
B & I_{m \times m}
\end{pmatrix} = \begin{pmatrix}
\lambda I_{n \times n} & A \\
0 & \lambda I_{m \times m} - BA
\end{pmatrix}
$$

Note that the determinants on the two left-hand sides are equal. Now, equating the determinants on the right-hand sides and cancelling $\lambda^m$ shows the desired identity:

$$
\det(\lambda I_{n \times n} - AB) = \lambda^{n-m} \det(\lambda I_{m \times m} - BA)
$$

for $\lambda \neq 0$. But since both sides here are polynomial (hence continuous) functions of $\lambda$, taking limits implies the identity for $\lambda = 0$ as well. (Alternately, $AB$ is singular if $n > m$, which shows the identity for $\lambda = 0$.) □

With Lemma 3.3 in hand, we can prove:

**Proposition 3.4.** For $A, B \in \mathbb{P}_n$, $AB$ has non-negative eigenvalues.
Proof. Let $X = \sqrt{A}$ and $Y = \sqrt{AB}$, where $A = U^T DU \implies\sqrt{A} = U^T \sqrt{D}U$. Then $XY = AB$ and $YX = \sqrt{AB} \sqrt{A}$. In fact, $YX$ is (symmetric and) positive semidefinite, since $x^T YX x = \|\sqrt{B} \sqrt{A} x\|^2$, $\forall x \in \mathbb{R}^n$.

It follows that $YX$ has non-negative eigenvalues, whence the same holds by Lemma 3.3 for $XY = AB$, even if $AB$ is not symmetric.

We next introduce a different multiplication operation on matrices (possibly rectangular, including row or column matrices), which features extensively below.

Definition 3.5. Given positive integers $m, n$, the Schur product of $A_{m\times n}$ and $B_{m\times n}$ is the matrix $C_{m\times n}$ with $c_{jk} = a_{jk} b_{jk}$ for $1 \leq j \leq m, 1 \leq k \leq n$. We denote the Schur product by $\circ$ (to distinguish it from the conventional matrix product).

Lemma 3.6. Given integers $m, n > 0$, the space $(\mathbb{R}^{m\times n}, +, \circ)$ forms a commutative associative algebra.

Proof. The easy proof is omitted. More formally, $\mathbb{R}^{m\times n}$ under coordinatewise addition and multiplication is simply the direct sum (or direct product) of copies of $\mathbb{R}$ under usual addition and multiplication.

We define another product operator on matrices of any dimensions:

Definition 3.7. Given matrices $A_{m\times n}, B_{p\times q}$, the Kronecker product of $A$ and $B$, denoted $A \otimes B$ is the $mp \times nq$ block matrix, described as:

$$A \otimes B = \begin{pmatrix}
a_{11}B & a_{12}B & \cdots & a_{1n}B \\
a_{21}B & a_{22}B & \cdots & a_{2n}B \\
\vdots & \vdots & \ddots & \vdots \\
a_{m1}B & a_{m2}B & \cdots & a_{mn}B
\end{pmatrix}$$

While the Kronecker product (as defined) is asymmetric in its arguments, it is easily seen that the analogous matrix $B \otimes A$ is obtained from $A \otimes B$ by permuting its rows and columns.

The next result, by Schur (1911), is important in later sections. We provide four proofs.

Theorem 3.8 (Schur product theorem). $\mathbb{P}_n$ is closed under $\circ$.

Proof. Suppose $A, B \in \mathbb{P}_n$; we present four proofs that $A \circ B \in \mathbb{P}_n$.

(1) Let $A, B \in \mathbb{P}_n$ have eigenbases $(\lambda_j, v_j)$ and $(\mu_k, w_k)$, respectively. Then, $(A \otimes B)(v_j \otimes w_k) = \lambda_j \mu_k (v_j \otimes w_k), \quad \forall 1 \leq j, k \leq n.$

It follows that the Kronecker product has spectrum $\{\lambda_j \mu_k\}$, and hence is positive (semi)definite if $A, B$ are positive (semi)definite. Hence every principal submatrix is also positive (semi)definite by Lemma 2.8. But now observe that the principal submatrix of $A \otimes B$ with entries $a_{jk} b_{jk}$ is precisely the Schur product $A \circ B$.

(2) By the spectral theorem and the bilinearity of the Schur product,

$$A = \sum_{j=1}^n \lambda_j v_j v_j^T, \quad B = \sum_{k=1}^n \mu_k w_k w_k^T \implies A \circ B = \sum_{j,k=1}^n \lambda_j \mu_k (v_j \circ w_k)(v_j \circ w_k)^T.$$

This is a non-negative linear combination of rank-one positive semidefinite matrices, hence lies in $\mathbb{P}_n$ by Lemma 3.1.

(3) This proof uses a clever computation. Given real matrices $A \in \mathbb{P}_n$ and $B \in \mathbb{R}^{n \times n}$, and a vector $v \in \mathbb{R}^n$, let $D_v$ denote the diagonal matrix with diagonal entries the coordinates of $v$ (in the same order), and compute:

$$v^T (A \circ B) v = \text{tr}(AD_vB^T D_v) = \text{tr}(A^{1/2}D_vB^T D_vA^{1/2}).$$

Now if $B$ is also positive semidefinite, then so is $A^{1/2}D_vB^T D_vA^{1/2}$, whence its trace is non-negative for each $v \in \mathbb{R}^n$.

(4) Given $t > 0$, let $X, Y$ be independent multivariate normal vectors centered at 0 and with covariance matrices $A + t \text{Id}_{n \times n}, B + t \text{Id}_{n \times n}$ respectively. The Schur product of $X$ and $Y$ is then a random vector with mean zero, and covariance matrix precisely $(A + t \text{Id}_{n \times n}) \circ (B + t \text{Id}_{n \times n})$. Now the result follows directly from the fact that covariance matrices are positive semidefinite, upon taking the limit as $t \to 0^+$.

For completeness, we mention another property of positive (semi)definite matrices – the Cholesky decomposition – although this is not used later.

**Theorem 3.9** (Cholesky decomposition). Given $A \in \mathbb{P}_n$, there exists a lower triangular matrix $L_{n \times n}$ such that $A = LL^T$ and $L$ has non-negative diagonal entries. If moreover $A$ is positive definite, then $L$ can be uniquely chosen to have all positive diagonal elements.

### 3.2. Totally Positive (TP) and Totally Non-negative (TN) matrices.

**Definition 3.10.** A matrix $A_{m \times n}$ is said to be totally positive (TP) (respectively totally non-negative (TN)) if all minors of $A$ are positive (respectively non-negative).

Here are some distinctions between TP/TN matrices and positive (semi)definite ones:

- For TP/TN matrices we consider all minors, not just the principal ones.
- As a consequence of considering the $1 \times 1$ minors, it follows that the entries of TP (TN) matrices are all positive (non-negative).
- TP/TN matrices are not required to be symmetric, unlike positive semidefinite matrices.

**Example 3.11.** The matrix $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ is totally non-negative but not totally positive, while the matrix $\begin{pmatrix} 1 & 2 \\ 3 & 16 \end{pmatrix}$ is totally positive.

Totally positive matrices have featured in the mathematics literature in a variety of classical and modern topics. We mention a few of these topics, as well as some of the experts who have worked/written on them.

- Interacting particle systems (Gantmakher, Krein, Deift).
- Analysis (Schoenberg, Loewner, Pólya).
- Probability (Karlin).
- Matrix theory and applications (Ando, Fallat, Johnson, Pinkus, Sokal).
- Combinatorics (Brenti).
- Representation theory and canonical bases (Lusztig).
- Cluster algebras (Berenstein, Fomin, Zelevinsky).
- Integrable systems (Kodama, Williams).

In the remainder of this section and the next, we discuss several other examples of TN matrices. Specifically, we show that the (positive semidefinite) Toeplitz cosine matrices and Hankel moment matrices considered above, are in fact totally non-negative!
Example 3.12 (Toeplitz cosine matrices). We claim that the matrices
\[ C(\theta) := (\cos(j-k)\theta)_{j,k=1}^n, \quad \text{where } \theta \in [0, \frac{\pi}{2(n-1)}] \]
are totally non-negative.

Indeed, all 1 \times 1 minors are non-negative, and as discussed above, \( C(\theta) \) has rank at most 2, and so all 3 \times 3 and larger minors vanish. It remains to consider all 2 \times 2 minors. Now a 2 \times 2 sub-matrix of \( C(\theta) \) is of the form
\[ C' = \begin{pmatrix} C_{ab} & C_{ac} \\ C_{db} & C_{dc} \end{pmatrix}, \]
where \( 1 \leq a < d \leq n, \) and \( 1 \leq b < c \leq n, \) and \( C_{ab} \) denotes the matrix entry \( \cos((a-b)\theta). \) Writing \( a, b, c, d \) in place of \( a\theta, b\theta, c\theta, d\theta \) in the next computation ease of exposition, the corresponding minor is:
\[
\det C' = \frac{1}{2} \{ 2 \cos(a-b) \cos(d-c) - 2 \cos(a-c) \cos(d-b) \} \\
= \frac{1}{2} \{ \cos(a-b+d-c) + \cos(a-b-d+c) - \cos(a-c+d-b) - \cos(a-c-d+b) \} \\
= \frac{1}{2} \{ \cos(a-b-d+c) - \cos(a-c-d+b) \} \\
= \frac{1}{2} (-2) \sin(a-d) \sin(c-b) \\
= - \sin(a-d) \sin(c-b). \tag{3.13}
\]
Thus, \( \det C' = \sin(d\theta - a\theta) \sin(c\theta - b\theta), \) which is non-negative because \( a < d, b < c, \) and \( \theta \in [0, \frac{\pi}{2(n-1)}]. \) This shows that \( C(\theta) \) is totally non-negative.
4. Fekete’s result. Hankel moment matrices are TN. Hankel psd and TN matrices.

Continuing from the previous section, we show that the Hankel moment matrices $H_\mu$ studied above are not only positive semidefinite, but more strongly, totally non-negative (TN). Akin to the Toeplitz cosine matrices (where the angle $\theta$ is restricted to ensure the entries are non-negative), we restrict the support of the measure to $[0, \infty)$, which implies that the entries are non-negative.

To achieve these goals, we prove the following result, which is crucial in relating preservers of positive semidefinite matrices to those of Hankel TN matrices, later in these notes.

**Theorem 4.1.** If $A_{n \times n}$ is a real Hankel matrix, then $A$ is TN (TP) if and only if both $A$ and $A^{(1)}$ are positive semidefinite (positive definite), where $A^{(1)}$ is obtained from $A$ by removing the first row and last column.

From this theorem, we derive the following two consequences, both of which are useful in later sections.

**Corollary 4.2.** The set of Hankel TN matrices is a closed convex cone, further closed under taking Schur products.

The second corollary of Theorem 4.1 provides a large class of examples of such Hankel TN matrices:

**Corollary 4.3.** Suppose $\mu$ is a non-negative measure supported in $[0, \infty)$, with all moments finite. Then $H_\mu$ is TN.

The proofs are left as easy exercises.

**Remark 4.4.** Akin to Lemma 2.21 and the remark following its proof, Corollary 4.3 is also the easy half of a well-known classical result on moment problems – this time, by Stieltjes. The harder converse result says (in particular) that if a semi-infinite Hankel matrix $H$ is TN, with $(j, k)$-entry $s_{j+k} \geq 0$ for $j, k \geq 0$, then there exists a non-negative Borel measure on $\mathbb{R}$ with support in $[0, \infty)$, whose $k$th moment is $s_k$ for all $k \geq 0$. By Theorem 4.1 this is equivalent to both $H$ as well as $H^{(1)}$ being positive semidefinite, where $H^{(1)}$ is obtained by truncating the first row (or the first column) of $H$.

The remainder of this section is devoted to showing Theorem 4.1. The proof uses a sequence of lemmas shown by Gantmakher, Krein, Fekete, and others. The first of these lemmas may be (morally) attributed to Laplace.

**Lemma 4.5.** Let $r \geq 1$ be an integer and $U = (u_{jk})$ an $(r+2) \times (r+1)$ matrix. Given subsets $[a, b], [c, d] \subset (0, \infty)$, let $U_{[a,b] \times [c,d]}$ denote the submatrix of $U$ with entries $u_{jk}$ such that $j, k$ are integers and $a \leq j \leq b$, $c \leq k \leq d$. Then the following identity holds:

$$
\begin{align*}
\det U_{[1,r] \cup [r+2] \times [1,r+1]} \cdot \det U_{[2,r+1] \times [1,r]} & = \det U_{[2,r+2] \times [1,r+1]} \cdot \det U_{[1,r] \times [1,r]} \\
& + \det U_{[1,r+1] \times [1,r+1]} \cdot \det U_{[2,r] \cup [r+2] \times [1,r]}. 
\end{align*}
$$

Note that in each of the three products of determinants, the second factor in the subscript for the first (respectively second) determinant terms is the same: $[1, r+1]$ (respectively $[1, r]$).

To give a feel for the result, the special case of $r = 1$ asserts that

$$
\begin{vmatrix}
u_{11} & \nu_{21} & \nu_{32} \\
u_{12} & \nu_{22} & \nu_{33} \end{vmatrix} - \nu_{21} \begin{vmatrix} \nu_{11} & \nu_{12} \\
u_{31} & \nu_{32} \end{vmatrix} + \nu_{31} \begin{vmatrix} \nu_{11} & \nu_{12} \\
u_{21} & \nu_{22} \end{vmatrix} = 0.
$$
But this is precisely the Laplace expansion along the third column of the singular matrix
\[
\det \begin{pmatrix} u_{11} & u_{12} & u_{11} \\ u_{21} & u_{22} & u_{21} \\ u_{31} & u_{32} & u_{31} \end{pmatrix} = 0.
\]

**Proof.** Consider the \((2r+1) \times (2r+1)\) block matrix of the form
\[
M = \begin{pmatrix}
    b^T & u_{1,r+1} & b^T \\
    A & a & 0_{r-1} \\
    c^T & u_{r+1,r+1} & c^T \\
    d^T & u_{r+2,r+1} & d^T \\
    A & a & 0_{(r-1)\times r}
\end{pmatrix} = \begin{pmatrix}
    (u_{jk})_{j\in[1,r+2],k\in[1,r+1]} \\
    (u_{jk})_{j\in[2,r],k\in[1,r+1]} \\
    0_{(r-1)\times r}
\end{pmatrix};
\]
that is, where
\[
a = (u_{2,r+1}, \ldots, u_{r,r+1})^T, \quad b = (u_{1,1}, \ldots, u_{1,r})^T, \quad c = (u_{r+1,1}, \ldots, u_{r+1,r})^T,
\]
\[
d = (u_{r+2,1}, \ldots, u_{r+2,r})^T, \quad A = (u_{jk})_{2\leq j\leq r, 1\leq k\leq r}.
\]

Notice that \(M\) is a square matrix whose first \(r+2\) rows have column space of dimension at most \(r+1\); hence \(\det M = 0\). Now we compute \(\det M\) using the (generalized) Laplace expansion by complementary minors: choose all possible \((r+1)\)-tuples of rows from the first \(r+1\) columns to obtain a submatrix \(M'_{(r+1)}\), and deleting these rows and columns from \(M\) yields the complementary \(r \times r\) submatrix \(M''_{(r)}\) from the final \(r\) columns. The generalized Laplace expansion says that if one multiplies \(\det M'_{(r+1)} \cdot \det M''_{(r)}\) by \((-1)^{\Sigma}\), with \(\Sigma\) the sum of the row numbers in \(M'_{(r+1)}\), then summing over all such products (running over subsets of rows) yields \(\det M\) – which vanishes for this particular matrix \(M\).

Now in the given matrix, to avoid obtaining zero terms, the rows in \(M'_{(r+1)}\) must include all entries from the final \(r-1\) rows (and the first \(r+1\) columns). But then it moreover cannot include entries from the rows of \(M\) labelled 2, \ldots, \(r\); and it must include two of the remaining three rows (and entries from only the first \(r+1\) columns).

Thus, we obtain three product terms that must sum to: \(\det M = 0\). Upon carefully examining the terms and computing the companion signs (by row permutations), we obtain \([4,6]\). \(\square\)

The next sequence of arguments, due to Gantmakher and Krein, aims to prove a classical result on totally positive matrices, first shown by Fekete (1912).

**Lemma 4.7.** Given integers \(m \geq n \geq 1\) and a real matrix \(A_{m \times n}\) such that
(a) all \((n-1) \times (n-1)\) minors \(\det A_{J \times [1,n-1]} > 0\) for \(J \subset [1,m]\) of size \(n-1\), and
(b) all \(n \times n\) minors \(\det A_{[j+1,j+n] \times [1,n]} > 0\) for \(0 \leq j \leq m-n\),
we have that all \(n \times n\) minors of \(A\) are positive.

**Proof.** Define the gap, or ‘index’ of a subset of integers \(J = \{j_1 < j_2 < \cdots < j_n\} \subset [1,m]\), to be \(g_J := j_n - j_1 - (n-1)\). Thus, the gap is zero if and only if \(J\) consists of successive integers, and in general it counts precisely the number of integers between \(j_1\) and \(j_n\) that are missing from \(J\).

We will prove that \(\det A_{J \times [1,n]} > 0\) for \(|J| = n\), by induction on the gap \(g_J \geq 0\); note that the base case \(g_J = 0\) is given as hypothesis. For the induction step, suppose \(j^0\) is a missing
4. Fekete’s result. Hankel moment matrices are TN. Hankel psd and TN matrices.

index (or row number) in \( J = \{j_1 < \cdots < j_n\} \). By suitably specializing the identity \((4.6)\), we obtain:

\[
\begin{align*}
\det A_{\{j_1, \ldots, j_n\} \times [1, n]} &= \det A_{\{j_1, \ldots, j_{n-1}, j^0\} \times [1, n-1]} \\
&= \det A_{\{j_1, \ldots, j_{n-1}, j^0\} \times [1, n]} - \det A_{\{j_2, \ldots, j_{n-1}, j^0\} \times [1, n]} - \det A_{\{j_1, \ldots, j_{n-1}\} \times [1, n-1]}.
\end{align*}
\]

(4.8)

Consider the six factors in serial order. The first, fourth, and sixth factors have indices listed in increasing order, while the other three factors have \( j^0 \) listed at the end, so their indices are not listed in increasing order. For each of the six factors, the number of ‘bubble sorts’ required to rearrange indices in increasing order (by moving \( j^0 \) down the list) equals the number of row exchanges in the corresponding determinants; label these numbers \( n_1, \ldots, n_6 \). Thus \( n_1 = n_4 = n_6 = 0 \) as above, while \( n_2 = n_3 \) (since \( j_1 < j^0 < j_n \)), and \( |n_2 - n_5| = 1 \). Now multiply the equation \((4.8)\) by \((-1)^{n_2}\), and divide both sides by

\[
c_0 := (-1)^{n_2} \det A_{\{j_2, \ldots, j_{n-1}, j^0\} \times [1, n-1]} > 0.
\]

Using the given hypotheses as well as the induction step (since all terms involving \( j^0 \) have a gap equal to \( g_J - 1 \)), it follows that

\[
\begin{align*}
\det A_{\{j_1, \ldots, j_n\} \times [1, n]} &= c_0^{-1} \left( (-1)^{n_2} \det A_{\{j_1, \ldots, j_{n-1}, j^0\} \times [1, n]} - \det A_{\{j_2, \ldots, j_{n-1}\} \times [1, n-1]} \\
&\quad + (-1)^{n_2+1} \det A_{\{j_2, \ldots, j_{n}, j^0\} \times [1, n]} - \det A_{\{j_1, \ldots, j_{n-1}\} \times [1, n-1]} \right) \\
&> 0.
\end{align*}
\]

This completes the induction step, and with it, the proof.

We can now state and prove (a slight extension of) Fekete’s 1912 result. First we introduce the following notation, which is also useful in later sections.

**Definition 4.9.** Given an integer \( r > 0 \), we say a matrix is **totally positive** (totally non-negative) of order \( r \), denoted \( TP_r \) (\( TN_r \)), if all its \( 1 \times 1, 2 \times 2, \ldots, r \times r \) minors are positive (non-negative). We will also abuse notation and write \( A \in TP_r \) (\( A \in TN_r \)) if \( A \equiv TP_r \) (\( A \equiv TN_r \)).

**Proposition 4.10.** Suppose \( m, n \geq r - 1 \) are integers, and \( A_{m \times n} \) is a real matrix such that all ‘contiguous’ minors of \( A \) of orders at most \( r \) are positive (namely, minors of square submatrices corresponding to successive rows and to successive columns). Then \( A \) is \( TP_r \).

**Proof.** We show that for any integer \( s \in [1, r] \), every \( s \times s \) minor of \( A \) is positive. The proof is by induction on \( s \) (and running over all real matrices satisfying the hypotheses); note that the base case of \( s = 1 \) is immediate from the assumptions. For the induction step, suppose

\[
2 \leq s = |J| = |K| \leq r, \quad J \subset \mathbb{Z} \cap [1, m], \quad K \subset \mathbb{Z} \cap [1, n].
\]

First fix a subset \( K \) that consists of consecutive rows, i.e. has gap \( g_K = 0 \) (as in the proof of Lemma \([4.7]\)). Let \( B \) denote the submatrix \( A_{[1, m] \times K} \). Then all \( s \times s \) minors of \( B \) are positive, by Lemma \([4.7]\). In particular, it follows for all \( J \) that all \( s \times s \) minors \( \det A_{J \times J'} \) are positive, whenever \( K' \subset [1, n] \) has size \( s \) and gap \( g_{K'} = 0 \). Now apply Lemma \([4.7]\) to the matrix \( B := (A_{J \times [1, n]})^T \) to obtain: \( \det(A_{J,K})^T > 0 \) for (possibly non-consecutive subsets) \( K \). This concludes the proof.

The final ingredient required to prove Theorem \([4.1]\) is the following observation.

Lemma 4.11. If $A_{n \times n}$ is a Hankel matrix, then every contiguous minor of $A$ (see Proposition 4.10) is a principal minor of $A$ or of $A^{(1)}$.

Recall that $A^{(1)}$ was defined in Theorem 4.1.

Proof. Let the first row (respectively, last column) of $A$ contain the entries $s_0, s_1, \ldots, s_{n-1}$ (respectively, $s_{n-1}, s_n, \ldots, s_{2n-2}$). Then every contiguous minor is the determinant of a submatrix of the form

$$M = \begin{pmatrix} s_j & \cdots & s_{j+m} \\ \vdots & \ddots & \vdots \\ s_{j+m} & \cdots & s_{j+2m} \end{pmatrix}, \quad 0 \leq j \leq j + m \leq n - 1.$$ 

It is now immediate that if $j$ is even then $M$ is a principal submatrix of $A$, while if $j$ is odd then $M$ is a principal submatrix of $A^{(1)}$. \qed

With these results in hand, we conclude by proving the above theorem.

Proof of Theorem 4.1. If the Hankel matrix $A$ is TN (TP) then $A, A^{(1)}$ are positive semidefinite (positive definite) by definition. To prove the converse, first suppose $A, A^{(1)}$ are positive definite. By Lemma 4.11, this implies every contiguous minor of $A$ is positive. By Fekete’s result (Proposition 4.10), $A$ is TP as desired.

Finally, suppose $A, A^{(1)}$ are positive semidefinite. It follows by Lemma 4.11 that every contiguous submatrix of $A$ is positive semidefinite. Also choose and fix a $n \times n$ Hankel TP matrix $B$ (note by (5.9) or Lemma 6.9 below that such matrices exist for all $n \geq 1$). Applying Lemma 4.11, $B, B^{(1)}$ are positive definite, whence so is every contiguous submatrix of $B$.

Now for $\epsilon > 0$, it follows (by Sylvester’s criterion, Theorem 2.7) that every contiguous minor of $A + \epsilon B$ is positive. Again applying Fekete’s Proposition 4.10, the Hankel matrix $A + \epsilon B$ is necessarily TP, and taking $\epsilon \to 0^+$ finishes the proof. \qed

The previous sections discussed examples (Toeplitz, Hankel) of totally non-negative (TN) matrices. These example consisted of symmetric matrices.

- We now see some examples of non-symmetric matrices that are totally positive (TP).
- Let \( H_\mu := (s_{j+k}(\mu))_{j,k\geq 0} \) denote the moment matrix associated to a non-negative measure \( \mu \) supported on \([0, \infty)\). We have already seen that this matrix is Hankel and positive semidefinite – in fact, TN. We will show (in this section and the next) that \( H_\mu \) is in fact TP in ‘many’ cases. The proof will use a ‘continuous’ generalization of the Cauchy–Binet formula.

5.1. Generalized Vandermonde Matrices. A generalized Vandermonde matrix (GVM) is a matrix \((x_1^{\alpha_k})_{j,k=1}^n\), where \(x_j > 0\) and \(\alpha_k \in \mathbb{R}\) for all \(j\). If the \(x_j\) are pairwise distinct, as are the \(\alpha_k\), then the corresponding GVM is non-singular. In fact, a stronger result holds:

**Theorem 5.1.** If \(0 < x_1 < \cdots < x_m\) and \(\alpha_1 < \cdots < \alpha_n\) are real numbers, then the generalized Vandermonde matrix \(V_{m \times n} := (x_j^{\alpha_k})\) is totally positive.

As an illustration, consider the special case \(m = n\) and \(\alpha_k = k - 1\), which recovers the usual Vandermonde matrix:

\[
V = \begin{pmatrix}
1 & x_1 & \cdots & x_1^{n-1} \\
1 & x_2 & \cdots & x_2^{n-1} \\
\vdots & \vdots & \ddots & \vdots \\
1 & x_n & \cdots & x_n^{n-1}
\end{pmatrix}.
\]

This matrix has determinant \(\prod_{1 \leq j < k \leq n} (x_k - x_j) > 0\). Thus if \(0 < x_1 < x_2 < \cdots < x_n\) then \(\det V > 0\). However, note this is not enough to prove that the matrix is totally positive. Thus, more work is required to prove total positivity, even for usual Vandermonde matrices.

**Lemma 5.2** (Descartes’ (Laguerre’s) rule of signs, weaker version). Fix \(n\) distinct real numbers \(\alpha_1, \alpha_2, \ldots, \alpha_n \in \mathbb{R}\) and \(n\) scalars \(c_1, c_2, \ldots, c_n \in \mathbb{R}\) such that not all scalars are 0. Then the function \(f(x) := \sum_{k=1}^n c_k x^{\alpha_k}\) can have at most \((n - 1)\) distinct positive roots.

**Proof.** By induction on \(n\). For \(n = 1\), clearly the function has no positive root.

For the induction step, without loss of generality we may assume \(\alpha_1 < \alpha_2 < \cdots < \alpha_n\) and that all \(c_j\) are nonzero. If \(f\) has \(n\) distinct positive roots, then so does the function

\[
g(x) := x^{-\alpha_1} f(x) = c_1 + \sum_{k=2}^n c_k x^{\alpha_k - \alpha_1}.
\]

But then Rolle’s theorem implies that \(g'(x) = \sum_{k=2}^n c_k (\alpha_k - \alpha_1) x^{\alpha_k - \alpha_1 - 1}\) has \((n - 1)\) distinct positive roots. This contradicts the induction hypothesis, completing the proof. \(\Box\)

With this result in hand, we can prove that generalized Vandermonde matrices are TP.

**Proof of Theorem 5.1.** As any minor of \(V\) is also a generalized Vandermonde matrix, it suffices to show that the determinant of \(V\) is positive when \(m = n\).
We first claim that \( \det V \neq 0 \). Indeed, suppose for contradiction that \( V \) is singular. Then there is a vector \( c = (c_1, c_2, \ldots, c_n)^T \) such that \( Vc = 0 \). If it is so, then we have \( n \) distinct positive numbers \( x_1, x_2, \ldots, x_n \) such that \( \sum_{k=1}^{n} c_k x_1^{\alpha_k} = 0 \). This contradicts Lemma 5.2 for \( f(x) = \sum_{k=1}^{n} c_k x^{\alpha_k} \). Thus, the claim follows.

We can now prove the theorem. Consider a (continuous) path \( \gamma : [0, 1] \to \mathbb{R}^n \) going from \( \gamma(0) = (0, 1, \ldots, n-l) \) to \( \gamma(1) = (\alpha_1, \alpha_2, \ldots, \alpha_n) \), such that at each timepoint \( t \in [0, 1] \), the coordinates of \( \gamma(t) \) are in increasing order. It is possible to choose such a path; indeed, the straight line ‘geodesic’ path is one such.

Now let \( W(t) := (x_j^{\gamma(t)})^{n}_{j,k=1} \). Then \( W : [0, 1] \to \mathbb{R} \) is a nonzero, continuous map. Since \( [0, 1] \) is connected and \( W(0) > 0 \) (see remarks above), it follows that \( W(1) = \det V > 0 \). □

**Remark 5.3.** If we have \( 0 < x_n < x_{n-1} < \cdots < x_1 \) and \( \alpha_n < \alpha_{n-1} < \cdots < \alpha_1 \), then observe that the corresponding generalized Vandermonde matrix \( V' := (x_j^{\alpha_k})^{n}_{j,k=} \) is also TP.

Indeed, once again we only need to show \( \det V' > 0 \), and this follows from applying the same permutation to the rows and to the columns of \( V' \) to reduce it to the situation in Theorem 5.1 (since then the determinant does not change in sign).

### 5.2. The Cauchy–Binet formula

The following is a recipe to construct ‘new’ examples of TP/TN matrices from known ones.

**Proposition 5.4.** If \( A_{m \times n}, B_{n \times k} \) are both TN, then so is the matrix \( (AB)_{m \times k} \). This assertion is also valid upon replacing ‘TN’ by ‘TP’, provided \( n \geq \max\{m, k\} \).

To prove this proposition, we require the following important result.

**Theorem 5.5** (Cauchy–Binet formula). Given matrices \( A_{m \times n} \) and \( B_{n \times m} \), we have

\[
\det(AB)_{m \times m} = \sum_{J \subseteq [n]} \det(A_{[m] \times J^\top}) \det(B_{J^\top \times [m]}),
\]

where \( J^\top \) reiterates the fact that the elements of \( J \) are arranged in increasing order.

For example, if \( m = n \), this theorem just reiterates the fact that the determinant map is multiplicative on square matrices. If \( m > n \), the theorem says that determinants of singular matrices are zero. If \( m = 1 \), we obtain the inner product of a row and column vector.

**Proof.** Notice that

\[
\det(AB) = \det \begin{pmatrix}
\sum_{j_1=1}^{n} a_{1j_1} b_{j_11} & \cdots & \sum_{j_m=1}^{n} a_{1j_m} b_{j_m1} \\
\vdots & \ddots & \vdots \\
\sum_{j_1=1}^{n} a_{mj_1} b_{j_11} & \cdots & \sum_{j_m=1}^{n} a_{mj_m} b_{j_m1}
\end{pmatrix}.
\]

By the multilinearity of the determinant, expanding \( \det(AB) \) as a sum over all \( j_1 \) yields:

\[
\det(AB) = \sum_{(j_1, j_2, \ldots, j_m) \in [n]^m} b_{j_11} b_{j_22} \cdots b_{j_m m} \cdot \det \begin{pmatrix}
\vdots & \cdots & \vdots \\
a_{1j_1} & \cdots & a_{1j_m} \\
a_{mj_1} & \cdots & a_{mj_m}
\end{pmatrix}.
\]

The determinant in the summand vanishes if \( j_k = j_m \) for any \( k \neq m \). Thus the above expression reduces to

\[
\det(AB) = \sum_{\{j_1, \ldots, j_m\} \subseteq [n]^m, \ all \ j_i \ are \ distinct} b_{j_1} b_{j_2} \cdots b_{j_m} m \cdot \det \begin{pmatrix}
    a_{1j_1} & \cdots & a_{1j_m} \\
    \vdots & \ddots & \vdots \\
    a_{mj_1} & \cdots & a_{mj_m}
\end{pmatrix}
\]

\[= \sum_{\{j_1, \ldots, j_m\} \subseteq [n]^m, \ all \ j_i \ are \ distinct} b_{j_1} b_{j_2} \cdots b_{j_m} \det A_{[m] \times (j_1, j_2, \ldots, j_m)}.
\]

We split this sum into two sub-summations. One part runs over all collections of indices, while the other runs over all possible orderings – that is, permutations – of each fixed collection of indices. Thus for each ordering \( j = (j_1, \ldots, j_m) \) of \( J = \{j_1, \ldots, j_m\} \), there exists a unique permutation \( \sigma_j \in S_m \) such that \( (j_1, \ldots, j_m) = \sigma_j(J^\uparrow) \). Now using the definition of the determinant:

\[
\det(AB) = \sum_{J = \{j_1, j_2, \ldots, j_m\} \subseteq [n], \ all \ j_i \ are \ distinct} \sum_{\sigma = \sigma_j \in S_m} b_{j_1} b_{j_2} \cdots b_{j_m} (-1)^{\sigma_j} \det A_{[m] \times J^\uparrow}
\]

\[= \sum_{\{J \subseteq [n], \ \text{of size} \ m\}} \det(A_{[m] \times J^\uparrow}) \det(B_{J^\uparrow \times [m]}).
\]

**Proof of Proposition 5.4.** Suppose two matrices \( A_{m \times n} \) and \( B_{n \times k} \) are both TN. Let \( I \subseteq [m] \) and \( K \subseteq [k] \) be index subsets of the same size; we are to show \( \det(AB)_{IJ \times K} \) is non-negative. Define matrices, \( A' := A_{I \times [n]} \) and \( B' := B_{[n] \times K} \). Now it is easy to show that \( (AB)_{IJ \times K} = A'B' \). In particular, \( \det(AB)_{IJ \times K} = \det(A'B') \). Hence, the Cauchy–Binet theorem implies:

\[
\det(AB)_{IJ \times K} = \sum_{\{J \subseteq [n], \ |J| = |K| = |I|\}} \det(A'_{I \times J^\uparrow}) \det(B'_{J^\uparrow \times K}) = \sum_{\{J \subseteq [n], \ |J| = |K| = |I|\}} \det(A_{I \times J^\uparrow}) \det(B_{J^\uparrow \times K} \geq 0.
\]

It follows that \( AB \) is TN if both \( A \) and \( B \) are TN. For the corresponding TP-version, the above proof works as long as the sums in the preceding equation are always over non-empty sets; but this happens whenever \( n \geq \max\{m, k\} \).

**Remark 5.8.** Now Proposition 5.4 shows that if \( A, B \) are rectangular matrices that are TP/TN of order \( r \), then so is \( AB \).

5.3. **Generalized Cauchy–Binet formula.** We showed in the previous section that generalized Vandermonde matrices are examples of totally positive but non-symmetric matrices. Using these, we can construct additional examples of totally positive symmetric matrices: let \( V = (x_j^{\alpha_k})_{j,k=1}^n \) be a generalized Vandermonde matrix with \( 0 < x_1 < x_2 < \cdots < x_n \) and \( \alpha_1 < \alpha_2 < \cdots < \alpha_n \). Then Proposition 5.4 implies that the symmetric matrices \( V^TV \) and \( VV^T \) are totally positive.

For instance, if we take \( n = 3 \) and \( \alpha_k = k - 1 \), then

\[
V = \begin{pmatrix}
1 & x_1 & x_1^2 \\
1 & x_2 & x_2^2 \\
1 & x_3 & x_3^2
\end{pmatrix}, \quad V^TV = \begin{pmatrix}
\sum_{j=1}^3 x_{1j} & \sum_{j=1}^3 x_{2j} & \sum_{j=1}^3 x_{3j} \\
\sum_{j=1}^3 x_{1j}^2 & \sum_{j=1}^3 x_{2j}^2 & \sum_{j=1}^3 x_{3j}^2 \\
\sum_{j=1}^3 x_{1j} x_{2j} & \sum_{j=1}^3 x_{2j} x_{3j} & \sum_{j=1}^3 x_{3j} x_{1j}
\end{pmatrix}
\]

(5.9)

This is clearly the Hankel moment matrix \( H_\mu \) for the counting measure on the set \( \{x_1, x_2, x_3\} \). Moreover, \( V^TV \) is (symmetric and) totally positive by the Cauchy–Binet theorem. More
generally, for all increasing \( \alpha_k \) which are in arithmetic progression, the matrix \( V^T V \) (defined similarly as above) is a totally positive Hankel moment matrix for some non-negative measure on \([0, \infty)\) – more precisely, supported on \( \{x_1^{\alpha_1-\alpha}, x_2^{\alpha_2-\alpha}, \ldots, x_n^{\alpha_n-\alpha} \} \).

The following discussion aims to show (among other things) that the moment matrices \( H_\mu \) defined in (2.20) are totally positive for 'most' non-negative measures \( \mu \) supported in \([0, \infty)\). We begin with by studying functions that are TP or TN.

**Definition 5.10.** Let \( X,Y \subseteq \mathbb{R} \) and \( K : X \times Y \to \mathbb{R} \) be a function. Given \( r \in \mathbb{N} \), we say \( K(x,y) \) is a totally non-negative/totally positive kernel of order \( r \) if for any integer \( 1 \leq n \leq r \) and elements \( x_1 < x_2 < \cdots < x_n \in X \) and \( y_1 < y_2 < \cdots < y_n \in Y \), we have \( \det K(x_j, y_k)_{j,k=1}^n \) is non-negative (positive). We denote this by writing \( K \) is \( TNR \) (or \( TP_r \)). Similarly, we say that the kernel \( K : X \times Y \to \mathbb{R} \) is totally non-negative/totally positive if \( K \) is \( TNR \) (or \( TP_r \)) for all \( r \geq 1 \).

**Example 5.11.** We claim that the kernel \( K(x,y) = e^{xy} \) is totally positive, with \( X = Y = \mathbb{R} \).

Indeed, choose any real numbers \( x_1 < x_2 < \cdots < x_n \) and any \( y_1 < y_2 < \cdots < y_n \). Then the matrix \( (e^{x_i y_j})_{i,j=1}^n \) is a generalized Vandermonde matrix that is TP, and hence its determinant is strictly positive.

We next generalize the Cauchy–Binet theorem to TP/TN kernels. Let \( X,Y,Z \subseteq \mathbb{R} \) and \( \mu \) be a non-negative Borel measure on \( Y \). Let \( K(x,y) \) and \( L(y,z) \) be 'nice' functions (i.e., Borel measurable with respect to \( Y \)), and assume the following function is well-defined:

\[
M : X \times Z \to \mathbb{R}, \quad M(x,z) := \int_Y K(x,y) L(y,z) d\mu(y).
\]

For example, consider \( K(x,y) = e^{xy} \) and \( L(y,z) = e^{yz} \). Take \( X = Z = \{\alpha_1, \alpha_2, \ldots, \alpha_n \} \) and \( Y = \{\log(x_1), \log(x_2), \ldots, \log(x_n)\} \) such that \( 0 < x_1 < x_2 < \cdots < x_n \) and \( \alpha_1 < \alpha_2 < \cdots < \alpha_n \). Finally, \( \mu \) denote the counting measure on \( Y \). Then \( M(\alpha_j, \alpha_k) = \sum_{i=1}^n x_j^{\alpha_i} x_j^{\alpha_i} \). So \( (M(\alpha_i, \alpha_k))_{i,k=1}^n = VTV \), where \( V = (x_j^{\alpha_k})_{j,k=1}^n \) is a generalized Vandermonde matrix.

In this 'discrete' example (i.e., where the support of \( \mu \) is a discrete set), \( \det M \) is shown to be positive using the total positivity of \( V, V^T \) and the Cauchy–Binet formula. Following Karlin, we note that this phenomenon extends to the above more general setting:

**Exercise 5.13** (Pólya–Szegő, Basic Composition Formula, or Generalized Cauchy–Binet formula). Suppose \( X,Y,Z \subseteq \mathbb{R} \) and \( K(x,y), L(y,z), M(x,z) \) are as above. Then using an argument similar to the above proof of the Cauchy–Binet formula, show that

\[
\det \begin{pmatrix}
M(x_1, z_1) & \cdots & M(x_1, z_m) \\
\vdots & \ddots & \vdots \\
M(x_m, z_1) & \cdots & M(x_m, z_m)
\end{pmatrix}
= \int_{y_1 < y_2 < \cdots < y_m} \cdots \int_{y_1 < y_2 < \cdots < y_m} \prod_{j=1}^m d\mu(y_j). \tag{5.14}
\]

**Remark 5.15.** In the right-hand side, we may also integrate over the region \( y_1 \leq \cdots \leq y_m \) in \( Y \), since matrices with equal rows or columns are singular.
6. Hankel moment matrices are TP. Andréief’s identity.
Density of TP matrices in TN matrices.

6. **Hankel moment matrices are TP. Andréief’s identity.** Density of TP in TN.

### 6.1. Total positivity of \(H_\mu\) for ‘most’ measures; Andréief’s identity.

Continuing from the previous section with the generalized Cauchy–Binet formula of Pólya–Szegő, from (5.14) we obtain the following consequence.

**Corollary 6.1.** *(Notation as in the previous section.)* If the kernels \(K\) and \(L\) are both TN, (TP, for some integer \(r > 0\), then so is \(M\), where \(M\) was defined in (5.12). In particular, if \(K\) and \(L\) are TP/TN kernels, then so is \(M\).

We will apply this result to the moment matrices \(H_\mu\) defined in (2.20). We begin more generally: suppose \(u : Y \to \mathbb{R}\) is a positive and strictly increasing function, all of whose moments exist:

\[
\int_Y u(y)^n \, d\mu(y) < \infty, \quad \forall n \geq 0.
\]

Then we claim:

**Proposition 6.2.** The kernel \(M : Z^\geq 0 \times Z^\geq 0 \to \mathbb{R}\), given by:

\[
M(n, m) := \int_Y u(y)^{n+m} \, d\mu(y)
\]

is TN as well as TP on \(Y_+\), where \(Y_+ := \text{supp}(\mu) \subseteq Y\).

**Proof.** To show \(M\) is TP on \(Y_+\), the first claim is that the kernel \(K(n, y) := u(y)^n\) is TN on \(Z^\geq 0 \times Y\). Indeed, we can rewrite \(K(n, y)\) as \(e^{n \log(u(y))}\). Now given increasing tuples of elements \(n_j \in \mathbb{Z}^\geq 0\) and \(y_k \in Y_+\), the matrix \(K(n_j, y_k)\) is TP, by the total positivity of \(e^{xy}\) (shown in the previous section).

Similarly, \(L(y, m) := u(y)^m\) is also TP on \(Y \times \mathbb{Z}^\geq 0\). The result now follows by Corollary 6.1. That \(M\) is TN follows from the same arguments, via Remark 5.15. \(\square\)

This result implies the total positivity of the Hankel moment matrices (2.20). Indeed, setting \(u(y) = y\) on domains \(Y \subseteq [0, \infty)\), we obtain:

**Corollary 6.3.** Suppose \(Y \subseteq [0, \infty)\) and \(\mu\) is a non-negative measure on \(Y\) with infinite support. Then the moment matrix \(H_\mu\) is totally positive (of all orders).

We now show a result that will be used to provide another proof of the preceding corollary.

**Theorem 6.4** *(Andréief’s Identity).* Suppose \(Y \subseteq \mathbb{R}\) is a bounded interval, \(n > 0\) is an integer, and \(f_1, f_2, \ldots, f_n; g_1, g_2, \ldots, g_n : Y \to \mathbb{R}\) are integrable functions with respect to a positive measure \(\mu\) on \(Y\). Define \(y := (y_1, \ldots, y_n)\), and

\[
K(y) := (f_i(y_j))_{i,j=1}^n, \quad L(y) := (g_k(y_j))_{j,k=1}^n, \quad M' := \left(\int_Y f_i(y)g_k(y) \, d\mu(y)\right)_{i,k=1}^n.
\]

Then,

\[
\det M' = \frac{1}{n!} \int_{Y^n} \cdots \int_{Y^n} \det(K(y)) \det(L(y)) \prod_{j=1}^n d\mu(y_j).
\]
6. Hankel moment matrices are TP. Andréief’s identity.

Density of TP matrices in TN matrices.

Proof. We compute, beginning with the right-hand side:
\[
\int \cdots \int_{Y^n} \det(K(y)) \det(L(y)) \prod_{j=1}^{n} d\mu(y_j)
\]
\[
= \sum_{\sigma, \tau \in S_n} \sgn(\sigma) \sgn(\tau) \prod_{j=1}^{n} \int_{Y} f_{\sigma(j)}(y_j) g_{\tau(j)}(y_j) \, d\mu(y_j).
\]
Let $\beta = \sigma \tau^{-1}$. Then a change of variables shows that this expression equals
\[
= \sum_{\tau, \beta \in S_n} \sgn(\beta) \prod_{j=1}^{n} \int_{Y} g_{\tau(j)}(y) f_{\beta \tau(j)}(y) \, d\mu(y)
\]
\[
= \sum_{\tau \in S_n} \det \left( \int_{Y} f_{\tau(i)}(y) g_{\tau(k)}(y) \, d\mu(y) \right)_{i,k=1}^{n}
\]
\[
= n! \det M'.
\]
\[
\square
\]
As a special case, let $u : Y \to \mathbb{R}$ be positive and strictly increasing, and set $f_i(y) = u(y)^{n_i}, g_k(y) = u(y)^{m_k}$ for all $1 \leq i, k \leq n$ and increasing sequences of integers $n_1 < n_2 < \cdots$ and $m_1 < m_2 < \cdots$. Then the matrix $M'$ has $(i, k)$ entry $\int_{Y} u(y)^{n_i+m_k} \, d\mu(y)$. Now using Andréief’s identity – and the analysis from earlier in this section – we obtain a second proof of Proposition 6.2.

In particular, specializing to $u(y) = y$ and $Y \subset [0, \infty)$ reproves the total positivity of moment matrices $H_{\mu}$ for measures $\mu \geq 0$ with infinite support in $Y$. In this case we have $n_j = m_j = j - 1$ for $j = 1, \ldots, n$. The advantage of this proof (over using the generalized Cauchy–Binet formula) is that we can compute $\det M'$ ‘explicitly’ using Andréief’s identity:
\[
M' = (s_{i+k-2}(\mu))_{i,k=1}^{n}, \quad s_{i+k-2}(\mu) = \int_{Y} y^{i-1} y^{k-1} \, d\mu(y),
\]
\[
\det M' = \frac{1}{n!} \int_{Y^n} \prod_{1 \leq r < s \leq n} (y_s - y_r)^2 \, d\mu(y_1) d\mu(y_2) \cdots d\mu(y_n).
\]
(This uses the Vandermonde determinant identity $\det K(y) = \det L(y) = \prod_{1 \leq r < s \leq n} (y_s - y_r)$.)

6.2. Density of TP matrices in TN matrices. We will now prove an important density result due to Whitney. Standard/well-known examples of such results are:

(1) Every square real matrix can be approximated by non-singular real matrices.
(2) Symmetric non-singular real matrices are dense in symmetric real matrices.
(3) $n \times n$ positive definite matrices are dense in $\mathbb{P}_n$.

Our goal in this section is to prove the following

Theorem 6.7 (Whitney, 1952). Given positive integers $m, n \geq p$, the set of $TP_p$ $m \times n$ matrices is dense in the set of $TN_p$ $m \times n$ matrices.

In order to prove this theorem, we first prove a lemma by Pólya.

Lemma 6.8 (Pólya). For all $\sigma > 0$, the kernel $F_{\sigma} : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ given by
\[
F_{\sigma}(x, y) := e^{-\sigma(x-y)^2}
\]
is totally positive.
Note that the function \( f(x) = e^{-\sigma x^2} \) is such that the TP kernel \( F_\sigma(x, y) \) can be rewritten as \( f(x - y) \). Such functions are known as Polya frequency (PF) functions.

**Proof.** Given real numbers \( x_1 < x_2 < \cdots < x_n \) and \( y_1 < y_2 < \cdots < y_n \), we have:

\[
(F_\sigma(x_j, y_k))_{j,k=1}^n = (e^{-\sigma x_j^2}e^{2\sigma x_j y_k}e^{-\sigma y_k^2})_{j,k=1}^n,
\]

\[
= \text{diag}(e^{-\sigma x_j^2})_{j=1}^n (e^{2\sigma x_j y_1} \cdots e^{2\sigma x_j y_n})(e^{-\sigma y_k^2})_{k=1}^n,
\]

and this has positive determinant by the previous section (see Example 5.11). □

In a similar vein, we have:

**Lemma 6.9.** For all \( \sigma > 0 \), the kernel \( G_\sigma : \mathbb{R} \times \mathbb{R} \to \mathbb{R} \), given by

\[
G_\sigma(x, y) := e^{\sigma(x+y)^2}
\]

is totally positive. In particular, the kernels \( F_\sigma \) (from Lemma 6.8) and \( G_\sigma \) provide examples of TP Toeplitz and Hankel matrices, respectively.

**Proof.** The proof of the total positivity of \( G_\sigma \) is similar to that of \( F_\sigma \) above, and hence left as an exercise. To obtain TP Toeplitz and Hankel matrices, akin to Example 3.12 we choose any arithmetic progression \( x_1, \ldots, x_n \) of finite length, and consider the matrices with \((j, k)\)th entry \( F_\sigma(x_j, x_k) \) and \( G_\sigma(x_j, x_k) \), respectively. □

Now we come to the main proof of this section.

**Proof of Theorem 6.7.** Let \( A_{m \times n} \) be \( TN_p \) of rank \( r \). Define for each integer \( m > 0 \) the matrix

\[
(F_\sigma,m)_{m \times m} = (e^{-\sigma(j-k)^2})_{j,k=1}^m,
\]

\[
A(\sigma) = F_\sigma,m AF_\sigma,n.
\]

Note that \( F_\sigma,m \) is TP by Lemma 6.8 and \( F_\sigma,m \to \text{Id}_{m \times m} \) as \( \sigma \to \infty \). Now as the product of totally non-negative matrices is totally non-negative (Proposition 5.4), and \( F_\sigma,m \) is non-singular for all \( m \), we have that \( A(\sigma) \) is \( TN_p \) of rank \( r \).

**Claim 6.11.** \( A(\sigma) \) is TP_{min(r,p)}.

**Proof.** For any \( s \leq \min(r,p) \), let \( J, K \subseteq [\min(r,p)] \) of size \( s \). Using the Cauchy–Binet Formula, we compute:

\[
\det A(\sigma)_{J \times K} = \sum_{L,M \subseteq [\min(r,p)], |L|=s} \det(F_\sigma,m)_{J \times L} \det A_{L \times M} \det(F_\sigma,n)_{M \times K}.
\]

Now note that all \( 1 \leq k \leq \min(r,p) \), at least one \( k \times k \) minor of \( A \) is positive, and all other \( k \times k \) minors are non-negative. Combined with the total positivity of \( F_\sigma,m \) and \( F_\sigma,n \), this shows that \( \det A(\sigma)_{J \times K} > 0 \). This concludes the proof. □

Returning to the proof of the theorem, if \( r \geq p \) then the TP matrices \( A(\sigma) \) approximate \( A \) as \( \sigma \to \infty \); thus the proof is complete.

For the remainder of the proof, assume that \( A \) and \( A(\sigma) \) both have rank \( r < p \). Define

\[
A^{(1)} := A(\sigma) + \frac{1}{\sigma} E_{11},
\]

where \( E_{11} \) is the elementary \( m \times n \) matrix with (1,1) entry 1, and all other entries 0.
Claim 6.12. $A^{(1)}$ is $TN_p$ of rank $r + 1$.

Proof. Fix an integer $1 \leq s \leq p$ and subsets $J \subseteq [m], K \subseteq [n]$ of size $s$. Now consider the $s \times s$ minor $A^{(1)}_{J \times K}$. If $1 \not\in J$ or $1 \not\in K$, then we have:

$$\det A^{(1)}_{J \times K} = \det A(\sigma)_{J \times K} \geq 0,$$

whereas if $1 \in J \cap K$, then expanding along the first row or column yields:

$$\det A^{(1)}_{J \times K} = \det A(\sigma)_{J \times K} + \frac{1}{\sigma} \det A(\sigma)_{J \setminus \{1\} \times K \setminus \{1\}} \geq 0.$$

This shows that $A^{(1)}$ is $TN_p$.

As $A, A(\sigma)$ have rank $r$, and we are changing only one entry, all the $(r+2) \times (r+2)$ minors of $A^{(1)}$ will have determinant 0. We now claim $\det A^{(1)}_{[r+1] \times [r+1]} > 0$. Indeed, we compute:

$$\det A^{(1)}_{[r+1] \times [r+1]} = \det A(\sigma)_{[r+1] \times [r+1]} + \frac{1}{\sigma} \det A(\sigma)_{[r+1] \setminus \{1\} \times [r+1] \setminus \{1\}} > 0,$$

where the last inequality occurs because $A(\sigma)$ is $TP_r$ (shown above) and $\det A(\sigma)_{[r+1] \times [r+1]} = 0$. Thus the rank of $A^{(1)}$ is $r + 1$. \qed

Returning to the proof: note that $A^{(1)}$ also converges to $A$ as $\sigma \to \infty$. Inductively repeating this procedure, after $(p - r)$ iterations we obtain a matrix $A^{(p-r)}$, via the procedure

$$A^{(k)}(\sigma) := F_{\sigma,m} A^{(k)} F_{\sigma,n}, \quad A^{(k+1)} := A^{(k)}(\sigma) + \frac{1}{\sigma} E_{11}. \quad (6.13)$$

Moreover, $A^{(p-r)}$ is a $TN_p$ matrix with rank $p$. As $\min(r, p) = p$ for this matrix, it follows that $A^{(p-r)}(\sigma)$ is $TP_p$ with $A^{(p-r)}(\sigma) \to A^{(p-r-1)}(\sigma) \to \cdots \to A$ as $\sigma \to \infty$. Thus, $A$ can be approximated by $TP_p$ matrices, and the proof is complete. \qed
7. Density of symmetric TP matrices. (Non-)Symmetric TP completion problems.

7.1. Density of TP matrices in TN matrices. Recall the steps in proving Whitney’s Theorem 6.7 on the density of TP matrices in TN matrices (all of size m × n):

- The m × m Pólya matrix $F_{\sigma,m} := (e^{-\sigma(j-k)^2})_{j,k=1}^m$ is TP and symmetric.
- Given $A_{m \times n} \in TN_p$ of rank r, define $A(\sigma) := F_{\sigma,m} AF_{\sigma,n}$. Then $A(\sigma) \in TP_{\min(r,p)}$. If $r \geq p$ then we are done.
- Else suppose $r < p$. Then $A^{(1)} := A(\sigma) + \frac{1}{\sigma}E_{11}$ is $TP_{\min(r,p)}$ as well as $TN_p$, and of rank $r + 1$.
- Repeating this construction until we get $A^{(p-r)}$, we have $A^{(p-r)}(\sigma)$ is in $TP_p$, and converges to $A$ as $\sigma \to \infty$.

We now make some observations that further Whitney’s theorem. First, this density phenomenon also holds upon restricting to symmetric matrices:

**Proposition 7.1.** The set of symmetric TP $p \times n \times n$ matrices is dense in the set of symmetric TN $p \times n \times n$ matrices.

**Proof.** The proof of Theorem 6.7 (in the previous section, and also sketched above) goes through verbatim.

Second, a careful analysis of the above proof further shows that

$$A(\sigma)_{jk} = \sum_{l,m} e^{-\sigma(j-l)^2} a_{lm} e^{-\sigma(m-k)^2} \geq a_{jk}. $$

Thus, given $A_{m \times n}$ that is $TN_p$ (possibly symmetric) and $\epsilon > 0$, there exists $B_{m \times n}$ that is $TP_p$ (possibly symmetric) such that $0 \leq b_{jk} - a_{jk} < \epsilon$ for all $j, k$.

**Remark 7.2.** In fact we can further refine this: by working with $B(\sigma)$ for such $B$, we can insist on $0 < b_{jk} - a_{jk} < \epsilon$. From this it follows that given a (possibly symmetric) $TN_p$ matrix $A_{m \times n}$, there exists a sequence $B_l$ of $m \times n$ matrices, all of them $TP_p$ (and symmetric if $A$ is), such that $B_l \to A$ entrywise as $l \to \infty$, and moreover for all $j, k$,

$$(B_1)_{jk} > (B_2)_{jk} > \cdots > (B_l)_{jk} > \cdots > a_{jk}. $$

7.2. Matrix completion problems. The main question in matrix completion problems is as follows. Given a partially filled matrix (that is, a partial matrix), do there exist choices for the ‘missing’ entries such that the resulting matrix has specified properties?

For example: is it possible to complete

$$\begin{pmatrix}
1 & 0 & ? \\
2 & ? & ? \\
? & ? & ?
\end{pmatrix}$$

to a Toeplitz matrix? The answer is clearly yes: $\begin{pmatrix}
1 & 0 & a \\
2 & 1 & 0 \\
b & 2 & 1
\end{pmatrix}$, for arbitrary $a, b$. Similarly, the above partial matrix can be completed to a non-singular, singular, or totally non-negative matrix. However, it cannot be completed to a Hankel or a symmetric Toeplitz matrix, nor to a positive (semi)definite or totally positive matrix. These are examples of some matrix completion problems.

In this section, we discuss three TP completion problems. The first is to understand which $2 \times 2$ matrices can ‘embed in’ (or ‘extend to’) TP matrices.
36 7. Density of symmetric TP matrices. (Non-)Symmetric TP completion problems.

**Theorem 7.3.** Given integers \( m, n \geq 2 \) and a partial \( m \times n \) matrix
\[
\begin{pmatrix}
a & b & ? & \cdots & ? \\
c & d & ? & \cdots & ? \\
\vdots & \vdots & \ddots & \ddots & \vdots \\
\end{pmatrix}
\]
with real entries, the matrix can be completed to a totally positive matrix if and only if \( \begin{pmatrix} a & b \\ c & d \end{pmatrix} \) is totally positive.

**Proof.** The ‘only if’ part is obvious. For the ‘if’ part, we will work (without reference henceforth) with the matrix \( \begin{pmatrix} \beta a & \beta b \\ \beta c & \beta d \end{pmatrix} \) for some scalar \( \beta > 0 \). We show that this matrix always embeds inside a generalized Vandermonde matrix if it is TP; to show this, we will repeatedly appeal below to the total positivity of generalized Vandermonde matrices \( (x_1^\alpha) \) with \( x_1 \) and \( \alpha \) either both increasing or both decreasing. See Theorem 5.1 and Remark 5.3.

**Case 1:** Suppose three of the four entries \( a, b, c, d \) are equal (note that all four cannot be equal). Then up to rescaling, the possible matrices are
\[
A_1 = \begin{pmatrix} \lambda & 1 \\ 1 & 1 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 1 & \mu \\ \mu & 1 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 1 & \mu \\ 1 & \lambda \end{pmatrix}, \quad A_4 = \begin{pmatrix} 1 & 1 \\ 1 & \lambda \end{pmatrix},
\]
where \( \lambda > 1 > \mu > 0 \). Now \( A_1, \ldots, A_4 \) are generalized Vandermonde matrices \( (x_j^\alpha) \) with (respectively)
\[
(x_1, x_2, \alpha_1, \alpha_2) = (\lambda, 1, 1, 0), \quad (1, \mu, 1, 0), \quad (\mu, 1, 0, 1), \quad (1, \lambda, 0, 1).
\]
For \( A_1, A_2 \), choosing any \( x_2 > x_3 > \cdots > x_m > 0 \) and \( 0 > \alpha_3 > \cdots > \alpha_n \), we are done. The other two cases are treated similarly.

**Case 2:** Suppose two entries in a row or column are equal (but three entries are not). Up to rescaling, the possible matrices are
\[
A'_1 = \begin{pmatrix} 1 & 1 \\ \gamma & \delta \end{pmatrix}, \quad A'_2 = \begin{pmatrix} \delta & 1 \\ 1 & \gamma \end{pmatrix}, \quad A'_3 = \begin{pmatrix} \delta & 1 \\ \gamma & 1 \end{pmatrix}, \quad A'_4 = \begin{pmatrix} 1 & \gamma \\ \delta & 1 \end{pmatrix},
\]
where \( 0 < \gamma < \delta \), and \( \gamma, \delta \neq 1 \). Now \( A'_1, \ldots, A'_4 \) are generalized Vandermonde matrices \( (x_j^\alpha) \) with (respectively)
\[
(x_1, x_2, \alpha_1, \alpha_2) = (1, e, \log \gamma, \log \delta), \quad (e, 1, \log \delta, \log \gamma), \quad (\delta, \gamma, 1, 0), \quad (\gamma, \delta, 0, 1).
\]
The result follows as in the previous case.

**Case 3:** In all remaining cases, \( \{a, d\} \) is disjoint from \( \{b, c\} \). We now set \( \alpha_1 = 1 \), and claim that there exist scalars \( \beta, x_1, x_2 > 0 \) and \( \alpha_2 \in \mathbb{R} \) such that
\[
\beta \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} x_1 & x_1^{\alpha_2} \\ x_2 & x_2^{\alpha_2} \end{pmatrix}.
\]
(7.4)

To see why, denote \( L := \log(\beta) \) for \( \beta > 0 \), as well as \( A = \log(a) \), \( B = \log(b) \) etc. Now applying \( \log \) entrywise to both sides of (7.4), we have
\[
\begin{pmatrix} L + A & L + B \\ L + C & L + D \end{pmatrix} = \begin{pmatrix} \log x_1 & \alpha_2 \log x_1 \\ \log x_2 & \alpha_2 \log x_2 \end{pmatrix}.
\]
Taking determinants, we obtain:

\[(L + A)(L + D) - (L + B)(L + C) = 0 \implies L = \frac{BC - AD}{(A + D) - (B + C)},\]

where \(A + D > B + C\) since \(ad > bc\). Now check that \(x_1 = e^{L}a, x_2 = e^{L}c, \alpha_2 = \frac{L + B}{L + A}\) satisfies the conditions in (7.4). (Note here that by the assumptions on \(a, b, c, d\), the sum \(L + A\) is nonzero, as are \(L + B, L + C, L + D\) also.) This shows the claim.

To complete the proof, we need to check that \(\beta \left( x_1 \ x_2 \atop x_1^\alpha \ x_2^\alpha \right)\) is a generalized Vandermonde matrix. Since \(x_1 \neq x_2\) by choice of \(a, c\), we need to verify two cases. If \(x_1 < x_2\), then \(a < c\), so

\[\left( \frac{x_1}{x_2} \right)^{\alpha_2} = \frac{b}{d} < \frac{a}{c} = \frac{x_1}{x_2} < 1 \implies \alpha_2 > 1.\]

Hence we indeed get a generalized Vandermonde matrix \((x_j^{\alpha_k})_{2 \times 2}\) with increasing \(x_j\) and increasing \(\alpha_k\). The case when \(x_1 > x_2\) is similarly verified. \(\square\)

**Remark 7.5.** The above proof in fact shows that one can start out by assuming the matrix

\[
\begin{pmatrix}
a & b \\
c & d
\end{pmatrix}
\]

to be in any position, for example \(
\begin{pmatrix}
? & a & ? & ? & \ldots & ? \\
\end{pmatrix}
\)

Now such a matrix can be completed to a totally positive \(m \times n\) matrix, if and only if \(\begin{pmatrix} a & b \\ c & d \end{pmatrix}\) is totally positive.

Our second TP completion result embeds symmetric \(2 \times 2\) TP matrices into symmetric TP matrices, again in ‘leading position’.

**Theorem 7.6.** Given \(n \geq 2\) and a matrix \(\begin{pmatrix} a & b & \ldots & ? \\ b & c & \ldots & ? \\ \vdots & \vdots & \ddots & \vdots \\ ? & ? & \ldots & ? \end{pmatrix}\) with real entries, it can be completed to a symmetric totally positive matrix if and only if \(\begin{pmatrix} a & b \\ b & c \end{pmatrix}\) is totally positive.

To prove the theorem, it would suffice to embed the matrix \(\begin{pmatrix} 1 & b \\ b & c \end{pmatrix}\), where \(0 < b < \sqrt{c}\), inside the square matrix

\[
\frac{1}{K} V^T V = \frac{1}{K} \begin{pmatrix}
1 & 1 & \cdots & 1 \\
x_1 & x_2 & \cdots & x_K \\
\vdots & \vdots & \ddots & \vdots \\
x_1^{n-1} & x_2^{n-1} & \cdots & x_K^{n-1}
\end{pmatrix} \begin{pmatrix}
1 & x_1 & \cdots & x_1^{n-1} \\
x_2 & x_2 & \cdots & x_2^{n-1} \\
\vdots & \vdots & \ddots & \vdots \\
x_K & x_K & \cdots & x_K^{n-1}
\end{pmatrix},
\]

where \(K \geq n\) and \(V_{K \times n}\) is (part of) a ‘usual’ Vandermonde matrix \((x_j^{k-1})\). In terms of probability, this amounts to finding a discrete uniform random variable, supported on \(\{x_1, \ldots, x_K\}\), whose mean and variance equal \(b\) and \(c - b^2\), respectively. Thus, one needs to solve an inverse moment problem, with a discrete uniform distribution supported on a set of size at least \(n\).
We will prove the result by a parallel approach involving continuous random variables, using the Generalized Cauchy–Binet theorem or the Andréief identity.

**Proof.** Once again, one implication is obvious. To prove the other, without loss of generality, by rescaling we may assume $a = 1$. We have $\begin{pmatrix} 1 & b \\ b & c \end{pmatrix}$, where $c > b^2$. Using Proposition 6.2 it suffices to produce an increasing function $u : [0, 1] \to [0, \infty)$ such that $u$ has first two moments $b, c$.

We now produce such a function. Let $u_s(y) = b(s + 1)y^s$ for $s > 0$. Then $u_s$ has first two moments

$$\int_0^1 u_s(y) dy = b, \quad \int_0^1 u_s(y)^2 dy = \frac{b^2(s + 1)^2}{2s + 1}.$$  

Notice that the second moment converges to $b^2$ as $s \to 0^+$, and grows to $+\infty$ as $s \to \infty$. By the intermediate value theorem, there exists $s_0 \in (0, \infty)$ such that $u(y) = u_{s_0}(y)$ has second moment $c$, as desired. Now using Proposition 6.2, the matrix $\begin{pmatrix} 1 & b \\ b & c \end{pmatrix}$ embeds inside (i.e., can be completed to) the symmetric TP matrix $(\int_0^1 u_{s_0}(y)^{j+k-2} dy)^n_{j,k=1}$.

Finally, we extend Theorem 7.3 and the subsequent Remark to completions of matrices of arbitrary sizes, that are totally positive of arbitrary order:

**Theorem 7.7.** Suppose $m, n \geq 1$ and $1 \leq p \leq \min(m, n)$ are integers, and $J \subset [m], \ K \subset [n]$ are ‘sub-intervals’ containing $m', n'$ consecutive integers, respectively. Then a real $m' \times n'$ matrix $A'$ can be completed to a TP real $m \times n$ matrix, in positions $J \times K$, if and only if $A'$ is TP.

**Proof.** One implication is obvious. Conversely, suppose $A'_m \times n'$ is TP. It suffices to show that one can add an extra row either above or below $A'$ and obtain a TP matrix. Then the result follows by induction and taking transposes.

We first show the result by adding a row $(a_1, \ldots, a_n)$ at the bottom. For $a_1$, choose any positive real number. Now having defined $a_1, \ldots, a_k > 0$ for some $0 < k < n'$, we inductively define $a_{k+1}$ as follows. Define the $(m' + 1) \times (k + 1)$ matrix $B_k := \begin{pmatrix} A'_{m' \times k+1} \\ a_1 \cdots a_{k+1} \end{pmatrix}$ (with unknown $a_{k+1} > 0$), and consider the $1 \times 1, \ldots, p \times p$ submatrices $B'$ of $B_k$ which contain entries from the last row and last column, whence $a_{k+1}$. Compute det$(B')$ by expanding along the last row, and from right to left. Requiring det$(B') > 0$ yields an inequality for $a_{k+1}$, which is in fact a strict lower bound since the cofactor corresponding to $a_{k+1}$ is a lower-order minor of $A'$, whence positive. Working over all such minors det$(B')$ yields a finite set of lower bounds, so that it is possible to define $a_{k+1}$ and obtain all minors with ‘bottom corner’ $a_{k+1}$ and size at most $p \times p$ to be positive. By the induction hypothesis, all other minors of $B_k$ of size at most $p \times p$ are positive, so that $B_k$ is TP. Proceeding inductively, we obtain the desired $(m' + 1) \times n'$ completion of $A'$ that is TP.

The argument is similar to add a row $(a_1, \ldots, a_n)$ on top of $A'$: this time we proceed sequentially from right to left. First, $a_{n'}$ is arbitrary; then to define $a_k$ (given $a_{k+1}, \ldots, a_n$), we require $a_k$ to satisfy a finite set of inequalities (obtained by expanding det$(B')$ along the first row from left to right), and each inequality is again a strict lower bound. □

We conclude this part by studying the spectra of TP/TN matrices. For real square symmetric matrices \( A_{n \times n} \), recall Sylvester's criterion \( 2.7 \) which says modulo Theorem \( 2.4 \) that such a matrix \( A_{n \times n} \) has all principal minors non-negative (or positive), if and only if all eigenvalues of \( A \) are non-negative (positive).

Our goal in this section is to show a similar result for TP/TN matrices. More precisely, we will show the same statement as above, removing the words ‘symmetric’ and ‘principal’ from the preceding paragraph. In other words, not only are all minors of TP and TN matrices positive and non-negative respectively, but moreover, so are their eigenvalues. A slightly more involved formulation of this result is:

**Theorem 8.1.** Given positive integers \( m \geq p \geq n \geq p \), and a real matrix \( A_{m \times n} \), the following are equivalent:

1. For every square submatrix \( B \) of \( A \) of size \( \leq p \), we have \( \det(B) \) is non-negative (respectively positive). In other words, \( A \) is TN\(_p\) (respectively TP\(_p\)).
2. For every square submatrix \( B \) of \( A \) of size \( \leq p \), the eigenvalues of \( B \) are non-negative (respectively positive and simple).

Note that the analogous statement for positive semidefinite matrices clearly holds (as mentioned above), by Sylvester’s theorem.

We will follow the approach taken by Gantmakher and Krein in their 1937 paper in Compositio Math. This approach is also found in Chapter XIII.9 of F.R. Gantmakher’s book *The theory of matrices* (translated from the Russian by K.A. Hirsch); and in an expository account by A. Pinkus found in the conference proceedings *Total positivity and its applications* (of the 1994 Jaca meeting), edited by M. Gasca and C.A. Micchelli.

This approach relies on two well-known theorems, which are interesting in their own right. The first was shown by O. Perron in his 1907 paper in Math. Ann.:

**Theorem 8.2 (Perron).** Let \( A_{n \times n} \) be a square, real matrix with all positive elements. Then \( A \) has a simple, positive eigenvalue \( \lambda \) with an eigenvector \( u_0 \in \mathbb{R}^n \), such that:

1. For the remaining \( n - 1 \) eigenvalues \( \mu \in \mathbb{C} \), we have \( |\mu| < \lambda \).
2. The coordinates of \( u_0 \) are all nonzero and of the same sign.

This result has been studied and extended by many authors in the literature; notably, the Perron–Frobenius theorem is a key tool used in one of the approaches to studying finite state-space discrete time Markov chains. As these extensions are not central to the present discussion, we do not pursue them further.

**Proof.** Write \( v \geq u \) (or \( v > u \)) for \( u, v \in \mathbb{R}^n \) to denote the (strict) coordinatewise ordering: \( v_j \geq u_j \) (or \( v_j > u_j \)) for all \( 1 \leq j \leq n \). A first observation, used below, is:

\[
Au < Av \implies u \leq v \quad \text{in} \quad \mathbb{R}^n; \quad u \neq v \quad \implies \quad Au < Av. \quad (8.3)
\]

We now proceed to the proof. Define

\[
\lambda := \sup\{\mu \in \mathbb{R} : Au \geq \mu u \text{ for some nonzero vector } 0 \leq u \in \mathbb{R}^n\}.
\]

Now verify that

\[
0 < n \min_{j,k} a_{jk} \leq \lambda \leq n \max_{j,k} a_{jk};
\]

in particular, \( \lambda \) is well-defined. Now for each \( k \geq 1 \), there exist (rescaled) vectors \( u_k \geq 0 \) in \( \mathbb{R}^n \) whose coordinates sum to 1, and such that \( Au_k \geq (\lambda - 1/k)u_k \). But then the \( u_k \) belong to
a compact simplex $S$, whence there exists a subsequence converging to some vector $u_0 \in S$. It follows that $Au_0 \geq \lambda u_0$; if $Au_0 \neq \lambda u_0$, then an application of (8.3) leads to a contradiction to the maximality of $\lambda$.

Thus $Au_0 = \lambda u_0$ for nonzero $u_0 \geq 0$. But then $Au_0 > 0$, whence $u_0 = \lambda^{-1}Au_0$ has all positive coordinates. This proves part (b).

It remains to show part (a) and the simplicity of $\lambda$. First if $Av = \mu v$ for any eigenvalue $\mu$ of $A$ (and $v \neq 0$), then defining $|v| := (|v_1|, \ldots, |v_n|)^T$, we have:

$$A|v| \geq |Av| = |\mu v| = |\mu||v|.$$  

From this, we obtain $\lambda \geq |\mu|$.

Suppose for the moment that $|\mu| = \lambda$. Then $A|v| = \lambda |v|$, else (as above) an application of (8.3) leads to a contradiction to the maximality of $\lambda$. But then $A|v| = |Av|$ from the preceding computation. By the triangle inequality over $\mathbb{C}$, this shows all coordinates of $v$ have the same argument, which we can take to be $e^{i\theta} = 1$ by normalizing $v$. It follows that $A\nu = \lambda v$ from above, since now $v = |v|$. Hence $\mu = \lambda$.

Thus we have shown that if $Av = \mu v$ for $|\mu| = \lambda$ (and $v \neq 0$), then $\mu = \lambda$ and we may rescale to get $v \geq 0$. In particular, this shows part (a) modulo the simplicity of the eigenvalue $\lambda$.

Moreover, if $u_0, u'_0$ are linearly independent $\lambda$-eigenvectors for $A$, then one can come up with a linear combination $v \in \mathbb{R}u_0 + \mathbb{R}u'_0$ with at least one positive and one negative coordinate. This contradicts the previous paragraph, so it follows that $\lambda$ has geometric multiplicity one.

The final remaining task is to show that $\lambda$ is a simple eigenvalue of $A$. If not, then by the preceding paragraph there exists $u_1 \notin \mathbb{R}u_0$ such that $(Au_0 - \lambda u_0)$ and $Au_1 = \lambda u_1 + \mu u_0$ for some nonzero scalar $\mu$. Now since $A^T$ has the same eigenvalues as $A$, the above analysis there exists $v_0 \in \mathbb{R}^n$ such that $v_0^T A = \lambda v_0^T$. Hence:

$$\lambda v_0^T u_1 = v_0^T Au_1 = v_0^T (\lambda u_1 + \mu u_0).$$

But then $\mu \cdot v_0^T u_0 = 0$, which is impossible since $\mu \neq 0$ and $u_0, v_0 > 0$. This shows that $\lambda$ is simple, and concludes the proof. \qed

The second result we require is folklore: Kronecker’s theorem on compound matrices. We begin by introducing this family of auxiliary matrices, associated to each given matrix.

**Definition 8.4.** Fix a matrix $A_{m \times n}$ (which we take to be real, but the entries can lie in any unital commutative ring), and an integer $1 \leq r \leq \min(m, n)$.

1. Let $S_1, \ldots, S_{\binom{m}{r}}$ denote the $r$-element subsets of $[m] = \{1, \ldots, m\}$, ordered lexicographically. (Thus $S_1 = \{1, \ldots, r\}$ and $S_{\binom{m}{r}} = \{m-r+1, \ldots, m\}$.) Similarly, let $T_1, \ldots, T_{\binom{n}{r}}$ denote the $r$-element subsets of $[n]$ in lexicographic order.

Now define the $r$th compound matrix of $A$ to be a matrix $C_r(A)$ of dimension $\binom{m}{r} \times \binom{n}{r}$, whose $(j, k)$th entry is the minor $\det(A_{S_j \times T_k})$.

2. For $r = 0$, define $C_0(A) := \text{Id}_{1 \times 1}$.

We now collect together some basic properties of compound matrices:

**Lemma 8.5.** Suppose $m, n \geq 1$ and $0 \leq r \leq \min(m, n)$ are integers, and $A_{m \times n}$ a matrix.

1. Then $C_1(A) = A$, and $C_r(cA) = c^r C_r(A)$ for all scalars $c$.
2. $C_r(A^T) = C_r(A)^T$.
3. $C_r(\text{Id}_{n \times n}) = \text{Id}_{\binom{n}{r} \times \binom{n}{r}}$. 


(4) The Cauchy–Binet formula essentially says:

\[ C_r(AB) = C_r(A)C_r(B) \]

for matrices \( A_{m \times n} \) and \( B_{n \times p} \), where \( p \geq 1 \).

(5) As a consequence, \( \det(C_r(AB)) = \det(C_r(A))\det(C_r(B)) \) when \( m = n = p \) (i.e., \( A, B \) are square).

(6) As another consequence of the multiplicativity of \( C_r \), if \( A \) has rank \( 0 \leq r \leq \min(m, n) \), then \( C_j(A) \) has rank \( \binom{r}{j} \) for \( j = 0, 1 \), \( r, r + 1, \ldots \), \( \min(m, n) \).

(7) If \( A \) is square, then \( C_n(A) = \det(A) \); if \( A \) is moreover invertible, then \( C_r(A)^{-1} = C_r(A^{-1}) \).

(8) If \( A \) is upper/lower triangular, diagonal, symmetric, orthogonal, or normal, then \( C_r(A) \) has the same property.

Proof. We only sketch a couple of the proofs, and leave the others as exercises. If \( A \) has rank \( r \), then one can write \( A = M_{m \times r}N_{r \times n} \), where the columns of \( M \) are linearly independent, as are the rows of \( N \). But then \( C_r(A) \) is the product of a nonzero column vector \( C_r(M) \) and a nonzero row vector \( C_r(N) \), hence has rank 1. (Here we require the underlying ground ring to be an integral domain.)

The other case we consider here is when \( A \) is \((n \times n)\) and upper triangular. In this case let \( J = \{ j_1 < \cdots < j_r \} \) and \( K = \{ k_1 < \cdots < k_r \} \) be subsets of \([n]\), with \( J > K \) in the lexicographic order. Hence there exists a unique \( l \in [1, r] \) such that

\[ j_1 = k_1, \quad \cdots, \quad j_{l-1} = k_{l-1}, \quad k_l < j_l < j_{l+1} < \cdots < j_r. \]

It follows that \( A_{J \times K} \) is a block triangular matrix of the form

\[
\begin{pmatrix}
C_{(l-1)\times(l-1)} & D \\
0 & E_{(r-l+1)\times(r-l+1)}
\end{pmatrix},
\]

and that the leftmost column of \( E \) is the zero vector. Hence \( \det(A_{J \times K}) = 0 \) if \( J > K \). □

With Lemma 8.5 in hand, one can state and prove

**Theorem 8.6** (Kronecker). Let \( n \geq 1 \) and suppose the complex matrix \( A_{n \times n} \) has the multiset of eigenvalues \( \{ \lambda_1, \ldots, \lambda_n \} \). For all \( 0 \leq r \leq n \), the \( \binom{n}{r} \) eigenvalues of \( C_r(A) \) are precisely of the form \( \prod_{j \in S} \lambda_j \), where \( S \) runs over all \( r \)-element subsets of \([n]\).

In words, the eigenvalues of \( C_r(A) \) are precisely the \( \binom{n}{r} \) products of \( r \) distinct eigenvalues of \( A \).

Proof. Let \( J \) denote an \((upper\, triangular)\) Jordan canonical form of \( A \). That is, there exists an invertible matrix \( M \) satisfying \( MJM^{-1} = A \), with the diagonal entries of \( J \) given by \( \lambda_1, \ldots, \lambda_n \). Applying various parts of Lemma 8.5

\[ C_r(A) = C_r(M) C_r(J) C_r(M)^{-1}, \]

with \( C_r(J) \) upper triangular. Thus the eigenvalues of \( C_r(A) \) are precisely the diagonal entries of \( C_r(J) \), and these are precisely the claimed set of scalars. □

Finally, these ingredients help show that TP square matrices have simple, positive eigenvalues.

**Proof of Theorem 8.7**. Clearly, (2) implies (1). Conversely, first note that by focussing on a fixed square submatrix \( B \) and all of its minors, the implication \( (1) \implies (2) \) for general \( m, n \geq p \) reduces to the special case \( m = p = n \), which we assume henceforth.
We first show the TP case; thus say $A_{n \times n}$ is TP. Relabel its eigenvalues $\lambda_1, \ldots, \lambda_n$ such that

$$|\lambda_1| \geq \cdots \geq |\lambda_n|.$$ 

Now let $1 \leq r \leq n$; then the compound matrix $C_r(A)$ has positive entries, so by Perron’s theorem \[8.2\] there exists a unique largest positive eigenvalue $\lambda_{\max,r}$, and all others are smaller in modulus. Hence by Kronecker’s theorem \[8.6\] $\lambda_{\max,r} = \lambda_1 \cdots \lambda_r$, and we have

$$\lambda_1 \cdots \lambda_r > 0, \quad \forall 1 \leq r \leq n.$$

It follows that each $\lambda_j$ is positive. Moreover, from Perron and Kronecker’s results it also follows for each $1 \leq r \leq n-1$ that

$$\lambda_1 \cdots \lambda_r > \lambda_1 \cdots \lambda_{r-1} \cdot \lambda_{r+1},$$

and so $\lambda_r > \lambda_{r+1}$, as desired.

This shows the result for TP matrices. Now suppose $A_{n \times n}$ is TN. By Whitney’s density theorem \[6.7\] we may approximate $A$ by a sequence $A_k$ of TP matrices. Hence the characteristic polynomials converge: $p_{A_k}(t) := \det(t \text{Id}_{n \times n} - A_k) \to p_{A}(t)$ coefficientwise, as $k \to \infty$. Since deg$(p_{A_k}) = n$ for all $k \geq 1$, it follows by the ‘continuity of roots’ – proved below – that the eigenvalues of $p_{A}$ also avoid the open set $C \setminus [0, \infty)$. This concludes the proof. \qed

Thus, it remains to show that the roots of a real or complex polynomial are continuous functions of its coefficients. This is in fact a consequence of Hurwitz’s theorem in complex analysis, but we restrict ourselves here to mentioning a simpler result. The following argument can be found online or in books.

**Proposition 8.7.** Suppose $p_k \in \mathbb{C}[t]$ is a sequence of polynomials, with deg$(p_k)$ uniformly bounded over all $k \geq 1$. If $U \subset \mathbb{C}$ is an open set on which no $p_k$ vanishes, and $p_k(t) \to p(t)$ coefficientwise, then either $p \equiv 0$ on $U$, or $p$ is nonvanishing on $U$.

**Proof.** We restrict ourselves to outlining this argument, as this direction is not our main focus. Suppose $p|_U$ is not identically zero, and $p(w) = 0$ for some $w \in U$. Choose $\delta > 0$ such that the closed disc $\overline{D} := \overline{D}(w, \delta) \subset U$ and $p(t)$ has no roots in $\overline{D} \setminus \{w\}$. Then each $p_k$ is uniformly continuous on the compact boundary $\partial D = \overline{D} \setminus D$, where $D = D(w, \delta)$ is the open disc. For sufficiently large $k$, deg$(p_k) = \deg(p) \geq 0$ by the hypotheses. This is used to show that the $p_k$ converge uniformly on $\partial D$ to $p$, and similarly, $p_k' \to p'$ uniformly on $\partial D$.

Since $p$ is nonvanishing on $\partial D$, we have

$$m := \min_{z \in \partial D} |p(z)| > 0,$$

and hence for sufficiently large $k$, we have

$$\min_{z \in \partial D} |p_k(z)| > \frac{m}{2}, \quad \forall k \gg 0.$$

Using this, one shows that the sequence $\{p_k'/p_k : k > 0\}$ converges uniformly on $\partial D$ to $p'/p$.

Now integrate on $\partial D$: since $p_k'/p_k$ equals $\sum_j 1/(z - \lambda_j(p_k))$ where one sums over the multiset of roots $\lambda_j$ of each $p_k$, and since $p_k$ does not vanish in $U \supset \overline{D}$, we have

$$0 = \oint_{\partial D} \frac{p_k'(z)}{p_k(z)} \, dz \to \oint_{\partial D} \frac{p'(z)}{p(z)} \, dz.$$

Hence the right-hand integral vanishes. On the other hand, that same integral equals a positive integer – namely, the multiplicity of the root $w$ of $p(t)$. This yields the desired contradiction, whence $p$ does not vanish on $U$. \qed
Part 2:

Entrywise powers preserving (total) positivity in fixed dimension
Part 2: Entrywise powers preserving (total) positivity in fixed dimension

9. ENTRYWISE POWERS PRESERVING POSITIVITY IN FIXED DIMENSION:

In the rest of these notes, we discuss operations that preserve the notions of positivity that have been discussed earlier. Specifically, we will study functions that preserve positive semidefiniteness, or TP/TN, when applied entrywise to various classes of matrices. This part of the notes discusses the important special case of entrywise powers preserving positivity; to understand some of the motivations behind this study, we refer the reader to Section 13.1 below.

We begin with some preliminary definitions.

Definition 9.1. Given a subset $I \subset \mathbb{R}$, define $\mathbb{P}_n(I) := \mathbb{P}_n \cap \mathbb{R}^{n \times n}$ to be the set of $n \times n$ positive semidefinite matrices, all of whose entries are in $I$.

A function $f : I \to \mathbb{R}$ acts entrywise on $\mathbb{P}_n(I)$ via: $A = (a_{jk}) \mapsto f[A] := (f(a_{jk}))$.

Note that the entrywise operator $f[-]$ differs from the usual holomorphic calculus (except when acting on diagonal matrices by functions that vanish at the origin).

We fix the following notation for future use. If $f(x) = x^\alpha$ for some $\alpha \geq 0$ and $I \subset [0, \infty)$, then we write $A^\alpha$ for $f[A]$, where $A$ is any vector or matrix. By convention we shall take $0^0 = 1$ whenever required, so that $A^0$ is the matrix $1$ of all ones, and this is positive semidefinite whenever $A$ is square.

At this point, one can ask the following question: Which entrywise power functions preserve positive semidefiniteness, total positivity or total-negativity on $n \times n$ matrices? (We will also study later, the case of general functions.) The first of these questions was studied by Loewner in connection with the Bieberbach conjecture. It was eventually answered by two of his students, C.H. FitzGerald and R.A. Horn:

Theorem 9.2 (FitzGerald–Horn). Given an integer $n \geq 2$ and a scalar $\alpha \in \mathbb{R}$, $f(x) = x^\alpha$ preserves positive semidefiniteness on $\mathbb{P}_n((0, \infty))$ if and only if $\alpha \in \mathbb{Z} \geq 0 \cup [n - 2, \infty)$.

Remark 9.3. We will in fact show that if $\alpha$ is not in this set, there exists a rank two Hankel TN matrix $A_{n \times n}$ such that $A^\alpha \not\in \mathbb{P}_n$. (In fact, it is the (partial) moment matrix of a non-negative measure on two points.) Also notice that Theorem 9.2 holds for entrywise powers applied to $\mathbb{P}_n([0, \infty))$, since as we show, $\alpha < 0$ never works while $\alpha = 0$ always does so by convention; and for $\alpha > 0$ the power $x^\alpha$ is continuous on $[0, \infty)$, and we use the density of $\mathbb{P}_n((0, \infty))$ in $\mathbb{P}_n([0, \infty))$.

This ‘phase transition’ at $n - 2$ is a remarkable and oft-repeating phenomenon in the entrywise calculus (we will see additional examples of such events in Section 14). The value $n - 2$ is called the critical exponent for the given problem of preserving positivity.

To prove Theorem 9.2 we require a preliminary lemma, also by FitzGerald and Horn. Recall the preliminaries in Section 2.4.

Lemma 9.4. Given a matrix $A \in \mathbb{P}_n(\mathbb{R})$ with last column $\zeta$, the matrix $A - a_{nn}^\dagger \zeta \zeta^T$ is positive semidefinite with last row and column zero.

Here, $a_{nn}^\dagger$ denotes the Moore–Penrose inverse of the $1 \times 1$ matrix $(a_{nn})$.

Proof. If $a_{nn} = 0$ then $\zeta = 0$ by positive semidefiniteness, and $a_{nn}^\dagger = 0$ as well. The result follows. Now suppose $a_{nn} > 0$, and write $A = \begin{pmatrix} B & \omega \\ \omega^T & a_{nn} \end{pmatrix}$. Then a straightforward
computation shows that
\[ A - a_{nn}^{-1} \zeta \zeta^T = \left( B - \frac{\omega^T}{a_{nn}} \right) .\]

Notice that \( B - \frac{\omega x}{a_{nn}} \) is the Schur complement of \( A \) with respect to \( a_{nn} > 0 \). Now since \( A \) is positive semidefinite, so is \( B - \frac{\omega x}{a_{nn}} \).

**Proof of Theorem 9.2.** Notice that \( x^\alpha \) preserves positivity on \( \mathbb{P}_n((0, \infty)) \) for all \( \alpha \in \mathbb{Z}^{\geq 0} \), by the Schur product theorem 3.8. Now we prove by induction on \( n \geq 2 \) that if \( \alpha \geq n - 2 \), then \( x^\alpha \) preserves positivity on \( \mathbb{P}_n((0, \infty)) \). If \( n = 2 \) and \( A = \begin{pmatrix} a & b \\ b & c \end{pmatrix} \in \mathbb{P}_2((0, \infty)) \), then \( ac \geq b^2 \implies (ac)^\alpha \geq b^{2\alpha} \) for all \( \alpha \geq 0 \). It follows that \( A^\alpha \in \mathbb{P}_2((0, \infty)) \), proving the base case.

For the induction step, assume that the result holds for \( n - 1 \geq 2 \). Suppose \( A \in \mathbb{P}_n((0, \infty)) \) and \( \alpha \geq n - 2 \). If \( a_{nn} = 0 \) then by the positivity (psd) of \( A \), the entries in the last row and last column must vanish (consider the minors \( \begin{pmatrix} a_{jj} & a_{jn} \\ a_{nj} & a_{nn} \end{pmatrix} \) for \( 1 \leq j < n \)). But then we are done by the induction hypothesis.

The final case is when \( a_{nn} > 0 \). Now consider the following elementary definite integral:

\[ x^\alpha - y^\alpha = \alpha(x - y) \int_{0}^{1} (\lambda x + (1 - \lambda)y)^{\alpha - 1} \, d\lambda. \]  

Let \( \zeta \) denote the final column of \( A \); applying 9.5 entrywise to \( x \) an entry of \( A \) and \( y \) the corresponding entry of \( B := \frac{\zeta \zeta^T}{a_{nn}} \) yields:

\[ A^\alpha - B^\alpha = \alpha \int_{0}^{1} (A - B) \circ (\lambda A + (1 - \lambda)B)^{\alpha - 1} \, d\lambda. \]  

By the induction hypothesis, the leading principal \((n - 1) \times (n - 1)\) submatrix of the matrix \((\lambda A + (1 - \lambda)B)^{\alpha - 1}\) is positive semidefinite (even though the entire matrix need not be psd). By Lemma 9.4, \( A - B \) is psd and has last row and column zero. It follows by the Schur product theorem that the integrand on the right is positive semidefinite. Since \( B^\alpha \) is a rank-one psd matrix (this is easy to verify), it follows that \( A^\alpha \) is also psd. This concludes one direction of the proof.

To prove the other half, suppose \( \alpha \notin \mathbb{Z}^{\geq 0} \cup \{n - 2, \infty\} \); now consider \( H_\mu \) where \( \mu = \delta_1 + \epsilon \delta_x \) for \( \epsilon, x > 0, \, x \neq 1 \). Note that \( (H_\mu)_{jk} = s_{j+k}(\mu) = 1 + \epsilon x^{j+k} \); and as shown previously, \( H_\mu \) is positive semidefinite of rank 2.

First suppose \( \alpha < 0 \). Then consider the leading principal \( 2 \times 2 \) minor of \( H_\mu^\alpha \), which equals

\[ B := \begin{pmatrix} (1 + \epsilon)^\alpha & (1 + \epsilon x)^\alpha \\ (1 + \epsilon x)^\alpha & (1 + \epsilon x^2)^\alpha \end{pmatrix}. \]

We claim that \( \det B < 0 \), which shows \( H_\mu^\alpha \) is not psd. Indeed, note that

\[ (1 + \epsilon)(1 + \epsilon x^2) - (1 + \epsilon x)^2 = \epsilon(x - 1)^2 > 0, \]

so \( \det B = (1 + \epsilon)^\alpha (1 + \epsilon x^2)^\alpha - (1 + \epsilon x)^{2\alpha} < 0 \) because \( \alpha < 0 \).

Next suppose that \( \alpha \in (0, n - 2) \setminus \mathbb{N} \). Given \( x > 0 \), for small \( \epsilon \) we know by the binomial theorem that

\[ (1 + \epsilon x)^\alpha = 1 + \sum_{k \geq 1} \binom{\alpha}{k} \epsilon^k x^k, \quad \text{where} \quad \binom{\alpha}{k} = \frac{\alpha(\alpha - 1) \cdots (\alpha - k + 1)}{k!}. \]
We will produce $u \in \mathbb{R}^n$ such that $u^TH_\mu^{\alpha o}u < 0$; note this shows that $H_\mu^{\alpha o} \not\in \mathbb{P}_n$.

Starting with the matrix $H_\mu = \mathbf{11}^T + \epsilon v v^T$ where $v = (1, x, \ldots, x^{n-1})^T$, we obtain:

$$H_\mu^{\alpha o} = \mathbf{11}^T + \sum_{k=1}^{\lfloor |\alpha|+2 \rfloor} \epsilon^k \binom{\alpha}{k} (v^o k)(v^o k)^T + o(\epsilon^{\lfloor |\alpha|+2 \rfloor}), \quad (9.7)$$

where $o(\epsilon^{\lfloor |\alpha|+2 \rfloor})$ is a matrix such that the quotient of any entry by $\epsilon^{\lfloor |\alpha|+2 \rfloor}$ goes to zero as $\epsilon \to 0^+$. Note that the first term and the sum together contain at most $n$ terms. Since the corresponding vectors $1, v, v^o, \ldots, v^{o(|\alpha|+2)}$ are linearly independent (by considering the – possibly partial – usual Vandermonde matrix formed by them), there exists a vector $u \in \mathbb{R}^n$ satisfying:

$$u^T \mathbf{1} = u^T v = u^T v^o = \cdots = u^T v^{o(|\alpha|+1)} = 0, \quad u^T v^{o(|\alpha|+2)} = 1.$$ 

Substituting these into the above computation, we obtain:

$$u^T H_\mu^{\alpha o} u = \epsilon^{\lfloor |\alpha|+2 \rfloor} \binom{\alpha}{|\alpha|+2} + u^T \cdot o(\epsilon^{\lfloor |\alpha|+2 \rfloor}) \cdot u.$$ 

Since $\binom{\alpha}{|\alpha|+2}$ is negative if $\alpha$ is not an integer, it follows that

$$\lim_{\epsilon \to 0^+} \frac{u^T H_\mu^{\alpha o} u}{\epsilon^{\lfloor |\alpha|+2 \rfloor}} < 0.$$ 

Hence one can choose a small $\epsilon > 0$ such that $u^T H_\mu^{\alpha o} u < 0$. It follows for this $\epsilon$ that $H_\mu^{\alpha o}$ is not positive semidefinite. \hfill \Box

**Remark 9.8.** As the above proof reveals, the following are equivalent for $n \geq 2$ and $\alpha \in \mathbb{R}$:

1. The entrywise map $x^\alpha$ preserves positivity on $\mathbb{P}_n((0, \infty))$ (or $\mathbb{P}_n([0, \infty))$).
2. $\alpha \in \mathbb{Z}^\geq 0 \cup [n-2, \infty)$.
3. The entrywise map $x^\alpha$ preserves positivity on the (leading principal $n \times n$ truncations of) Hankel moment matrices of non-negative measures supported on $\{1, x\}$, for any fixed $x > 0$, $x \neq 1$.

The use of the Hankel moment matrix ‘counterexample’ $\mathbf{11}^T + \epsilon v v^T$ for $v = (1, x, \ldots, x^{n-1})^T$ and small $\epsilon > 0$ was not due to FitzGerald and Horn – who used $v = (1, 2, \ldots, n)^T$ instead – but due to Fallat, Johnson, and Sokal. In fact, the above proof can be made to work if one uses any vector $v$ with distinct positive real coordinates, and small enough $\epsilon > 0$.

As these remarks show, to isolate the entrywise powers preserving positivity on $\mathbb{P}_n((0, \infty))$, it suffices to consider a much smaller family – namely, the one-parameter family of truncated moment matrices of the measures $\delta_1 + \epsilon \delta_x$ – or the one-parameter family $\mathbf{1}_{n \times n} + \epsilon v v^T$, where $v = (x_1, \ldots, x_n)^T$ for pairwise distinct $x_j > 0$. In fact a stronger result is true. In her 2017 paper in *Linear Algebra Appl.*, Jain was able to eliminate the dependence on $\epsilon$.

**Theorem 9.9** (Jain). Suppose $n > 0$ is an integer, and $x_1, x_2, \ldots, x_n$ are pairwise distinct positive real numbers. Let $C := (1 + x_j x_k)^n_{j,k=1}$. Then $C^{\alpha o}$ is positive semidefinite if and only if $\alpha \in \mathbb{Z}^\geq 0 \cup [n-2, \infty)$.

In other words, this result identifies a multiparameter family of matrices, each one of which encodes the positivity preserving powers in the original result of FitzGerald–Horn. (The proof of this result is beyond the scope of this text.) We conclude by highlighting the power and applicability of the ‘integration trick’ (9.6) of FitzGerald and Horn. First, it in fact applies to general functions, not just to powers. The
The following observation (by the author and Tao) will be useful later, in the final part of these notes.

**Theorem 9.10 (Extension Principle).** Let $0 < \rho \leq \infty$ and $I = (0, \rho)$ or $(-\rho, \rho)$ or its closure. Fix an integer $n \geq 2$ and a continuously differentiable function $h : (0, \rho) \to \mathbb{R}$. If $h[-]$ preserves positivity on rank-one matrices in $\mathbb{P}_n(I)$ and $h'[-]$ preserves positivity on $\mathbb{P}_{n-1}(I)$, then $h[-]$ preserves positivity on $\mathbb{P}_n(I)$.

The proof is exactly as before, but now using the more general integral identity:

$$h(x) - h(y) = \int_x^y h'(t) \, dt = \int_0^1 (x - y)h'(\lambda x + (1 - \lambda)y) \, d\lambda.$$

Second, this integration trick is even more powerful, in that it further applies to classify the entrywise powers that preserve other properties of $\mathbb{P}_n$, including monotonicity and super-additivity. See Section 14 for details on these properties, their power-preservers, and their further application to positivity preservers on the distinguished sub-cones $\mathbb{P}_G$ for non-complete graphs $G$. 
10. ENTRYWISE POWERS PRESERVING TOTAL POSITIVITY: I.

Our next goal is to study which entrywise power functions preserve total positivity and total non-negativity. The present section is devoted to proving:

**Theorem 10.1.** If $A_{m \times n}$ is TP$_3$, then so is $A^t$ for all $t \geq 1$.

The proof relies on Descartes’ rule of signs (also known as Laguerre’s rule of signs). Recall that we had shown a ‘weak’ variant of this in Lemma 5.2. The next variant is stronger, and relies on the following notion.

**Definition 10.2.** Suppose $F : \mathbb{R} \to \mathbb{R}$ is infinitely differentiable. Given an integer $k \geq 0$, we say $F$ has a zero of order $k$ at $t_0 \in \mathbb{R}$, if $F(t_0) = F'(t_0) = \cdots = F^{(k-1)}(t_0) = 0$ and $F^{(k)}(t_0) \neq 0$. (Note that a zero of order 0 means that $F(t_0) \neq 0$.)

Descartes’ rule of signs bounds the number of real zeros of generalized Dirichlet polynomials, which are functions of the form

$$F : \mathbb{R} \to \mathbb{R}, \quad F(t) = \sum_{j=1}^{n} c_j e^{\alpha_j t}, \quad c_j, \alpha_j \in \mathbb{R}.$$  

These functions are so named because changing variables to $x = e^t$ gives

$$f(x) = \sum_{j=1}^{n} c_j x^{\alpha_j} : (0, \infty) \to \mathbb{R},$$

which are known as generalized polynomials. Another special case of $F(t)$ is when one uses $\alpha_j = -\log(j)$, to obtain $F(t) = \sum_{j=1}^{n} c_j/j^j$; these are called Dirichlet polynomials. The generalized Dirichlet polynomials subsume both of these families of examples.

We can now state

**Theorem 10.3** (Descartes’ rule of signs). Suppose $F(t) = \sum_{j=1}^{n} c_j e^{\alpha_j t}$ as above, with $c_j \in \mathbb{R}$ not all zero, and $\alpha_1 > \alpha_2 > \cdots > \alpha_n$ also real. Then the number of real zeros of $F$, counting multiplicities, is at most the number of sign changes in the sequence $c_1, c_2, \ldots, c_n$ (after removing all zero terms).

For instance, the polynomial $x^6 - 8 = (x^2 - 2)(x^4 + 2x^2 + 4)$ has only one sign change, so at most one positive root – which is at $x = e^{t} = 2$.

To prove Theorem 10.3 we require a couple of preliminary lemmas.

**Lemma 10.4** (Generalized Rolle’s theorem). Given an open interval $I$ and a smooth function $F : I \to \mathbb{R}$, let $Z(F, I)$ denote the number of zeros of $F$ in $I$, counting orders. If $Z(F, I)$ is finite, then we have $Z(F', I) \geq Z(F, I) - 1$.

**Proof.** Suppose $F$ has a zero of order $k_r > 0$ at $x_r, 1 \leq r \leq n$. Then $F'$ has a zero of order $k_r - 1 \geq 0$ at $x_r$. These add up to:

$$\sum_{r=1}^{n} (k_r - 1) = Z(F, I) - n$$

We may also suppose $x_1 < x_2 < \cdots < x_n$. Now by Rolle’s theorem, $F'$ also has at least $n - 1$ zeros in the intervals $(x_r, x_{r+1})$ between the points $x_r$. Together, we obtain: $Z(F', I) \geq Z(F, I) - 1$.  

**Lemma 10.5.** Let $F, G : I \to \mathbb{R}$ be smooth and $G \neq 0$ on $I$. If $F$ has a zero of order $k$ at $t_0$ then so does $F \cdot G$.  

Proof. This is straightforward: use Leibnitz’s rule to compute \((F \cdot G)^{(j)}(t_0)\) for \(0 \leq j \leq k\). □

With these lemmas in hand, we can prove Descartes’ rule of signs.

Proof of Theorem 10.3 The proof is by induction on the number \(s\) of sign changes in the sequence \(c_1, c_2, \ldots, c_n\). (Note that not all \(c_j\) are zero.) The base case is \(s = 0\), in which case \(F(t) = \sum_{j=1}^{n} c_j e^{\alpha_j t}\) has all nonzero coefficients of the same sign, and hence never vanishes.

For the induction step, we first assume without loss of generality that all \(c_j\) are nonzero. Suppose the last sign change occurs at \(c_k\), i.e., \(c_k c_{k-1} < 0\). Choose and fix \(\alpha \in (\alpha_k, \alpha_{k-1})\), and define \(G(t) := e^{-\alpha t}\). Then, \(H(t) := \sum_{j=1}^{n} c_j e^{(\alpha_j - \alpha) t}\) has the same zeros (with orders) as \(F(t)\), by Lemma 10.5. Moreover, \(H'(t) = \sum_{j=1}^{n} c_j (\alpha_j - \alpha) e^{(\alpha_j - \alpha) t}\) has exactly one less sign change than \(F(t)\), namely, \(s - 1\). It follows by the induction hypothesis that \(Z(H', \mathbb{R}) \leq s - 1\). Hence by Lemma 10.4 \(Z(F, \mathbb{R}) = Z(H, \mathbb{R}) \leq 1 + Z(H', \mathbb{R}) \leq s\), and the proof is complete by induction. □

We remark that there are even stronger versions of Descartes’ rule; see for instance G.J.O. Jameson, Counting zeros of generalised polynomials: Descartes’ rule of signs and Laguerre’s extensions, The Mathematical Gazette vol. 90 no. 518 (2006), pp. 223–234. Here we restrict ourselves to mentioning some of these variants without proofs (although we remark that their proofs are quite accessible). As above, let \(F(t) = \sum_{j=1}^{n} c_j e^{\alpha_j t}\) with \(\alpha_1 > \alpha_2 > \cdots > \alpha_n\).

1. Then not only is \(Z(F, \mathbb{R}) \leq s\), but \(s - Z(F, \mathbb{R})\) is an even integer.
2. Define the partial sums
   \[
   C_1 := c_1, \quad C_2 := c_1 + c_2, \quad \ldots, \quad C_n := c_1 + c_2 + \cdots + c_n.
   \]
Then the number of positive roots of \(F(t)\) is at most the number of sign changes in \(C_1, C_2, \ldots, C_n\).
3. Similarly, the number of negative roots of \(F(t)\) is the number of positive roots of \(F(-t)\), hence at most the number of sign changes in the ‘reverse sequence’
   \[
   D_1 := c_n, \quad D_2 := c_n + c_{n-1}, \quad \ldots, \quad D_n := c_n + c_{n-1} + \cdots + c_1.
   \]
Finally, we use Descartes’ rule of signs to show the result stated above: that all powers \(\geq 1\) preserve total positivity of order 3.

Proof of Theorem 10.1 It is easy to check that all entrywise powers \(\alpha \geq 1\) preserve the TP2 property. We now show that the positivity of all \(3 \times 3\) minors is also preserved by entrywise applying \(x^t\), \(t \geq 1\). Without loss of generality, we may assume \(m = n = 3\) and work with a TP matrix \(B_{3 \times 3} = (b_{jk})_{j,k=1}^{3}\).
The first claim is that we may assume $B = \begin{pmatrix} 1 & 1 & 1 \\ 1 & a & b \\ 1 & c & d \end{pmatrix}$. Indeed, define the diagonal matrices

$$D_1 := \begin{pmatrix} 1/b_{11} & 0 & 0 \\ 0 & 1/b_{21} & 0 \\ 0 & 0 & 1/b_{31} \end{pmatrix}, \quad D_2 := \begin{pmatrix} 1 & 0 & 0 \\ 0 & b_{11}/b_{12} & 0 \\ 0 & 0 & b_{11}/b_{13} \end{pmatrix}.$$ Using the Cauchy–Binet formula, one shows that $B_jk_{j,k=1}^3$ is totally positive if and only if $D_1BD_2$ is TP. But check that $D_1BD_2$ has only ones in its first row and column, as desired.

The next observation is that a matrix $A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & a & b \\ 1 & c & d \end{pmatrix}$ is TP if and only if $b,c > a > 1$ and $ad > bc$. (This is easily verified, and in turn implies $d > b,c$.)

Now consider $A^t$. For $t \geq 1$ this is TP by above, so we only need to consider $\det A^t$ for $t \geq 1$. Define the generalized Dirichlet polynomial

$$F(t) := \det(A^t) = (ad)^t - d^t - (bc)^t + b^t + c^t - a^t, \quad t \in \mathbb{R}.$$ Notice from the above inequalities involving $a,b,c,d$ that regardless of whether or not $d > bc$, and whether or not $b > c$, the sign sequence remains unchanged when arranging the exponents in $F$ (namely, $\log ad, \log d, \log bc, \log b, \ldots$) in decreasing order. It follows by Theorem \ref{thm:10.3} that $F$ has at most three real roots.

As $t \to \infty$, $F(t) \to \infty$. Now one can carry out a Taylor expansion of $F$ and check that the constant and linear terms vanish, yielding:

$$F(t) = e^{t \log(ad)} - e^{t \log(d)} - \cdots = t^2(\log(a) \log(d) - \log(b) \log(c)) + o(t^2).$$ It follows that $F$ has (at least) a double root at 0. We claim $F$ is indeed positive on $(1, \infty)$, as desired. For if $F$ is negative on $(1, \infty)$, then since $F(1) = \det A > 0$, it follows by continuity that $F$ has at least two more roots in $(1, \infty)$, which is false. Hence $F \geq 0$ on $(1, \infty)$. If $F(t_0) = 0$ for some $t_0 > 1$, then $t_0$ is a global minimum point in $[1, t_0 + 1]$ for $F$, whence $F'(t_0) = 0$. But then $F$ has at least two zeros at $t_0 \in (1, \infty)$, which is false. □
11. ENTRYWISE POWERS PRESERVING TOTAL POSITIVITY: II.

In the previous section, we used Descartes’ rule of signs to show that \( x^\alpha \) entrywise preserves the \( 3 \times 3 \) TP matrices, for all \( \alpha \geq 1 \). Our goal here is twofold: first, to completely classify the entrywise powers that preserve TP/TN for \( m \times n \) matrices for each fixed \( m, n \geq 1 \); and second, to then classify all continuous functions that do the same (at present, only for TN).

**Corollary 11.1.** If \( \alpha \geq 1 \), then \( x^\alpha \) entrywise preserves the \( 3 \times 3 \) TN matrices.

**Proof.** Let \( A_{3\times3} \) be TN and \( \alpha \geq 1 \). By Whitney’s density theorem \( 6.7 \), there exist \( 3 \times 3 \) TP matrices \( B_m \) that entrywise converge to \( A \), as \( m \to \infty \). Hence \( B_m^\alpha \to A^\alpha \) for \( \alpha \geq 1 \). Since \( B_m^\alpha \) is TP by Theorem \( 10.1 \), it follows that \( A^\alpha \) is TN, as claimed. \( \square \)

The next result classifies all entrywise powers preserving total non-negativity for matrices of any fixed size.

**Theorem 11.2.** Given integers \( m, n > 0 \), define \( d := \min(m, n) \). The following are equivalent for \( \alpha \in \mathbb{R} \).

1. \( x^\alpha \) preserves (entrywise) the \( m \times n \) TN matrices.
2. \( x^\alpha \) preserves (entrywise) the \( d \times d \) TN matrices.
3. Either \( \alpha = 0 \) (where we set \( 0^0 := 1 \)), or
   - (a) For \( d = 1, 2 \): \( \alpha \geq 0 \).
   - (b) For \( d = 3 \): \( \alpha \geq 1 \).
   - (c) For \( d \geq 4 \): \( \alpha = 1 \).

Thus we see that in contrast to the entrywise preservers of positive semidefiniteness (see Theorem \( 9.2 \)), almost no powers preserve the TN matrices – nor the TP matrices, as we show presently.

**Proof.** That (2) \( \implies \) (1) is straightforward, as is (1) \( \implies \) (2) by padding by zeros – noting that negative powers are not allowed (given zero entries of TN matrices). To show (3) \( \implies \) (2), we use Theorem \( 10.1 \) as well as that \( x^0 \) applied to any TN matrix yields the matrix of all ones.

It remains to prove (2) \( \implies \) (3). We may rule out negative powers since \( (0_{d\times d})^\alpha \) is not defined for \( d \geq 1 \). Similarly, \( x^0 \) always preserves total non-negativity. This shows (3) for \( d = 1, 2 \). For \( d = 3 \), suppose \( \alpha \in (0, 1) \) and consider the matrix \( A = \begin{pmatrix} 1 & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & 1 & \frac{1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} & 1 \end{pmatrix} \). This is a Toeplitz cosine matrix, hence TN (see Example \( 3.12 \) or verify directly). Now compute:

\[
\det A^\alpha = \det \begin{pmatrix} 1 & (\sqrt{2})^{-\alpha} & 0 \\ (\sqrt{2})^{-\alpha} & 1 & (\sqrt{2})^{-\alpha} \\ 0 & (\sqrt{2})^{-\alpha} & 1 \end{pmatrix} = 1 - 2^{1-\alpha},
\]

which is negative if \( \alpha < 1 \). So \( A^\alpha \) is not TN (not even positive semidefinite, in fact), for \( \alpha < 1 \), which shows (3) for \( d = 3 \).

Next suppose \( d = 4 \) and consider the matrix

\[
N(x) = 1_{4\times4} + x \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 2 & 3 \\ 0 & 2 & 4 & 6 \\ 0 & 3 & 8 & 14 \end{pmatrix}, \quad x \geq 0.
\]
One verifies that: all $2 \times 2$ minors are of the form $ax + bx^2$, where $a > 0, b \geq 0$; all $3 \times 3$ minors are of the form $cx^2$, where $c \geq 0$; and $\det N(x) = 0$. This implies $N(x)$ is TN for $x \geq 0$. Moreover, for small $x > 0$, computations similar to the proof of Theorem 10.1 show that

$$
\det N(x)^{ct} = 2(t^3 - t^4)x^4 + o(x^4),
$$

so given $t > 1$, it follows that $\det N(x)^{ct} < 0$ for sufficiently small $x > 0$. We conclude that $N(x)^{ct}$ is not TN, whence $x^\alpha$ does not preserve $4 \times 4$ TN matrices for $\alpha > 1$. If on the other hand $\alpha \in (0, 1)$, then we work with the $4 \times 4$ TN matrix $C = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$, proceeding as in the $d = 3$ case. This concludes the proof for $d = 4$.

Finally if $d > 4$ then we use the TN matrices $\begin{pmatrix} N(x) & 0 \\ 0 & 0 \end{pmatrix}_{d \times d}$; for small $x > 0$ this rules out the powers $\alpha > 1$ as above. Similarly, using the TN matrix $\begin{pmatrix} C & 0 \\ 0 & 0 \end{pmatrix}_{d \times d}$ rules out the powers in $(0, 1)$. □

In turn, Theorem 11.2 helps classify the powers preserving total positivity in each fixed size.

**Corollary 11.3.** Given $m, n > 0$, define $d := \min(m, n)$ as in Theorem 11.2. The following are equivalent for $\alpha \in \mathbb{R}$:

1. $x^\alpha$ preserves entrywise the $m \times n$ TP matrices.
2. $x^\alpha$ preserves entrywise the $d \times d$ TP matrices.
3. We have:
   a. For $d = 1$: $\alpha \in \mathbb{R}$.
   b. For $d = 2$: $\alpha > 0$.
   c. For $d = 3$: $\alpha \geq 1$.
   d. For $d \geq 4$: $\alpha = 1$.

**Proof.** That (2) $\implies$ (1) is straightforward, as is (1) $\implies$ (2) (as above) by now using Theorem 7.7. That (3) $\implies$ (2) was shown in Theorem 10.1 for $d = 3$, and is obvious for $d \neq 3$. Finally, we show (2) $\implies$ (3). The $d = 1$ case is trivial, while the $d = 2$ case follows by considering $\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$, say. Next, if $d \geq 3$ and if $x^\alpha$ preserves the $d \times d$ TP matrices, then $\alpha > 0$, by considering any TP matrix and applying $x^\alpha$ to any of its $2 \times 2$ minors. Hence $x^\alpha$ extends continuously to $x = 0$; now $x^\alpha$ preserves the $d \times d$ TN matrices by continuity. We are then done by Theorem 11.2. □

Next, we tackle the more challenging question of classifying all functions that entrywise preserve total positivity or total non-negativity in fixed dimension $m \times n$. We will show that (i) every such function must be continuous (barring a one-parameter exceptional family of TN preservers), and in turn, this implies that (ii) it must be a power function. We first show (ii), beginning with an observation on the (additive) Cauchy functional equation.

**Remark 11.4 (Additive continuous functions).** Suppose $g : \mathbb{R} \to \mathbb{R}$ is continuous and satisfies the Cauchy functional equation $g(x + y) = g(x) + g(y)$ for all $x, y \in \mathbb{R}$. Then we claim that $g(x) = cx$ for some $c \in \mathbb{R}$ (and all $x$). Indeed, $g(0 + 0) = g(0) + g(0)$, whence $g(0) = 0$. Next,
one shows by induction that \( g(n) = ng(1) \) for integers \( n > 0 \), and hence for all integers \( n < 0 \) as well. Now one shows that \( pg(1) = g(p) = g(q \cdot p/q) = q \cdot g(p/q) \) for integers \( p, q \) with \( q \neq 0 \), from which it follows that \( g(p/q) = (p/q)g(1) \) for all rationals \( p/q \). Finally, using continuity we conclude that \( g(x) = xg(1) \) for all \( x \in \mathbb{R} \).

**Proposition 11.5.** Suppose \( f : [0, \infty) \to \mathbb{R} \) is continuous and entrywise preserves the \( 2 \times 2 \) TN matrices. Then \( f(x) = f(1)x^\alpha \) for some \( \alpha \geq 0 \).

We recall here that \( 0^0 := 1 \) by convention.

**Proof.** Define the matrices

\[
A(x, y) = \begin{pmatrix} x & xy \\ 1 & y \end{pmatrix}, \quad B(x, y) = \begin{pmatrix} xy & x \\ y & 1 \end{pmatrix}, \quad x, y \geq 0.
\]

Clearly, these matrices are TN, whence by the hypotheses,

\[
\det f[A(x, y)] = f(x)f(y) - f(1)f(xy) \geq 0,
\]

\[
\det f[B(x, y)] = f(1)f(xy) - f(x)f(y) \geq 0.
\]

It follows that

\[
f(x)f(y) = f(1)f(xy), \quad \forall x, y \geq 0. \tag{11.6}
\]

There are two cases. First if \( f(1) = 0 \) then choosing \( x = y \geq 0 \) in (11.6) gives \( f \equiv 0 \) on \([0, \infty)\). Else if \( f(1) > 0 \) then we claim that \( f \) is always positive on \((0, \infty)\). Indeed, if \( f(x_0) = 0 \) for \( x_0 > 0 \), then set \( x = x_0, y = 1/x_0 \) in (11.6) to get: \( 0 = f(1)^2 \), which is false.

Now define the functions

\[
g(x) := f(x)/f(1), \quad x > 0, \quad h(y) := \log g(e^y), \quad y \in \mathbb{R}.
\]

Then (11.6) can be successively reformulated as:

\[
g(xy) = g(x)g(y), \quad \forall x, y > 0,
\]

\[
h(a + b) = h(a) + h(b), \quad \forall a, b \in \mathbb{R}. \tag{11.7}
\]

Moreover, both \( g, h \) are continuous. Since \( h \) satisfies the additive Cauchy functional equation, it follows by Remark 11.4 that \( h(y) = yh(1) \) for all \( y \in \mathbb{R} \). Translating back, we get \( g(x) = x^{h(1)} \) for all \( x > 0 \). It follows that \( f(x) = f(1)x^\alpha \) for \( x > 0 \), where \( \alpha = h(1) \). Finally, since \( f \) is also continuous at \( 0^+ \), it follows that \( \alpha \geq 0 \); and either \( \alpha = 0 \) and \( f \equiv 1 \) (so we set \( 0^0 := 1 \)), or \( f(0) = 0 < \alpha \). (Note that \( \alpha \) cannot be negative, since \( f[-] \) preserves TN on the zero matrix, say.) \( \square \)

**Corollary 11.8.** Suppose \( f : [0, \infty) \to \mathbb{R} \) is continuous and entrywise preserves the \( m \times n \) TN matrices, for some \( m, n \geq 2 \). Then \( f(x) = f(1)x^\alpha \) for some \( \alpha \geq 0 \), with \( f(1) \geq 0 \).

**Proof.** Given \( m, n \geq 2 \), every \( 2 \times 2 \) TN matrix can be embedded as a leading principal submatrix in a \( m \times n \) TN matrix, by padding it with (all other) zero entries. Hence the hypotheses imply that \( f[-] \) preserves the \( 2 \times 2 \) TN matrices, and we are done by the above Proposition 11.5 \( \square \)


We continue working toward the classification of all entrywise functions preserving $m \times n$ TP/TN matrices. Thus far, we have classified the power functions among these preservers; and we also showed that every continuous map that preserves $m \times n$ TN matrices is a multiple of a power function.

We now show that every function that entrywise preserves the $m \times n$ TP/TN matrices is automatically continuous on $(0, \infty)$ – which allows us to classify all such preservers. The continuity will follow from a variant of a 1929 result by Ostrowski on mid-convex functions on normed linear spaces, and we begin by proving this result.

12.1. Mid-convex functions.

Definition 12.1. Given a convex subset $U$ of a real vector space, a function $f : U \to \mathbb{R}$ is said to be mid-convex if

$$f\left(\frac{x+y}{2}\right) \leq \frac{f(x) + f(y)}{2}, \quad \forall x, y \in U.$$ 

We say that $f$ is convex if $f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y)$ for all $x, y \in U$ and $\lambda \in (0, 1)$.

Notice that convex functions are automatically mid-convex. The converse need not be true in general. However, if a mid-convex function is continuous then it is easy to see that it is also convex. Thus, a natural question for mid-convex functions is to find sufficient conditions under which they are continuous. One such condition is that the function is measurable, see e.g. the 1919 paper of Blumberg in Trans. Amer. Math. Soc. or the 1920 paper of Sierpinsky in Fund. Math. In these notes, we will use a different such condition: $f$ is locally bounded, on one neighborhood of one point.

Theorem 12.2. Let $\mathbb{B}$ be a normed linear space (over $\mathbb{R}$) and let $U$ be a convex open subset. Suppose $f : U \to \mathbb{R}$ is mid-convex and $f$ is bounded above in an open neighborhood of a single point $x_0 \in U$. Then $f$ is continuous on $U$, and hence convex.

This generalizes to normed linear spaces a special case of a result by Ostrowski, who showed in Jber. Deut. Math. Ver. (1929) the same conclusion, but over $\mathbb{B} = \mathbb{R}$ and assuming $f$ is bounded above in a measurable subset.

The proof requires the following useful observation:

Lemma 12.3. If $f : U \to \mathbb{R}$ is mid-convex, then $f$ is rationally convex, i.e., $f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y)$ for all $x, y \in U$ and $\lambda \in (0, 1) \cap \mathbb{Q}$.

Proof. Inductively using mid-convexity, it follows that

$$f\left(\frac{x_1 + x_2 + \cdots + x_{2^n}}{2^n}\right) \leq \frac{f(x_1) + \cdots + f(x_{2^n})}{2^n}, \quad \forall n \in \mathbb{N}, \ x_1, \ldots, x_{2^n} \in U.$$
Proof of Theorem 12.2. We may assume without loss of generality that \( \lambda = \frac{p}{q} \in (0,1) \), where \( p, q > 0 \) are integers and \( 2^{n-1} \leq q < 2^n \) for some \( n \in \mathbb{N} \). Let \( x_1, \ldots, x_q \in U \), and define \( \pi = \frac{x_1 + \cdots + x_q}{q} \). Then,

\[
   f \left( \frac{x_1 + \cdots + x_q + (2^n - q)x}{2^n} \right) \leq \frac{f(x_1 + \cdots + f(x_q) + (2^n - q)f(\pi)}{2^n}
\]

\[
   \implies 2^n f(\pi) \leq f(x_1) + \cdots + f(x_q) + (2^n - q)f(\pi)
\]

\[
   \implies q f(\pi) \leq f(x_1) + \cdots + f(x_q)
\]

\[
   \implies f \left( \frac{x_1 + \cdots + x_q}{q} \right) \leq \frac{f(x_1) + \cdots + f(x_q)}{q}
\]

In this inequality, set \( x_1 = \cdots = x_p = x \) and \( x_{p+1} = \cdots = x_q = y \) to get:

\[
   f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y),
\]
as desired. \( \Box \)

With Lemma 12.3 in hand, we can prove the theorem above.

**Proof of Theorem 12.2.** We may assume without loss of generality that \( x_0 = 0 \in U \subset \mathbb{B} \), and also that \( f(x_0) = f(0) = 0 \).

We claim that \( f \) is continuous at \( 0 \), where \( f \) was assumed to be bounded above in an open neighborhood of \( 0 \). Write this as: \( f(B(0,r)) < M \) for some \( r, M > 0 \), where \( B(x,r) \subset \mathbb{B} \) denotes the open ball of radius \( r \) centered at \( x \in \mathbb{B} \). Now given \( \epsilon \in (0,1) \cap \mathbb{Q} \) rational, and \( x \in B(0,\epsilon/2) \), we compute using Lemma 12.3:

\[
   x = \epsilon \left( \frac{x}{\epsilon} \right) + (1 - \epsilon)0 \implies f(x) \leq \epsilon f \left( \frac{x}{\epsilon} \right) + 0 < \epsilon M.
\]

Moreover,

\[
   0 = \left( \frac{\epsilon}{1 + \epsilon} \right) \left( \frac{x}{\epsilon} \right) + \frac{x}{1 + \epsilon},
\]

so applying Lemma 12.3 once again, we obtain:

\[
   0 \leq \left( \frac{\epsilon}{1 + \epsilon} \right) f \left( \frac{-x}{\epsilon} \right) + \frac{f(x)}{1 + \epsilon} < \frac{\epsilon M}{1 + \epsilon} + \frac{f(x)}{1 + \epsilon} \implies f(x) > -\epsilon M.
\]

Therefore, we have \( x \in B(0,\epsilon/2) \implies |f(x)| < \epsilon M \).

Now given \( \epsilon > 0 \), choose \( 0 < \epsilon' < \min(M, \epsilon) \) such that \( \epsilon'/M \) is rational, and set \( \delta := \epsilon'/M \). Then \( \delta < r \), so

\[
   x \in B(0,\delta) \implies |f(x)| < \delta M/r = \epsilon < \epsilon.
\]

This proves the continuity of \( f \) at \( x_0 \).

We have shown that if \( f \) is bounded above in some open neighborhood of \( x_0 \in U \), then \( f \) is continuous at \( x_0 \). To finish the proof, we claim that for all \( y \in U \), \( f \) is bounded above on some open neighborhood of \( y \). This would show that \( f \) is continuous on \( U \), which combined with mid-convexity implies convexity.

To show the claim, choose a rational \( \rho > 1 \) such that \( \rho y \in U \) (this is possible as \( U \) is open) and set \( U_y := B(y, (1 - 1/\rho)r) \). Note that \( U_y \subset U \) since for every \( v \in U_y \) there exists \( x \in B(0, r) \) such that

\[
   v = y + (1 - 1/\rho)x = \frac{1}{\rho}(\rho y) + \left( 1 - \frac{1}{\rho} \right)x,
\]

The test set of Hankel TN matrices.

and hence $v$ is a convex combination of $\rho y \in U$ as well as $x \in B(0,r) \subset U$. Hence $U_y \subset U$; now by Lemma 12.3

$$f(v) \leq \frac{1}{\rho} f(\rho y) + \left( 1 - \frac{1}{\rho} \right) f(x) \leq \frac{f(\rho y)}{\rho} + \left( 1 - \frac{1}{\rho} \right) M, \quad \forall v \in U_y.$$  

Since the right-hand side is independent of $v \in U_y$, the above claim follows. Hence by the first claim, $f$ is indeed continuous at every point in $U$. □

12.2. Functions preserving total non-negativity. With Theorem 12.2 in hand, it is possible to classify all entrywise functions that preserve total non-negativity or total positivity in a fixed size, or even positive semidefiniteness on $2 \times 2$ matrices. A major portion of the work is carried out by the next result. To state this result, we need the following notion.

Definition 12.4. Suppose $I \subset [0, \infty)$ is an interval. A function $f: I \to [0, \infty)$ is multiplicatively mid-convex on $I$ if and only if $f(\sqrt{xy}) \leq \sqrt{f(x)f(y)}$ for all $x,y \in I$.

Remark 12.5. A straightforward computation yields that if $f: I \to \mathbb{R}$ is always positive, and $0 \notin I$, then $f$ is multiplicatively mid-convex on $I$ if and only if the auxiliary function $g(y) := \log f(e^y)$ is mid-convex on $\log(I)$.

We now prove the following important result, which is also crucial later.

Theorem 12.6. Suppose $I = [0, \infty)$ and $I^+ := I \setminus \{0\}$. A function $f: I \to \mathbb{R}$ satisfies $(f(a) \quad f(b) \\ f(b) \quad f(c))$ is positive semidefinite whenever $a,b,c \in I$ and $(a \quad b \\ b \quad c)$ is TN, if and only if $f$ is non-negative, non-decreasing, and multiplicatively mid-convex on $I$. In particular,

1. $f|_{I^+}$ is never zero or always zero.
2. $f|_{I^+}$ is continuous.

The same results hold if $I = [0, \infty)$ is replaced by $I = (0, \infty)$, $[0, \rho)$, or $(0, \rho)$ for $0 < \rho < \infty$.

This result was essentially proved by H.L. Vasudeva, under some reformulation. In the result, note that TN is the same as ‘positive semidefinite with non-negative entries’, since we are dealing with $2 \times 2$ matrices; thus, the test set of matrices is precisely $P_2(I)$, and the hypothesis can be rephrased as:

$$f[\cdot]: P_2(I) \to P_2 = P_2(\mathbb{R}).$$

Moreover, all of these matrices are clearly Hankel. This result will therefore also play an important role later, when we classify the entrywise preservers of positive semidefiniteness on low-rank Hankel matrices.

Proof. Let $I$ be any of the domains mentioned in the theorem. We begin by showing the equivalence. Given a TN matrix

$$\begin{pmatrix} a & b \\ b & c \end{pmatrix}, \quad a,b,c \in I, \quad 0 \leq b \leq \sqrt{ac},$$

we compute using the non-negativity, monotonicity, and multiplicative mid-convexity respectively:

$$0 \leq f(b) \leq f(\sqrt{ac}) \leq \sqrt{f(a)f(c)}.$$  

It follows that $(f(a) \quad f(b) \\ f(b) \quad f(c))$ is TN and hence positive semidefinite.
Conversely, if (via the above remarks) \( f[-] : \mathbb{P}_2(I) \to \mathbb{P}_2 \), then apply the entrywise \( f[-] \) to the matrices

\[
\begin{pmatrix}
  a & a/2 \\
  a/2 & a
\end{pmatrix}, \quad \begin{pmatrix}
  a & b \\
  b & a
\end{pmatrix}, \quad \begin{pmatrix}
  a & \sqrt{ac} \\
  \sqrt{ac} & c
\end{pmatrix},
\]

with \( 0 \leq b \leq a \) and \( a, b, c \in I \). From the hypotheses, it successively (and respectively) follows that \( f \) is non-negative, non-decreasing, and multiplicatively mid-convex. This proves the equivalence.

We now show the two assertions in turn, again on \( I^+ \) for any of the domains \( I \) above; in other words, \( I^+ = (0, \rho) \) for \( 0 < \rho \leq \infty \). For (1), suppose \( f(x) = 0 \) for some \( x \in I^+ \). Since \( f \) is non-negative and non-decreasing on \( I^+ \), it follows that \( f \equiv 0 \) on \((0, x)\). We claim that \( f(y) = 0 \) if \( y > x \), \( y \in I^+ = (0, \rho) \). Indeed, choose \( n \gg 0 \) such that \( y \sqrt{y/x} < \rho \). Set \( \zeta := \sqrt{y/x} > 1 \), and consider the following rank-one matrices in \( \mathbb{P}_2(I^+) \):

\[
A_1 := \begin{pmatrix} x & x_\zeta \\ x_\zeta & x_\zeta^2 \end{pmatrix}, \quad A_2 := \begin{pmatrix} x_\zeta & x_\zeta x_\zeta^2 \\ x_\zeta^2 & x_\zeta x_\zeta^3 \end{pmatrix}, \quad \ldots, \quad A_n := \begin{pmatrix} x_\zeta^{n-1} & x_\zeta^n \\ x_\zeta^n & x_\zeta^{n+1} \end{pmatrix}.
\]

The inequalities \( \det f[A_k] \geq 0 \), \( 1 \leq k \leq n \) yield:

\[
0 \leq f(x_\zeta^k) \leq \sqrt{f(x_\zeta^{k-1})f(x_\zeta^{k+1})}, \quad k = 1, 2, \ldots, n.
\]

From this inequality for \( k = 1 \), it follows that \( f(x_\zeta) = 0 \). Similarly, these inequalities inductively yield: \( f(x_\zeta^k) = 0 \) for all \( 1 \leq k \leq n \). In particular, we have \( f(y) = f(x_\zeta^n) = 0 \). This shows that \( f \equiv 0 \) on \( I^+ \), as claimed.

To show (2), if \( f \equiv 0 \) on \( I^+ \), then \( f \) is continuous on \( I^+ \). Otherwise by (1), \( f \) is strictly positive on \((0, \rho) = I^+ \). Now define the function \( g : \log I^+ := (-\infty, \log \rho) \to \mathbb{R} \) via:

\[
g(y) := \log f(e^y), \quad y < \log \rho.
\]

By the assumptions on \( f \) and the observation in Remark 12.5, \( g \) is mid-convex and non-decreasing on \((-\infty, \log \rho) \). In particular, \( g \) is bounded above on compact sets. Now apply Theorem 12.2 to deduce that \( g \) is continuous. It follows that \( f \) is continuous on \((0, \rho) \).

As an application, Theorem 12.6 allows us to complete the classification of all entrywise maps that preserve total non-negativity in each fixed size.

**Theorem 12.7.** Suppose \( f : [0, \infty) \to \mathbb{R} \) entrywise preserves the \( m \times n \) TN matrices, for some \( m, n \geq 2 \). Then either \( f(x) = f(1)x^\alpha \) for \( f(1) \neq 0 \), \( \alpha \geq 0 \) and all \( x \geq 0 \) (and these powers were classified in Theorem 11.2), or \( \min(m, n) = 2 \) and \( f(x) = f(1)\text{sgn}(x) \) for \( x \geq 0 \) and \( f(1) > 0 \).

If instead \( \min(m, n) = 1 \), then \( f \) can be any function that maps \([0, \infty)\) into itself.

**Proof.** The result is trivial for \( \min(m, n) = 1 \), so we assume henceforth that \( m, n \geq 2 \). By embedding \( 2 \times 2 \) TN matrices inside \( m \times n \) TN matrices, it follows that \( f[-] \) preserves the \( 2 \times 2 \) TN matrices. In particular, \( f \) is continuous on \((0, \infty) \) by Theorem 12.6 and nonnegative and non-decreasing on \([0, \infty) \). Now one can repeat the proof of Proposition 11.5 above, to show that

\[
f(x)f(y) = f(xy)f(1), \quad \forall x, y \geq 0,
\]

and moreover, either \( f \equiv 0 \) on \([0, \infty) \), or \( f(x) = f(1)x^\alpha \) for \( x > 0 \) and some \( \alpha \geq 0 \).

We assume henceforth that \( f \neq 0 \) on \([0, \infty) \), whence \( f(x) = f(1)x^\alpha \) as above with \( f(1) > 0 \). If now \( f(0) \neq 0 \), then substituting \( x = 0, y \neq 1 \) in \( (12.8) \) shows that \( \alpha = 0 \), and now using \( x = y = 0 \) in \( (12.8) \) shows that \( f(0) = f(1) \), i.e., \( f|_{[0, \infty)} \) is constant (and positive).

The test set of Hankel TN matrices.

Otherwise \( f(0) = 0 \). Now if \( \alpha > 0 \) then \( f(x) = f(1)x^\alpha \) for all \( x \geq 0 \) and \( f \) is continuous on \([0, \infty)\). The final case is where \( f(0) = 0 = \alpha \), but \( f \neq 0 \). Then \( f(0) = 0 \) while \( f(x) = f(1) > 0 \) for all \( x > 0 \). Now if \( \min(m, n) = 2 \) then it is easy to verify that \( f[-] \) preserves \( TP \) matrices. On the other hand, if \( m, n \geq 3 \), then computing \( \det f[A] \) for the matrix

\[
A = \begin{pmatrix}
1 & \frac{1}{\sqrt{2}} & 0 \\
\frac{1}{\sqrt{2}} & 1 & \frac{1}{\sqrt{2}} \\
0 & \frac{1}{\sqrt{2}} & 1
\end{pmatrix}
\]

shows that \( f[-] \) is not a positivity preserver on \( A \oplus 0_{(m-3) \times (n-3)} \in TN_{m \times n} \).

12.3. Functions preserving total positivity. Akin to the above results, we can also classify the entrywise functions preserving total positivity in any fixed size, and they too are essentially power functions.

**Theorem 12.9.** Suppose \( f : (0, \infty) \to \mathbb{R} \) is such that \( f[-] \) preserves the \( m \times n \) \( TP \) matrices for some \( m, n \geq 2 \). Then \( f \) is continuous and \( f(x) = f(1)\alpha \forall x > 0 \), for some \( \alpha \geq 0 \) and \( f(1) > 0 \).

Recall that the powers preserving the \( m \times n \) \( TP \) matrices were classified in Corollary [11.3]

**Proof.** The proof is divided into steps; in the first four steps, we work in the case \( m = n = 2 \). The final step (Step 5) proves the result for general \( m, n \geq 2 \).

**Step 1:** Unless specified otherwise, we suppose henceforth that \( m = n = 2 \). We claim that \( f \) is positive and strictly increasing on \((0, \infty)\). Indeed, given \( 0 < x < y < \infty \), we have \( f[A] \) is totally positive, for \( A = \begin{pmatrix} y & x \\ x & x \end{pmatrix} \). This implies \( f(y) > f(x) > 0 \).

**Step 2:** By the previous step, \( f : (0, \infty) \to (0, \infty) \) has at most countably many discontinuities (and they are all jump discontinuities). Let \( f^+(x) := \lim_{y \to x^+} f(y) \), for \( x > 0 \). Then \( f^+(x) \geq f(x) \forall x \), and \( f^+ \) coincides with \( f \) at all points of right continuity and has the same jumps as \( f \). Thus, it suffices to show that \( f^+ \) is continuous (since this implies \( f \) is also continuous).

**Step 3:** Apply \( f[-] \) to the matrices

\[
M(x, y, \epsilon) := \begin{pmatrix} x + \epsilon & \sqrt{x}y + \epsilon \\ \sqrt{x}y + \epsilon & y + \epsilon \end{pmatrix}, \quad x, y, \epsilon > 0, \ x \neq y.
\]

Note that \( M(x, y, \epsilon) \) is TP, since

\[
\det M(x, y, \epsilon) = xy + \epsilon^2 + (x + y)\epsilon - (xy + \epsilon^2 + 2\sqrt{xy}\epsilon) = (\sqrt{x} - \sqrt{y})^2 \epsilon > 0.
\]

It follows that \( \det [M(x, y, \epsilon)] > 0 \) for \( \epsilon > 0 \), i.e., \( f(x + \epsilon)f(y + \epsilon) > f(\sqrt{xy} + \epsilon)^2 \). Taking \( \epsilon \to 0^+ \), we obtain:

\[
f^+(x)f^+(y) \geq f^+(\sqrt{xy})^2, \quad \forall x, y > 0.
\]

Therefore \( f^+ \) is positive, non-decreasing and multiplicatively mid-convex on \((0, \infty)\). From the proof of Theorem 12.6(2), we conclude that \( f^+ \) is continuous on \((0, \infty)\), whence so is \( f \).

**Step 4:** We claim that \( f(x) = f(1)x^\alpha \) for all \( x > 0 \) (and some \( \alpha > 0 \)). Consider the matrices

\[
A(x, y, \epsilon) := \begin{pmatrix} x & xy \\ 1 - \epsilon & y \end{pmatrix}, \quad B(x, y, \epsilon) := \begin{pmatrix} xy & y \\ x & 1 + \epsilon \end{pmatrix}, \quad \text{where} \ x, y, \epsilon > 0.
\]

These are both TP matrices, whence so are \( f[A(x, y, \epsilon)] \) and \( f[B(x, y, \epsilon)] \). The positivity of both determinants yields:

\[
f(x)f(y) > f(xy)f(1 - \epsilon), \quad f(xy)f(1 + \epsilon) > f(x)f(y), \quad \forall x, y, \epsilon > 0.
\]

The test set of Hankel TN matrices.

Taking \( \epsilon \to 0^+ \), the continuity of \( f \) and the assumptions imply that \( \frac{f(x)}{x} \) is multiplicative, continuous, positive, non-constant, and strictly increasing on \((0, \infty)\). Hence (e.g. as in the proof of Proposition 11.5), \( f(x) = f(1)x^\alpha \) for all \( x > 0 \), where \( \alpha > 0 \) and \( f(1) > 0 \).

**Step 5:** The above steps completed the proof for \( m = n = 2 \). Now suppose more generally that \( m, n \geq 2 \). Recall by a TP completion problem (see Theorem 7.3) that every \( 2 \times 2 \) TP matrix can be completed to an \( m \times n \) TP matrix. It follows from the assumptions that \( f \) must preserve the \( 2 \times 2 \) TP matrices, and we are done by the previous steps. \( \square \)

12.4. **Totally non-negative Hankel matrices – entrywise preservers.** We have seen that if the entrywise map \( f[-] \) preserves the \( m \times n \) TP/TN matrices for \( m, n \geq 4 \), then \( f \) is either constant on \((0, \infty)\) (and \( f(0) \) equals either this constant or zero), or \( f(x) = f(1)x \) for all \( x \). In contrast, the powers \( x^\alpha \) that entrywise preserve positive semidefiniteness on \( \mathbb{P}_n((0, \infty)) \) (for fixed \( n \geq 2 \)) are \( \mathbb{Z} \geq 0 \cup [n-2, \infty) \).

This discrepancy is also supported by the fact that \( \mathbb{P}_n \) is closed under taking the Schur (or entrywise) product, but already the \( 3 \times 3 \) TN matrices are not. (Hence neither are the \( m \times n \) TN or TP matrices for \( m, n \geq 3 \), by using completions and density arguments.) For example,

\[
A := \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix}, \quad B := A^T = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}
\]

are both TN, but \( A \circ B = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix} \) has determinant \(-1\), and hence cannot be TN.

Thus, a more refined (albeit technical) question would be to isolate and work with a class of TN matrices that is a closed convex cone, and which is further closed under the Schur product. In fact such a class has already been discussed earlier: the family of Hankel TN matrices (see Corollary 4.2 above). With those results in mind, and for future use, we introduce the following notation.

**Definition 12.10.** Given an integer \( n > 0 \), let \( \mathbb{HTN}_n \) denote the \( n \times n \) Hankel TN matrices.

We will study in the next two parts of these notes, the notion of entrywise preservers of TN on \( \mathbb{HTN}_n \) for a fixed \( n \) and for all \( n \). This study turns out to be remarkably similar (and related) to the study of positivity preservers on \( \mathbb{P}_n \) – which is perhaps not surprising, given Theorem 4.1. For now, we work in the setting under current consideration: entrywise power-preservers.

**Theorem 12.11.** For \( n \geq 2 \) and \( \alpha \in \mathbb{R} \), \( x^\alpha \) entrywise preserves TN on \( \mathbb{HTN}_n \), if and only if \( \alpha \in \mathbb{Z} \geq 0 \cup [n-2, \infty) \).

In other words, \( x^\alpha \) preserves total non-negativity on \( \mathbb{HTN}_n \) if and only if it preserves positive semidefiniteness on \( \mathbb{P}_n((0, \infty)) \).

**Proof.** If \( \alpha \in \mathbb{Z} \geq 0 \cup [n-2, \infty) \) then we use Theorem 4.1 together with Theorem 9.2. Conversely, suppose \( \alpha \in (0, n-2) \setminus \mathbb{Z} \). We have previously shown that the ‘moment matrix’ \( H := \sum_{j, k=1}^{n} (1 + \epsilon x^{j+k-2})^n \) lies in \( \mathbb{HTN}_n \) for \( x, \epsilon > 0 \); but if \( x \neq 1 \) and \( \epsilon > 0 \) is small then \( H^{\circ \alpha} \not\in \mathbb{P}_n \), as shown in the proof of Theorem 9.2. (Alternately, this holds for all \( \epsilon > 0 \) by Theorem 9.9.) It follows that \( H^{\circ \alpha} \not\in \mathbb{TN}_n \). \( \square \)

Having completed the classification of entrywise functions preserving the TP/TN matrices in any fixed size, we henceforth restrict ourselves to the entrywise functions preserving positive semidefiniteness – which will henceforth be termed *positivity* – either in a fixed dimension or in all dimensions. (As mentioned above, there will be minor detours studying the related notion of entrywise preservers of HTN.)

In this section and the next, which concludes the present component of these notes (on entrywise powers), we continue to study entrywise powers preserving positivity in a fixed dimension, by refining the test set of positive semidefinite matrices. The plan for the next two sections is as follows:

1. We begin by recalling the test set $\mathcal{P}_G((0,\infty))$ associated to any graph $G$, and discussing some of the modern-day motivations in studying entrywise functions (including powers) that preserve positivity.

2. We then prove some results on general entrywise functions preserving positivity on $\mathcal{P}_G$ for arbitrary non-complete graphs. (The case of complete graphs is the subject of the remainder of these notes.) As a consequence, the powers – in fact, the functions – preserving $\mathcal{P}_G((0,\infty))$ for $G$ any tree (or collection of trees) are completely classified.

3. We show how the ‘integration trick’ of FitzGerald–Horn (see the discussion around Equations (9.5) and (9.6)) extends to help classify the entrywise powers preserving other Loewner properties, including monotonicity, and in turn, super-additivity.

4. Using these results, we classify the powers preserving $\mathcal{P}_G$ for $G$ the *almost complete graph* (i.e., the complete graph minus any one edge).

5. We then state some recent results on powers preserving $\mathcal{P}_G$ for other $G$ (all chordal graphs; cycles; bipartite graphs). This part of the notes concludes with some questions for general graphs $G$, which naturally arise from these results.

13.1. Modern-day motivations: graphical models and high-dimensional covariance estimation. As we discuss in the next part, the question of which functions preserve positivity when applied entrywise has a long history, having been studied for the better part of a century in the analysis literature. For now we explain why this question has attracted renewed attention owing to its importance in high-dimensional covariance estimation.

In modern-day scientific applications, one of the most important challenges involves understanding complex multivariate structure and dependencies. Such questions naturally arise in various domains: understanding the interactions of financial instruments, studying markers of climate parameters to understand climate patterns, and modeling gene-gene associations in cancer and cardiovascular disease, to name a few. In such applications, one works with very large random vectors $X \in \mathbb{R}^p$, and a fundamental measure of dependency that is commonly used (given a sample of vectors) is the covariance (or correlation) matrix and its inverse. Unlike traditional regimes, where the sample size $n$ far exceeds the dimension of the problem $p$ (i.e., the number of random variables in the model), these modern applications – among others – involve the reverse situation: $n \ll p$. This is due to the high cost of making, storing, and working with observations, for instance; but moreover, an immediate consequence is that the corresponding covariance matrix built out of the samples $x_1, \ldots, x_n \in \mathbb{R}^p$:

$$\hat{\Sigma} := \frac{1}{n-1} \sum_{j=1}^{n} (x_j - \bar{x})(x_j - \bar{x})^T,$$
is highly singular. (Its rank is bounded above by the sample size $n \ll p$.) This makes $\hat{\Sigma}$ a poor estimator of the true underlying covariance matrix $\Sigma$.

A second shortcoming of the sample covariance matrix has to do with zero patterns. In the underlying model, there is often additional domain-specific knowledge which leads to sparsity. In other words, certain pairs of variables are known to be independent, or conditionally independent given other variables. For instance, in probability theory one has the notion of a Markov random field, or graphical model, in which the nodes of a graph represent random variables, and edges the dependency structure between them. Or in the aforementioned climate-related application – specifically, temperature reconstruction – the temperature at one location is assumed to not influence that at another (perhaps faraway) location, at least when conditioned on the neighboring points. Such (conditional) independences are reflected in zero entries in the associated (inverse) covariance matrix. In fact, in the aforementioned applications, several models assume most of the entries ($\sim 90\%$ or more) to be zero.

However, in the observed sample covariance matrix, there is almost always some noise, as a result of which very few entries are zero. This is another of the reasons why the sample covariance is a poor estimator in modern applications.

For such reasons, it is common for statistical practitioners to regularize the sample covariance matrix (or other estimators), in order to improve its properties for a given application. Popular state-of-the-art methods involve inducing sparsity – i.e., zero entries – via convex optimization techniques that impose an $\ell^1$-penalty (since $\ell^0$-penalties are not amenable to such techniques). While these methods induce sparsity and are statistically consistent as $n, p \to \infty$, they are iterative and hence require solving computationally expensive optimization problems. In particular, they are not scalable to ultra high-dimensional data, say for $p \sim 100,000$ or $500,000$, as one often encounters in the aforementioned situations.

A recent promising alternative is to apply entrywise functions on the entries of sample covariance matrices. For example, the hard and soft thresholding functions set ‘very small’ entries to zero (operating under the assumption that these often come from noise, and do not represent the most important associations). Another popular family of functions used in applications consists of entrywise powers. Indeed, powering up the entries provides an effective way in applications to separate signal from noise.

Note that these ‘entrywise’ operations do not suffer from the same drawback of scalability, since they operate directly on the entries of the matrix, and do not involve optimization-based techniques. The key question now, is to understand when such entrywise operations preserve positive semidefiniteness. Indeed, the regularized matrix that these operations yield, must serve as a proxy for the sample covariance matrix in further statistical analyses, and hence is required to be positive semidefinite.

It is thus crucial to understand when these entrywise operations preserve positivity – and in fixed dimension, since in a given application one knows the dimension of the problem. Note that while the motivation here comes from downstream applications, the heart of the issue is very much a mathematical question involving analysis on the cone $\mathbb{P}_n$.

With these motivations, the current and last few sections deal with entrywise powers preserving positivity in fixed dimension; progress on these questions impacts applied fields. At the same time, the question of when entrywise powers and functions preserve positivity, has been studied in the mathematics literature for almost a century. Thus (looking slightly ahead), in the next two parts – i.e. the remainder of these notes – we return to the mathematical advances, both classical and recent. This includes proving some of the celebrated
13. Entrywise powers (and functions) preserving positivity: II.
Matrices with zero patterns. 65

characterization results in this area – by Schoenberg, Rudin, Loewner and Horn, and Va-
sudeva – using fairly accessible mathematical machinery.

13.2. Entrywise functions preserving positivity on $\mathbb{P}_G$ for non-complete graphs.
In this section and the next, we continue with the theme of entrywise powers and functions
preserving positivity in a fixed dimension, now under additional sparsity constraints – i.e.,
on $\mathbb{P}_G$ for a fixed graph $G$. In this section, we obtain certain necessary conditions on general
functions preserving positivity on $\mathbb{P}_G$.

As we will see in the next part, the functions preserving positive semidefiniteness on $\mathbb{P}_n$
for all $n$ (and those preserving TN on $\mathbb{HTN}_n$) for all integers $n \geq 1$ can be classified, and they
are precisely the power series with non-negative coefficients:

$$f(x) = \sum_{k=0}^{\infty} c_k x^k, \quad \text{with } c_k \geq 0 \ \forall \ k.$$  

This is a celebrated result of Schoenberg and Rudin. However, the situation is markedly
different for entrywise preservers of $\mathbb{P}_n$ for a fixed dimension $n \geq 1$:

- For $n = 1$, clearly any $f : [0, \infty) \to [0, \infty)$ works.
- For $n = 2$, the entrywise preservers of positive semidefiniteness (or of total non-
egativity) on $\mathbb{P}_2((0, \infty))$ have been classified by Vasudeva (1979): see Theorem [12.6].
- For $n \geq 3$, the problem remains open to date.

Given the open (and challenging!) nature of the problem in fixed dimension, efforts along
this direction have tended to work on refinements of the problem: either restricting the
class of entrywise functions (to e.g. power functions, or polynomials as we study later), or
restricting the class of matrices: to TP/TN matrices, to Toeplitz matrices (done by Rudin),
or Hankel TN matrices, or to matrices with rank bounded above (by Schoenberg, Rudin,
Loewner and Horn, and subsequent authors), or to matrices with a given sparsity pattern –
i.e., $\mathbb{P}_G$ for fixed $G$. It is this last approach that we focus on, in this section and the next.

Given a (finite simple) graph $G = (V,E)$, with $V = [n] = \{1, \ldots, n\}$ for some $n \geq 1,$
and a subset $0 \in I \subset \mathbb{R}$, the subset $\mathbb{P}_G(I)$ is defined to be:

$$\mathbb{P}_G(I) := \{A \in \mathbb{P}_n(I) : a_{jk} = 0 \text{ if } j \neq k \text{ and } (j,k) \notin E\}. \quad (13.1)$$

For example, when $G = A_3$ (the path graph on three nodes), $\mathbb{P}_G = \left\{ \begin{pmatrix} a & b & e \\ d & b & 0 \\ e & 0 & c \end{pmatrix} \in \mathbb{P}_3 \right\},$
and when $G = K_n$ (the complete graph on $n$ vertices), we have $\mathbb{P}_G(I) = \mathbb{P}_n(I)$.

We now study the entrywise preservers of $\mathbb{P}_G$ for a graph $G$. To begin, we extend the
notion of entrywise functions to $\mathbb{P}_G$, by acting only on the ‘unconstrained’ entries:

**Definition 13.2.** Let $0 \in I \subset \mathbb{R}$. Given a graph $G$ with vertex set $[n]$, and $f : I \to \mathbb{R}$, define

$f_G[-] : \mathbb{P}_G(I) \to \mathbb{R}^{n \times n}$ via

$$(f_G[A])_{jk} := \begin{cases} 0, & \text{if } j \neq k, \ (j,k) \notin E, \\ f(a_{jk}), & \text{otherwise}. \end{cases}$$

Here are some straightforward observations on entrywise preservers of $\mathbb{P}_G((0, \infty))$.

1. When $G$ is the empty graph, i.e., $G = (V, \emptyset)$, the functions $f$ such that $f_G[-]$ preserves
$\mathbb{P}_G$ are precisely the functions sending $[0, \infty)$ to itself.
13. Entrywise powers (and functions) preserving positivity: II.
Matrices with zero patterns.

(2) When $G$ is the disjoint union of a positive number of disconnected copies of $K_2$ and isolated nodes, $P_G$ consists of block diagonal matrices of the form $\bigoplus_{j=1}^k A_j$, where the $A_j$ are either $2 \times 2$ or $1 \times 1$ matrices (blocks), corresponding to copies of $K_2$ or isolated points respectively, and $\bigoplus$ denotes a block diagonal matrix of the form:

$$
\begin{pmatrix}
A_1 & & \\
& A_2 & \\
& & \ddots \\
& & & A_k
\end{pmatrix}.
$$

(The remaining entries are zero.) By assumption, at least one of the $A_j$ must be a $2 \times 2$ block. For such graphs, we conclude by Theorem $[12.6]$ that $f_G[-] : P_G([0, \infty)) \to P_G([0, \infty))$ if and only if $f$ is non-negative, non-decreasing, multiplicatively mid-convex, and $0 \leq f(0) \leq \lim_{x \to 0^+} f(x)$.

(3) More generally, if $G$ is a disconnected union of graphs: $G = \bigcup_{j \in J} G_j$ then $f_G[-] : P_G([0, \infty)) \to P_G([0, \infty))$ if and only if the entrywise map $f_{G_j}[-]$ preserves $P_{G_j}([0, \infty))$ for all $j$.

In light of these examples, we shall henceforth consider only connected, non-complete graphs $G$, and the functions $f$ such that $f_G[-]$ preserves $P_G([0, \infty))$. We begin with the following necessary conditions:

**Proposition 13.3.** Let $I = [0, \infty)$ and $G$ be a connected, non-complete graph. Suppose $f : I \to \mathbb{R}$ is such that $f_G[-] : P_G(I) \to P_G(I)$. Then the following statements hold:

1. $f(0) = 0$.
2. $f$ is continuous on $I$.
3. $f$ is super-additive on $I$, i.e., $f(x + y) \geq f(x) + f(y) \forall x, y \geq 0$.

**Remark 13.4.** In particular, $f_G[-] = f[-]$ for (non-)complete graphs $G$. Thus, following the proof of Proposition $[13.3]$ we use $f[-]$ in the sequel.

**Proof.** Clearly $f : I \to I$. Assume that $G$ has at least 3 nodes, since for connected graphs with two nodes, the proposition is vacuous. A small observation – made by Horn, if not earlier – reveals that there exist three nodes, which we may relabel as 1, 2, 3 without loss of generality, such that 2, 3 are adjacent to 1 but not to each other. Since $P_2(I) \to P_G(I)$ via

$$
\begin{pmatrix}
a & b \\
b & c
\end{pmatrix} \mapsto \begin{pmatrix} a & b \\
b & c \end{pmatrix} \oplus \mathbf{0}_{(|V|-2) \times (|V|-2)},
$$

it follows from Theorem $[12.6]$ that $f|_{[0,\infty)}$ is non-negative, non-decreasing, and multiplicatively mid-convex; moreover, $f|_{[0,\infty)}$ is continuous and is identically zero or never zero.

To prove (1), define

$$
B(\alpha, \beta) := 
\begin{pmatrix}
\alpha + \beta & \alpha & \beta \\
\alpha & \alpha & 0 \\
\beta & 0 & \beta
\end{pmatrix}, \quad \alpha, \beta \geq 0.
$$

Note that $B(\alpha, \beta) \oplus \mathbf{0}_{(|V|-3) \times (|V|-3)} \in P_G(I)$. Hence $f_G[B(\alpha, \beta) \oplus \mathbf{0}] \in P_G(I)$, from which we obtain:

$$
f_G[B(\alpha, \beta)] = 
\begin{pmatrix}
f(\alpha + \beta) & f(\alpha) & f(\beta) \\
f(\alpha) & f(\alpha) & 0 \\
f(\beta) & 0 & f(\beta)
\end{pmatrix} \in P_3(I), \quad \forall \alpha, \beta \geq 0. \quad (13.5)
$$
For \( \alpha = \beta = 0 \), (13.5) yields that \( \det f_G[B(0, 0)] = -f(0)^3 \geq 0 \). But since \( f \) is non-negative, it follows that \( f(0) = 0 \), proving (1).

Now if \( f |_{(0, \infty)} \equiv 0 \), the remaining assertions immediately follow. Thus, we assume in the sequel that \( f |_{(0, \infty)} \) is always positive.

To prove (2), let \( \alpha = \beta > 0 \). Then (13.5) gives:

\[
det f_G[B(\alpha, \alpha)] \geq 0 \implies f(\alpha)^2(f(2\alpha) - 2f(\alpha)) \geq 0 \implies f(2\alpha) - 2f(\alpha) \geq 0.
\]

Taking the limit as \( t \to 0^+ \), we obtain \(-f^+(0) \geq 0 \). Since \( f \) is non-negative, it follows that \( f^+(0) = 0 = f(0) \), whence \( f \) is continuous at 0. The continuity of \( f \) on \( I \) now follows from the above discussion.

Finally, to prove (3), let \( \alpha, \beta > 0 \). Invoking (13.5) and again starting with \( \det f_G[B(\alpha, \beta)] \geq 0 \), we obtain

\[
f(\alpha)f(\beta)(f(\alpha + \beta) - f(\alpha) - f(\beta)) \geq 0 \implies f(\alpha + \beta) \geq f(\alpha) + f(\beta).
\]

This shows that \( f \) is super-additive on \( (0, \infty) \); since \( f(0) = 0 \), we obtain super-additivity on all of \( I \).

\[\square\]

Proposition 13.3 is the key step in classifying all entrywise functions preserving positivity on \( \mathbb{P}_G \) for every tree \( G \). In fact, apart from the case of \( \mathbb{P}_2 = \mathbb{P}_{K_2} \), this is perhaps the only known case (i.e., family of individual graphs) for which a complete classification of the entrywise preservers of \( \mathbb{P}_G \) is available – and proved in our next result.

Recall that a tree is a connected graph in which there is a unique path between any two vertices; equivalently, where the number of edges is one less than the number of nodes; or also where there are no cycle subgraphs. For example, the graph \( A_3 \) considered above (with \( V = \{1, 2, 3\} \) and \( E = \{(1, 2), (1, 3)\} \)) is a tree.

**Theorem 13.6.** Suppose \( I = [0, \infty) \) and a function \( f : I \to I \). Let \( G \) be a tree on at least 3 vertices. Then the following are equivalent:

1. \( f[\cdot] : \mathbb{P}_G(I) \to \mathbb{P}_G(I) \).
2. \( f[\cdot] : \mathbb{P}_T(I) \to \mathbb{P}_T(I) \) for all trees \( T \).
3. \( f[\cdot] : \mathbb{P}_{A_3}(I) \to \mathbb{P}_{A_3}(I) \).
4. \( f \) is multiplicatively mid-convex and super-additive on \( I \).

**Proof.** Note that \( G \) contains three vertices on which the induced subgraph is \( A_3 \) (consider any induced connected subgraph on three vertices). By padding \( \mathbb{P}_{A_3} \) by zeros to embed inside \( \mathbb{P}_{|G|} \), we obtain (1) \( \implies \) (3). Moreover, that (2) \( \implies \) (1) is clear.

To prove that (3) \( \implies \) (4), note that \( K_2 \hookrightarrow A_3 \). Hence, \( f \) is multiplicatively mid-convex on \( (0, \infty) \) by Theorem 12.6. By Proposition 13.3, \( f(0) = 0 \) and \( f \) is super-additive on \( I \). In particular, \( f \) is also multiplicatively mid-convex on all of \( I \).

Finally, we show that (4) \( \implies \) (2) by induction on \( n \) for all trees \( T \) with at least \( n \geq 2 \) nodes. For the case \( n = 2 \) by Theorem 12.6, it suffices to show that \( f \) is non-decreasing. Given \( \gamma \geq \alpha \geq 0 \), let \( \beta \geq 0 \) be such that \( \gamma = \alpha + \beta \). Then we have

\[
f(\gamma) = f(\alpha + \beta) \geq f(\alpha) + f(\beta) \geq f(\alpha),
\]

proving the result.

For the induction step, suppose that (2) holds for all trees on \( n \) nodes and let \( G' = (V, E) \) be a tree on \( n + 1 \) nodes. Without loss of generality, let \( V = [n + 1] = \{1, \ldots, n + 1\} \), such that node \( n + 1 \) is adjacent only to node \( n \). (Note: there always exists such a node in every
Moreover, by multiplicative mid-convexity, we obtain that $f$ is non-negative and super-additive, $f(0) = f(0 + 0) \geq 2f(0)$, whence $f(0) = 0$. If $f \equiv 0$, we are done. Thus, we assume that $f \neq 0$, whence $f|_{(0,\infty)}$ is positive by Theorem 12.6.

If $c = 0$, then $b_{nn}c - b^2 \geq 0$, whence $b = 0$ and $f[A] = \begin{pmatrix} f[B] & 0_{n \times 1} \\ 0_{1 \times n} & 0 \end{pmatrix} \in \mathbb{P}_G \oplus 0_{1 \times 1}$ by the induction hypothesis. Otherwise $c > 0$, whence $f(c) > 0$. From the properties of Schur complements we obtain that $A$ is positive semidefinite (psd) if and only if $B - \frac{b^2}{c} E_{nn}$ is psd, where $E_{nn}$ is the elementary $n \times n$ matrix with $(j,k)$ entry $\delta_{j,n}\delta_{k,n}$; and similarly, $f[A]$ is psd if and only if $[f[B] - \frac{f(b)^2}{f(c)} E_{nn}]$ is psd.

By the induction hypothesis, we have that $f[B - \frac{b^2}{c} E_{nn}]$ is psd. Thus, it suffices to prove that $f[B] - \frac{f(b)^2}{f(c)} E_{nn} - f[B - \frac{b^2}{c} E_{nn}]$ is psd. Now compute:

\[
f[B] - \frac{f(b)^2}{f(c)} E_{nn} - f[B - \frac{b^2}{c} E_{nn}] = \alpha E_{nn}, \quad \text{where} \quad \alpha = f(b_{nn}) - \frac{f(b)^2}{f(c)} - f(b_{nn} - \frac{b^2}{c}).
\]

Therefore, it suffices to show that $\alpha \geq 0$. But by super-additivity, we have

\[
\alpha = f(b_{nn}) - \frac{f(b)^2}{f(c)} - f(b_{nn} - \frac{b^2}{c}) = f(b_{nn} - \frac{b^2}{c}) + f(b_{nn} - \frac{b^2}{c}) - \frac{f(b)^2}{f(c)} - f(b_{nn} - \frac{b^2}{c}) \geq f(b_{nn} - \frac{b^2}{c}) - \frac{f(b)^2}{f(c)} - f(b_{nn} - \frac{b^2}{c}) = f(b_{nn} - \frac{b^2}{c}) - \frac{f(b)^2}{f(c)}.
\]

Moreover, by multiplicative mid-convexity, we obtain that $f(b_{nn} - \frac{b^2}{c}) f(c) \geq f(b)^2$. Hence $\alpha \geq 0$ and $f[A]$ is psd, as desired. \hfill \Box

An immediate consequence is the complete classification of entrywise powers preserving positivity on $\mathbb{P}_T([0,\infty))$ for $T$ a tree.

**Corollary 13.7.** $f(x) = x^\alpha$ preserves $\mathbb{P}_T([0,\infty))$ for a tree on at least three nodes, if and only if $\alpha \geq 1$.

The proof follows from the observation that $x^\alpha$ is super-additive on $[0,\infty)$ if and only if $\alpha \geq 1$. 

14. ENTRYWISE POWERS PRESERVING POSITIVITY: III. CHORDAL AND OTHER GRAPHS.
LOEWNER MONOTONICITY AND SUPER-ADDITIVITY.

We conclude this part of the notes by studying the set of entrywise powers preserving positivity on matrices with zero patterns. Recall the closed convex cone $P_G([0, \infty))$ studied in the previous section, for a (finite simple) graph $G$. Now define

$$
\mathcal{H}_G := \{ \alpha \geq 0 : A^\alpha \in P_G([0, \infty)) \ \forall A \in P_G([0, \infty)) \},
$$

with the convention that $0^0 := 1$. Thus $\mathcal{H}_G$ is the set of entrywise, or Hadamard, powers that preserve positivity on $P_G$.

Observe that if $G \subset H$ are graphs, then $\mathcal{H}_G \supset \mathcal{H}_H$. In particular, by the FitzGerald–Horn classification in Theorem 9.2,

$$
\mathcal{H}_G \supset \mathcal{H}_{K_n} = \mathbb{Z}_{\geq 0} \cup [n-2, \infty),
$$

whenever $G$ has $n$ vertices. In particular, there is always a point $\beta \geq 0$ beyond which every real power preserves positivity on $P_G$. We are interested in the smallest such point, which leads us to the next definition (following the FitzGerald–Horn theorem 9.2 in the special case $G = K_n$):

**Definition 14.3.** The critical exponent of a graph $G$ is

$$
\alpha_G := \min \{ \beta \geq 0 : \alpha \in \mathcal{H}_G \ \forall \alpha \geq \beta \}.
$$

**Example 14.4.** We saw earlier that if $G$ is a tree (but not a disjoint union of copies of $K_2$), then $\alpha_G = 1$; and FitzGerald–Horn showed that $\alpha_{K_n} = n - 2$ for all $n \geq 2$.

In this section we are interested in closed-form expressions for $\alpha_G$ and $\mathcal{H}_G$. Not only is this a natural mathematical refinement of Theorem 9.2 but as discussed in Section 13.1 this moreover impacts applied fields, providing modern motivation to study the question. Somewhat remarkably, the above examples were the only known cases until very recently.

On a more mathematical note, we are also interested in understanding a combinatorial interpretation of the critical exponent $\alpha_G$. This is a graph invariant that arises out of positivity; it is natural to ask if it is related to previously known (combinatorial) graph invariants, and more broadly, how it relates to the geometry of the graph.

We explain in this section that there is a uniform answer for a large family of graphs, which includes complete graphs, trees, split graphs, banded graphs, cycles, bipartite graphs, and other classes; and moreover, there are no known counterexamples to this answer. Before stating the results, we remark that the question of computing $\mathcal{H}_G, \alpha_G$ for a given graph is easy to formulate, and one can carry out easy numerical simulations by running (software code) over large sets of matrices in $P_G$ (possibly chosen randomly), to better understand which powers preserve $P_G$. This naturally leads to accessible research problems for various classes of graphs: say triangle-free graphs, or graphs with small numbers of vertices. For instance, there is a graph on five vertices for which the critical exponent is not known!

Now on to the known results. We begin by computing the critical exponent $\alpha_G$ and $\mathcal{H}_G$, more generally – for a family of graphs that turns out to be crucial in understanding several other families (split, Apollonian, banded, and in fact all chordal graphs):

**Definition 14.5.** The almost complete graph $K_n^{(1)}$ is the complete graph on $n$ nodes, with one edge missing.
We will choose a specific labeling of the nodes in \( K_n^{(1)} \); note this does not affect the set \( \mathcal{H}_G \) or the threshold \( \alpha_G \). Specifically, we set the \((1,n)\) and \((n,1)\) entries to be zero, so that \( \mathbb{P}_{K_n^{(1)}} \) consists of matrices of the form 
\[
\begin{pmatrix}
\vdots & \vdots & \vdots \\
0 & \cdots & 0 \\
\vdots & \vdots & \vdots 
\end{pmatrix} \in \mathbb{P}_n.
\]
Our first goal is to prove:

**Theorem 14.6.** For all \( n \geq 2 \), we have \( \mathcal{H}_{K_n^{(1)}} = \mathcal{H}_{K_n} = \mathbb{Z}^{\geq 0} \cup [n-2, \infty) \).

14.1. **Other Loewner properties.** In order to prove Theorem 14.6, we need to understand the powers that preserve super-additivity on \( n \times n \) matrices under the positive semidefinite ordering. We now define these notions, as well as a related notion of monotonicity.

**Definition 14.7.** Let \( I \subset \mathbb{R} \) and \( n \in \mathbb{N} \). A function \( f : I \to \mathbb{R} \) is said to be

1. **Loewner monotone** on \( \mathbb{P}_n(I) \) if we have \( A \geq B \geq 0_{n \times n} \implies f(A) \geq f(B) \).
2. **Loewner super-additive** on \( \mathbb{P}_n(I) \) if \( f(A + B) \geq f(A) + f(B) \) for all \( A, B \in \mathbb{P}_n(I) \).

In these definitions, we are using the Loewner ordering (or positive semidefinite ordering) on \( n \times n \) matrices: \( A \geq B \) if \( A - B \in \mathbb{P}_n(\mathbb{R}) \).

**Remark 14.8.** A few comments to clarify these definitions are in order. First, if \( n = 1 \) then these notions both reduce to their ‘usual’ counterparts for real functions defined on \([0, \infty)\). Second, if \( f(0) \geq 0 \), then Loewner monotonicity implies Loewner positivity. Third, a Loewner monotone function differs from – in fact is the entrywise analogue of – the more commonly studied operator monotone functions, which have the same property but for the functional calculus: \( A \geq B \geq 0 \implies f(A) \geq f(B) \geq 0 \).

Note that if \( n = 1 \) and \( f \) is differentiable, then \( f \) is monotonically increasing if and only if \( f' \) is positive. The following result generalizes this fact to powers acting entrywise on \( \mathbb{P}_n \), and classifies the Loewner monotone powers.

**Theorem 14.9** (FitzGerald–Horn). *Given an integer \( n \geq 2 \) and a scalar \( \alpha \in \mathbb{R} \), the power \( x^\alpha \) is Loewner monotone on \( \mathbb{P}_n([0, \infty)) \) if and only if \( \alpha \in \mathbb{Z}^{\geq 0} \cup [n-1, \infty) \).*

**Proof.** The proof strategy is similar to that of Theorem 9.2: use the Schur product theorem for non-negative integer powers, induction on \( n \) for the powers above the critical exponent, and (the same) rank two Hankel moment matrix counterexample for the remaining powers. First, if \( \alpha \in \mathbb{N} \) and \( 0 \leq B \leq A \), then repeated application of the Schur product theorem yields:

\[
0_{n \times n} \leq B^\alpha \leq B^{\alpha(\alpha-1)} \circ A \leq B^{\alpha(\alpha-2)} \circ A^2 \leq \cdots \leq A^\alpha.
\]

Now suppose \( \alpha \geq n - 1 \). We prove that \( x^\alpha \) is Loewner monotone on \( \mathbb{P}_n \) by induction on \( n \); the base case of \( n = 1 \) is clear. For the induction step, if \( \alpha \geq n - 1 \), then recall the integration trick 9.6 of FitzGerald and Horn:

\[
A^\alpha - B^\alpha = \alpha \int_0^1 (A - B) \circ (\lambda A + (1-\lambda)B)^{\alpha(\alpha-1)} \, d\lambda.
\]

Since \( \alpha - 1 \geq n - 2 \), the matrix \( (\lambda A + (1-\lambda)B)^{\alpha(\alpha-1)} \) is positive semidefinite by Theorem 9.2 and thus \( A^\alpha - B^\alpha \in \mathbb{P}_n \). Therefore \( A^\alpha \geq B^\alpha \), and we are done by induction.

Finally, to show that the threshold \( \alpha = n - 1 \) is sharp, suppose \( \alpha \in (0, n-1) \setminus \mathbb{N} \) (we leave the case of \( \alpha < 0 \) as an easy exercise). We will again consider the Hankel moment matrices

\[
A(\epsilon) := H_\mu \text{ for } \mu = \delta_1 + \epsilon \delta_2, \quad B := A(0) = 1_{n \times n},
\]

Loewner monotonicity and super-additivity.

where $x, \epsilon > 0$, $x \neq 1$, and $H_\mu$ is understood to denote the leading principal $n \times n$ submatrix of the Hankel moment matrix for $\mu$. Clearly $A(\epsilon) \geq B \geq 0_{n \times n}$. As above, let $v = (1, x, \ldots, x^{n-1})^T$, so that $A(\epsilon) = 11^T + \epsilon v v^T$. Choose a vector $u \in \mathbb{R}^n$ that is orthogonal to $v, v^2, \ldots, v^{[\alpha]+1}$, and $u^T v^{[\alpha]+2} = 1$. (Note, this is possible since the vectors $v, v^2, \ldots, v^{[\alpha]+2}$ are linearly independent, forming the columns of a possibly partial generalized Vandermonde matrix.)

We claim that $u^T (A(\epsilon) - B) u < 0$ for small $\epsilon > 0$, which will show that $x^\alpha$ is not Loewner monotone on $\mathbb{P}_n([0, \infty))$. Indeed, compute using the binomial series for $(1 + x)^\alpha$:

\[
u^T A(\epsilon)^\alpha u - u^T B u = u^T (11^T + \epsilon v v^T)^\alpha u - u^T 11^T u = u^T \sum_{k=0}^{[\alpha]+2} \binom{\alpha}{k} \epsilon^k v^{[\alpha]^k} T \cdot u + u^T o(\epsilon^{[\alpha]+2}) \cdot u \leq u^T \left( \binom{\alpha}{[\alpha]+2} \epsilon^{[\alpha]+2} + u^T o(\epsilon^{[\alpha]+2}) u \right),
\]

and this is negative for small $\epsilon > 0$. (Here, $o(\cdot)$ always denotes a matrix, as in (9.7) above.) $\square$

Theorem 14.9 is now used to classify the powers preserving Loewner super-additivity. Note that if $n = 1$ then $x^\alpha$ is super-additive on $\mathbb{P}_n([0, \infty)) = [0, \infty)$, if and only if $\alpha \geq n = 1$. The following result generalizes this to all integers $n \geq 1$.

**Theorem 14.10** (Guillot, Khare, and Rajaratnam). Given an integer $n \geq 2$ and a scalar $\alpha \in \mathbb{R}$, the power $x^\alpha$ is Loewner super-additive on $\mathbb{P}_n([0, \infty))$ if and only if $\alpha \in \mathbb{N} \cup [n, \infty)$. Moreover, for each $\alpha \in (0, n) \setminus \mathbb{N}$ and $x \in (0, 1)$, for $\epsilon > 0$ small enough the matrix

\[
(11^T + \epsilon v v^T)^\alpha - 11^T - (\epsilon v v^T)^\alpha
\]

is not positive semidefinite, where $v = (1, x, \ldots, x^{n-1})^T$.

Thus, once again the same rank two Hankel moment matrices provide the desired counterexamples, for non-integer powers $\alpha$ below the critical exponent.

**Proof.** As above, we leave the proof of the case $\alpha < 0$ to the reader. Next, if $\alpha = 0$ then super-additivity fails, since we always get $-11^T$ from the super-additivity condition, and this is not positive semidefinite.

If $\alpha > 0$ is an integer, then by the binomial theorem and the Schur product theorem,

\[
(A + B)^\alpha = \sum_{k=0}^{\alpha} \binom{\alpha}{k} A^k \circ B^{\alpha-k} \geq A^{\alpha} + B^{\alpha}, \quad \forall A, B \in \mathbb{P}_n.
\]

Next, if $\alpha \geq n$ and $A, B \in \mathbb{P}_n([0, \infty))$, then $x^{\alpha-1}$ preserves Loewner monotonicity on $\mathbb{P}_n$, by Theorem 14.9. Again use the integration trick (9.6) to compute:

\[
(A + B)^\alpha - A^{\alpha} = \alpha \int_0^1 B \circ (\lambda (A + B) + (1 - \lambda)A)^{\alpha-1} \ d\lambda \\
\geq \alpha \int_0^1 B \circ (\lambda B)^{\alpha-1} \ d\lambda = B^{\alpha}.
\]

The final case is if $\alpha \in (0, n) \setminus \mathbb{N}$. As above, we fix $x > 0$, $x \neq 1$, and define $v := (1, x, \ldots, x^{n-1})^T$, $A(\epsilon) := \epsilon v v^T$ ($\epsilon > 0$), $B := A(0) = 1_{n \times n}$. 


Clearly $A(\epsilon), B \geq 0_{n \times n}$. Now since $\alpha \in (0, n)$, the vectors $v, v^{\alpha}, \ldots, v^{\alpha(\alpha)}$ are linearly independent (since the matrix with these columns is part of a generalized Vandermonde matrix). Thus, we may choose $u \in \mathbb{R}^n$ that is orthogonal to $v, \ldots, v^{\alpha(\alpha)}$ (if $\alpha (0, 1)$, this is vacuous) and such that $u^T v^{\alpha} = 1$. Now compute as in the previous proof, using the binomial theorem:

$$
(A(\epsilon) + B)^{\alpha} - A(\epsilon)^{\alpha} - B^{\alpha} = \sum_{k=1}^{[\alpha]} \binom{\alpha}{k} \epsilon^k v^{\alpha k}(v^{\alpha k})^T - \epsilon^\alpha v^{\alpha}(v^{\alpha})^T + o(\epsilon^\alpha);
$$

the point here is that the last term shrinks at least as fast as $\epsilon^{[\alpha]+1}$. Hence by choice of $u$,

$$
u^T ((A(\epsilon) + B)^{\alpha} - A(\epsilon)^{\alpha} - B^{\alpha}) u = -\epsilon^\alpha + u^T \cdot o(\epsilon^\alpha) \cdot u,
$$

and this is negative for small $\epsilon > 0$. Hence $x^\alpha$ is not Loewner super-additive even on rank-one matrices in $\mathbb{P}_n((0, 1))$. □

14.2. **Entrywise powers preserving positivity** $\mathbb{P}_G$. We now apply the above results to compute the set of entrywise powers preserving positivity on $\mathbb{P}_G$ (the almost complete graph).

*Proof of Theorem 14.6.* The result is straightforward for $n = 2$, so we assume henceforth that $n \geq 3$. It suffices to show that $\mathcal{H}_{G} \subset \mathbb{Z}^{\geq 0} \cup [n - 2, \infty)$, since the reverse inclusion follows from Theorem 9.2 via (14.2). Fix $x > 0$, $x \neq 1$, and define

$$
v := (1, x, \ldots, x^{n-3})^T \in \mathbb{R}^{n-2}, \quad A(\epsilon) := \begin{pmatrix} 1 & 1^T & 0 \\ 1 & 11^T + \epsilon v v^T & \sqrt{\epsilon v} \\ 0 & \sqrt{\epsilon v} & 1 \end{pmatrix}_{n \times n}, \quad \epsilon > 0.
$$

Note that if $p, q > 0$ are scalars, $a, b \in \mathbb{R}^{n-2}$ are vectors, and $B$ is an $(n - 2) \times (n - 2)$ matrix, then using Schur complements,

$$
\begin{pmatrix} p & a^T & 0 \\ a & B & b \\ 0 & b^T & q \end{pmatrix} \in \mathbb{P}_n \iff \begin{pmatrix} p & a^T \\ a & B - q^{-1}bb^T \end{pmatrix} \in \mathbb{P}_{n-1}
$$

$$
\iff B - p^{-1}aa^T - q^{-1}bb^T \in \mathbb{P}_{n-2}. \quad (14.11)
$$

Applying this to the matrices $A(\epsilon)^{\alpha}$, we obtain:

$$
A(\epsilon)^{\alpha} \in \mathbb{P}_n \iff (11^T + \epsilon v v^T)^{\alpha} - (11^T)^{\alpha} - (\epsilon v v^T)^{\alpha} \in \mathbb{P}_{n-2}.
$$

For small $\epsilon > 0$, Theorem 14.10 now shows that $\alpha \in \mathbb{Z}^{\geq 0} \cup [n - 2, \infty)$, as desired. □

In the remainder of this section, we present what is known about the critical exponents $\alpha_G$ and power-preserver sets $\mathcal{H}_G$ for various graphs. We do not provide proofs below, instead referring the reader to D. Guillot, A. Khare, and B. Rajaratnam, *Critical exponents of graphs*, Journal of Combinatorial Theory Series A vol. 139 (2016), pp. 30–58.

The first family of graphs is that of *chordal graphs*, and it subsumes not only the complete graphs, trees, and almost complete graphs (for all of which we have computed $\mathcal{H}_G, \alpha_G$ with full proofs above), but also other graphs including split, banded, and Apollonian graphs, which are discussed presently.

**Definition 14.12.** A graph is *chordal* if it has no induced cycle of length $\geq 4$.  

Chordal graphs are important in many fields. They are also known as triangulated graphs, decomposable graphs, and rigid circuit graphs. They occur in spectral graph theory, but also in network theory, optimization, and Gaussian graphical models. Chordal graphs play a fundamental role in areas including maximum likelihood estimation in Markov random fields, perfect Gaussian elimination, and the matrix completion problem.

The following is the main result of the aforementioned 2016 paper, and it computes $\mathcal{H}_G$ for every chordal graph.

**Theorem 14.13.** Let $G$ be a chordal graph with $n \geq 2$ nodes and at least one edge. Let $r$ denote the largest integer such that $K_r$ or $K_r^{(1)} \subset G$. Then $\mathcal{H}_G = \mathbb{Z}^{\geq 0} \cup [r-2, \infty)$.

The point of the theorem is that the study of powers preserving positivity reduces solely to the geometry of the graph, and can be understood combinatorially rather than through matrix analysis (given the theorem). While we do not prove this result here, we remark that the proof crucially uses Theorem 14.6 and the ‘clique-tree decomposition’ of a chordal graph.

As applications of Theorem 14.13, we mention several examples of chordal graphs $G$ and their critical exponents $\alpha_G$; by the preceding theorem, the only powers below $\alpha_G$ that preserve positivity on $\mathbb{P}_G$ are the non-negative integers.

1. The complete and almost complete graph on $n$ vertices are chordal, and have critical exponent $n - 2$.
2. Trees are chordal, and have critical exponent 1.
3. Let $C_n$ denote a cycle graph (for $n \geq 4$), which is clearly not chordal. Any minimal planar triangulation $G$ of $C_n$ is chordal, and one can check that $\alpha_G = 2$ regardless of the size of the original cycle graph or the locations of the additional chords drawn.
4. A banded graph with bandwidth $d > 0$ is a graph with vertex set $[n] = \{1, \ldots, n\}$, and edges $(j, j + x)$ for $x \in \{-d, -d + 1, \ldots, d - 1, d\}$ such that $1 \leq j, j + x \leq n$. One checks (combinatorially) that $\alpha_G = \min(d, n - 2)$ if $n > d$.
5. A split graph consists of a clique $C \subset V$ and an independent (i.e., pairwise disconnected) set $V \setminus C$, whose nodes are connected to various nodes of $C$. Split graphs are an important class of chordal graphs because it can be shown that the proportion of (connected) chordal graphs with $n$ nodes that are split graphs, grows to 1 as $n \to \infty$. Theorem 14.13 implies that for a split graph $G$,

$$\alpha_G = \max(|C| - 2, \max \deg(V \setminus C))$$

6. Apollonian graphs are constructed as follows: start with a triangle as the first iteration. Given any iteration, which is a subdivision of the original triangle by triangles, choose an interior point of any of these ‘atomic’ triangles, and connect it to the three vertices of the corresponding atomic triangle. This increases the number of atomic triangles by 2 at each step. If $G$ is an Apollonian graph on $n \geq 3$ nodes, one shows that $\alpha_G = \min(2, n - 2)$.

It is natural to ask what is known for non-chordal graphs. In this case there are very few results (also shown in the aforementioned 2016 paper). These are now stated.

**Theorem 14.14.** Let $C_n$ denote the cycle graph on $n$ vertices (which is non-chordal for $n \geq 4$). Then $\mathcal{H}_{C_n} = [1, \infty)$ for all $n \geq 4$.

Remarkably, this is the same combinatorial recipe as for chordal graphs (in Theorem 14.13)!
vertex set $V = V_1 \sqcup V_2$ such that $V_1$ is nonempty and consists of mutually non-adjacent vertices, and the same holds for $V_2$.

**Theorem 14.15.** If $G$ is a connected bipartite graph on $n \geq 3$ nodes, then $\mathcal{H}_G = [1, \infty)$.

Once again, this is the same combinatorial recipe as for chordal and cycle graphs.

**Remark 14.16.** From the above results, one notices that for trees, cycles, bipartite graphs, and Apollonian graphs with at least 3 nodes, the critical exponent – as also the set of powers preserving positivity on $\mathbb{P}_G$ – is independent of the number of vertices (or edges), and equals a constant. (For banded graphs it is almost as good, in that $\alpha_G$ is bounded above by $d$.) On the other hand, as one increases the number of edges in a graph of $n$ vertices, the critical exponent ‘grows’ to $n - 2$, which is linear in $|V|$. Thus, one may suppose the sparsity – i.e., order of magnitude – of the edge set to influence the order of magnitude (in $n$) of the critical exponent.

However, this coarse, ad hoc principle fails to hold for bipartite graphs! Namely, suppose $n = 2m$ is even and one considers the complete bipartite graph $G = K_{m,m}$. Then $G$ has $m^2$ edges, whence matrices in $\mathbb{P}_G$ can have up to $2m^2 + 2m = (n^2/2) + n$ nonzero entries (including the diagonal entries). Said differently, matrices in $\mathbb{P}_{K_{m,m}}$ are dense, in that they can have $O(m^2)$ nonzero entries. Nevertheless, $\alpha_G = 1$. This surprising phenomenon has the following implication in applications: in regularizing covariance matrices in $\mathbb{P}_{K_{m,m}}$, it is safe to raise entries to any power $\geq 1$, and be assured that positivity stays preserved.

We end with some questions, which can be avenues for further research into this nascent topic.

**Question 14.17.**

1. Compute the critical exponent (and set of powers preserving positivity) for graphs other than the ones discussed above. In particular, compute $\alpha_G$ for all $G = (V,E)$ with $|V| \leq 5$.

2. For all graphs $G$ with known critical exponent, the critical exponent turns out to be $r - 2$, where $r$ is the largest integer such that $G$ contains either $K_r$ or $K_{r}^{(1)}$. Does the same result hold for all graphs?

3. In fact more is true in all known cases: $\mathcal{H}_G = \mathbb{Z}^{\geq 0} \cup [\alpha_G, \infty)$. Is this true for all graphs?

4. ‘Taking a step back’: can one show that the critical exponent of a graph is an integer (perhaps without computing it explicitly)?

5. Does the critical exponent have connections to – or can it be expressed in terms of – other, purely combinatorial graph invariants?
Part 3:
Entrywise functions preserving positivity in all dimensions

Part 3: Entrywise functions preserving positivity in all dimensions


In this chapter, we take a step back and explore the original results that laid the foundations of the study of entrywise preservers of positive semidefiniteness. The study of these preservers has a rich history in the analysis literature.

15.1. History of the problem. In the next few sections, we will answer the question: “Which functions, when applied entrywise, preserve positivity (positive semidefiniteness)?” (Henceforth we use the word ‘positivity’ to denote ‘positive semidefiniteness’.) This question has been the focus of a concerted effort and significant research activity over the past century. It began with the Schur product theorem (1911) and the following consequence:

Lemma 15.1 (Pólya–Szegö, 1925). Suppose a power series \( f(x) = \sum_{k=0}^{\infty} c_k x^k \) is convergent on \( I \subset \mathbb{R} \) and \( c_k \geq 0 \) for all \( k \geq 0 \). Then \( f[-] : \mathbb{P}_n(I) \to \mathbb{P}_n(\mathbb{R}) \) for all \( n \geq 1 \).

Proof. By induction and the Schur product theorem 3.8, \( f(x) = x^k \) preserves positivity on \( \mathbb{P}_n(\mathbb{R}) \) for all integers \( k \geq 0 \) and \( n \geq 1 \), and hence sends \( \mathbb{P}_n(I) \) to \( \mathbb{P}_n(\mathbb{R}) \). From this the lemma follows, using that \( \mathbb{P}_n(\mathbb{R}) \) is a closed convex cone. \( \square \)

With Lemma 15.1 in hand, Pólya and Szegö asked if there exist any other functions that preserve positivity on \( \mathbb{P}_n \) for all \( n \geq 1 \). A negative answer would essentially constitute the converse result to the Schur product theorem; and indeed, this was shown by Schur’s student Schoenberg (who is well-known for inventing ‘splines’), for continuous functions:

Theorem 15.2 (Schoenberg, 1942). Suppose \( I = [-1,1] \) and \( f : I \to \mathbb{R} \) is continuous. The following are equivalent:

1. The entrywise map \( f[-] \) preserves positivity on \( \mathbb{P}_n(I) \) for all \( n \geq 1 \).
2. The function \( f \) equals a convergent power series \( \sum_{k=0}^{\infty} c_k x^k \) for all \( x \in I \), with the Maclaurin coefficients \( c_k \geq 0 \) for all \( k \geq 0 \).

Schoenberg’s 1942 paper (in Duke Math. J.) is well-known in the analysis literature. In a sense, his Theorem 15.2 is the (harder) converse to the Schur product theorem, i.e. Lemma 15.1, which is the implication \( (2) \implies (1) \). The reader may recall that some of these points were discussed in Section 13.2.

Schoenberg’s theorem can also be stated for \( I = (-1,1) \). In this setting, the continuity hypothesis was subsequently removed from assertion (1) by Rudin, who moreover showed that in order to prove assertion (2) in Theorem 15.2, one does not need to work with the full test set \( \bigcup_{n \geq 1} \mathbb{P}_n(I) \). Instead, it is possible to work with only the low-rank Toeplitz matrices:

Theorem 15.3 (Rudin, 1959). Suppose \( I = (-1,1) \) and \( f : I \to \mathbb{R} \). Then the assertions in Schoenberg’s theorem 15.2 are equivalent on \( I \), and further equivalent to:

3. The entrywise map \( f[-] \) preserves positivity on the Toeplitz matrices in \( \mathbb{P}_n(I) \) of rank at most 3, for all \( n \geq 1 \).

Schoenberg’s theorem has been studied in other settings. A notable ‘one-sided’ variant is over the semi-axis \( I = (0, \infty) \):

Theorem 15.4 (Vasudeva, 1979). Suppose \( I = (0, \infty) \) and \( f : I \to \mathbb{R} \). Then the two assertions of Schoenberg’s theorem 15.2 are equivalent on \( I \) as well.

Our goal in this part of the notes is to prove stronger versions of the theorems of Schoenberg and Vasudeva. Specifically, we will (i) remove the continuity hypothesis, and (ii) work with severely reduced test sets in each dimension, consisting of only the Hankel matrices of rank at most 3. For instance, we will show Theorem 15.3, but with the word ‘Toeplitz’ replaced by ‘Hankel’. Similarly, we will show a strengthening of Theorem 15.4 using totally non-negative Hankel matrices of rank at most 2. These results are stated and proved in the coming sections.

15.2. Digression: the complex case. In the aforementioned 1959 paper in Duke Math. J., Rudin made some observations about the complex case a la Pólya–Szegö, and presented a conjecture. First observe that the Schur product theorem holds for complex Hermitian matrices as well, with the same proof coming from the spectral theorem:

“If $A, B$ are $n \times n$ complex (Hermitian) positive semidefinite matrices, then so is $A \circ B$.”

As a consequence, every monomial $z \mapsto z^m$ preserves positivity on $\mathbb{P}_n(\mathbb{C})$ for all integers $m \geq 0$ and $n \geq 1$. (Here $\mathbb{P}_n(\mathbb{C})$ comprises the complex Hermitian matrices $A_{n \times n}$ such that $u^*Au \geq 0$ for all $u \in \mathbb{C}^n$.) But more is true: the (entrywise) conjugation map also preserves positivity on $\mathbb{P}_n(\mathbb{C})$ for all $n \geq 1$. Now using the Schur product theorem, the functions

$$z \mapsto z^m(z)^k, \quad m, k \geq 0$$

each preserve positivity on $\mathbb{P}_n(\mathbb{C})$, for all $n \geq 1$. Since $\mathbb{P}_n(\mathbb{C})$ is easily seen to be a closed convex cone as well, Rudin observed that if a series

$$f(z) = \sum_{m, k \geq 0} c_{m,k} z^m(z)^k,$$

is convergent on a disc $D(0, \rho) \subset \mathbb{C}$, then $f[-]$ entrywise preserves positivity on $\mathbb{P}_n(D(0, \rho))$ for all $n \geq 1$. He then conjectured that there are no other such functions. This conjecture was proved soon after:

**Theorem 15.5** (Herz, Ann. Inst. Fourier, 1963). Suppose $I = D(0, 1) \subset \mathbb{C}$ and $f : I \to \mathbb{C}$. The following are equivalent:

1. The entrywise map $f[-]$ preserves positivity on $\mathbb{P}_n(I)$ for all $n \geq 1$.
2. The function $f$ is of the form $f(z) = \sum_{m, k \geq 0} c_{m,k} z^m(z)^k$ on $I$, with $c_{m,k} \geq 0$ for all $m, k \geq 0$.

As a final remark, the Schoenberg–Rudin / Vasudeva / Herz results are reminiscent of an earlier, famous result, by Loewner in the parallel (and well-studied) setting of the matrix functional calculus. Namely, given a complex Hermitian matrix $A$ with eigenvalues in a real interval $(a, b)$, a function $f : (a, b) \to \mathbb{R}$ acts on $A$ as follows: let $A = UDU^*$ be a spectral decomposition of $A$; then $f(A) := Uf(D)U^*$, where $f(D)$ is the diagonal matrix with diagonal entries $f(d_{jj})$. Now Loewner showed in Math. Z., even before Schoenberg:

**Theorem 15.6** (Loewner, 1934). Let $-\infty \leq a < b \leq \infty$, and $f : (a, b) \to \mathbb{R}$. The following are equivalent:

1. $f$ is matrix monotone: if $A \leq B$ are square matrices with eigenvalues in $(a, b)$, then $f(A) \leq f(B)$.
2. There exist real constants $p \geq 0$ and $q$, and a finite measure $\mu$ on $\mathbb{R} \setminus (a, b)$, such that

$$f(x) = q + px + \int_{\mathbb{R} \setminus (a, b)} \frac{1 + xy}{y - x} d\mu(y).$$

(3) There exists a function \( \tilde{f} \) that is analytic on \((\mathbb{C} \setminus \mathbb{R}) \cup (a, b)\), such that (a) \( f \equiv \tilde{f}(a, b) \) and (b) if \( \Im z > 0 \) then \( \Im f(z) > 0 \).

Notice similar to the preceding results, the emergence of analyticity from the dimension-free preservation of a matrix property. This is also the case with a prior result of Rudin with Helson, Kahane, and Katznelson in Acta Math. in 1958, which directly motivated Rudin’s 1959 paper discussed above.

15.3. Origins of positive matrices: Menger, Fréchet, Schoenberg, and metric geometry. In this subsection and the next two, we study some of the historical origins of positive (semi)definite matrices. This class of matrices of course arises as Hessians of twice-differentiable functions at local minima; however, the branch of early 20th century mathematics that led to the development of positivity preservers is metric geometry. More precisely, the notion of a metric space – emerging from the works of Fréchet and Hausdorff – and isometric embeddings of such structures into Euclidean and Hilbert spaces, spheres, hyperbolic and homogeneous spaces, were studied by Schoenberg, Bochner, and von Neumann among others; and it is this work that led to the study of matrix positivity and its preservation.

Definition 15.7. Recall that a metric space is a set \( X \) together with a metric \( d : X \times X \to \mathbb{R} \), satisfying:

1. Positivity: \( d(x, y) \geq 0 \) for all \( x, y \in X \), with equality if and only if \( x = y \).
2. Symmetry: \( d(x, y) = d(y, x) \) for all \( x, y \in X \).
3. Triangle inequality: \( d(x, z) \leq d(x, y) + d(y, z) \) for all \( x, y, z \in X \).

In this section, we will state and prove three results by Schoenberg, which explain his motivations in studying positivity and its preservers, and serve to illustrate the (by now well-explored) connection between metric geometry and matrix positivity. We begin with a sample result on metric embeddings, shown by Fréchet in Math. Ann. in 1910: If \( (X, d) \) is a metric space with \( |X| = n + 1 \), then \( (X, d) \) isometrically embeds into \( (\mathbb{R}^n, \ell_\infty) \).

Such results led to exploring which metric spaces isometrically embed into Euclidean spaces. Specifically, in Menger’s 1931 paper in Amer. J. Math., and Fréchet’s 1935 paper in Ann. of Math., the authors explored the following question: Given integers \( n, r \geq 1 \), characterize the tuples of \( \binom{n+1}{2} \) positive real numbers that can denote the distances between the vertices of an \((n+1)\)-simplex in \( \mathbb{R}^r \) but not in \( \mathbb{R}^{r-1} \). In other words, given a finite metric space \( X \), what is the smallest \( r \), if any, such that \( X \) isometrically embeds into \( \mathbb{R}^r \)?

In his 1935 paper in Ann. of Math., Schoenberg gave an alternate characterization of all such ‘admissible’ tuples of distances. This characterization used...matrix positivity!

Theorem 15.8 (Schoenberg, 1935). Fix integers \( n, r \geq 1 \), and a finite set \( X = \{x_0, \ldots, x_n\} \) together with a metric \( d \) on \( X \). Then \( (X, d) \) isometrically embeds into \( \mathbb{R}^r \) (with the Euclidean distance/norm) but not into \( \mathbb{R}^{r-1} \) if and only if the \( n \times n \) matrix

\[
A := (d(x_0, x_j)^2 + d(x_0, x_k)^2 - d(x_j, x_k)^2)_{j,k=1}^n
\]

is positive semidefinite of rank \( r \).

Proof. If \( (X, d) \) isometrically embeds into \( (\mathbb{R}^r, \| \cdot \|) \), then

\[
d(x_0, x_j)^2 + d(x_0, x_k)^2 - d(x_j, x_k)^2 = \|x_0 - x_j\|^2 + \|x_0 - x_k\|^2 - \|(x_0 - x_j) - (x_0 - x_k)\|^2
\]

\[
= 2(x_0 - x_j, x_0 - x_k).
\]

But then the matrix \( A \) in \((15.9)\) is the Gram matrix of a set of vectors in \( \mathbb{R}^r \), and hence is positive semidefinite. Here and below in this section, we use Theorem 2.4 and Proposition 2.14.
Consider the case when $r = R$ and their proofs) without further reference. Thus $A = B^T B$, where the columns of $B$ are $x_0 - x_j \in \mathbb{R}^r$. But then $A$ has rank at most the rank of $B$, whence at most $r$. Since $(X, d)$ does not embed in $\mathbb{R}^{r-1}$, by the same argument $A$ has rank precisely $r$.

Conversely, suppose the matrix $A$ in (15.9) is positive semidefinite of rank $r$. We first consider the case when $r = n$, i.e. $A$ is positive definite. By Theorem 2.4 $\frac{1}{2} A = B^T B$ for a square invertible matrix $B$. Thus left-multiplication by $B$ sends the $r$-simplex with vertices $0, e_1, \ldots, e_r$ to an $r$-simplex, where $e_j$ comprise the standard basis of $\mathbb{R}^r$.

We now claim that the assignment $x_0 \mapsto 0$, $x_j \mapsto Be_j$ for $1 \leq j \leq r$, is an isometry $X \to \mathbb{R}^r$ whose image, being the vertex set of an $r$-simplex, necessarily cannot embed inside $\mathbb{R}^{r-1}$. Indeed, we compute distances:

\[
d(Be_j, 0)^2 = \|B^T e_j\|^2 = \frac{1}{2} e_j^T A e_j = \frac{a_{jj}}{2} = d(x_0, x_j)^2,
\]
\[
d(Be_j, Be_k)^2 = \|Be_j - Be_k\|^2 = \frac{a_{jj} + a_{kk} - a_{jk}}{2}
\]
\[= \frac{d(x_0, x_j)^2 + d(x_0, x_k)^2 - (d(x_0, x_j)^2 + d(x_0, x_k)^2 - d(x_j, x_k)^2)}{2}.
\]

Since all distances here are non-negative, taking square roots of the above equations proves the claim (and with it, the theorem for $r = n$).

Next suppose $r < n$. Then $A = B^T PB$ for some invertible matrix $B$, where $P$ is the projection operator $\begin{pmatrix} \text{Id}_{n-r} & 0 \\ 0 & 0 \end{pmatrix}^{(n-r) \times (n-r)}$. Let $\Delta := \{PB e_j, 1 \leq j \leq n\} \cup \{0\}$ denote the projection under $P$ of the vertices of an $(n + 1)$-simplex. Repeating the above proof shows that the map $x_0 \mapsto 0$, $x_j \mapsto PB e_j$ for $1 \leq j \leq r$, is an isometry from $X$ onto $\Delta$. By construction, $\Delta$ lies in the image of the projection $P$, whence in a copy of $\mathbb{R}^r$. But being the image under $P$ of the vertex set of an $n$-simplex, $\Delta$ cannot lie in a copy of $\mathbb{R}^{r-1}$ (else so would its span, which is all of $P(\mathbb{R}^n) \cong \mathbb{R}^r$).

We end this part with an observation. A (real symmetric square) matrix $A^{(n+1) \times (n+1)}$ is said to be conditionally positive semidefinite if $u^T A u \geq 0$ whenever $\sum_j u_j = 0$. Such matrices are also studied in the literature (though not as much as positive semidefinite matrices). The following lemma reformulates Theorem 15.8 into the conditional positivity of a related matrix:

**Lemma 15.11.** Let $X = \{x_0, \ldots, x_n\}$ be a finite set equipped with a metric $d$. Then the matrix $A_{n \times n}$ as in (15.9) is positive semidefinite if and only if the $(n + 1) \times (n + 1)$ matrix

\[
A' := \begin{pmatrix} -d(x_j, x_k)^2 \end{pmatrix}_{j, k=0}^{n}
\]

is conditionally positive semidefinite.

In particular, Schoenberg’s papers in the 1930s feature both positive semidefinite matrices (Theorem 15.8 above) and conditionally positive semidefinite matrices (Theorem 15.11). Certainly the former class of matrices were a popular and recurrent theme in the analysis literature, with contributions from Carathéodory, Hausdorff, Hermite, Nevanlinna, Pick, Schur, and many others.

**Proof of Lemma 15.11.** Let $u_1, \ldots, u_n \in \mathbb{R}$ be arbitrary, and set $u_0 := -(u_1 + \cdots + u_n)$. Defining $u := (u_1, \ldots, u_n)^T$ and $u' := (u_0, \ldots, u_n)^T$, we compute using that the diagonal

15.4. Origins of positivity preservers: Schoenberg, Bochner, and positive definite functions. We continue with our historical journey, this time into the origins of the entrywise calculus on positive matrices. As Theorem 15.8 and Lemma 15.11 show, applying entrywise the function \(-x^2\) to any distance matrix \((d(x_j, x_k))_{j,k=0}^n\) from Euclidean space yields a conditionally positive semidefinite matrix \(A'\).

It is natural to want to remove the word ‘conditionally’ from the above result. Namely: which entrywise maps send distance matrices to positive semidefinite matrices? These are precisely the positive definite functions:

**Definition 15.13.** Given a metric space \((X, d)\), a function \(f : [0, \infty) \to \mathbb{R}\) is positive definite on \(X\) if for any finite set of points \(x_1, x_2, \ldots, x_n \in X\), the matrix \(f([d(x_j, x_k)])_{j,k=1}^n\) is psd.

Note here that positive definite functions are not positivity preservers, since no distance matrix is psd unless all \(x_j\) are equal (in which case we get the zero matrix). On a different note, given any metric space \((X, d)\), the positive definite functions on \(X\) form a closed convex cone, by Lemma 3.1.

In arriving at Theorem 15.2 Schoenberg was motivated by metric geometry – as we just studied – as well as the study of positive definite functions. This latter was also of interest to other mathematicians in that era: Bochner, Pólya, and von Neumann, to name a few. In fact, positive definite functions are what originally led to the development of the entrywise calculus. Note that Bochner – and previously Carathéodory, Herglotz, Mathias, and others – studied functions on groups \(G\) that were positive definite in the ‘more standard’ sense – namely, where in the above definition \(f : G \to \mathbb{C}\), and one substitutes \(d(x_j, x_k)\) by \(x_j^{-1}x_k\).


We now present – from this paper – another characterization by Schoenberg of metric embeddings into a Euclidean space \(\mathbb{R}^r\), this time via positive definite functions. This requires a preliminary observation involving the positive definiteness of an even kernel:

**Lemma 15.14.** The Gaussian kernel \(\kappa(x, y) := \exp(-||x - y||^2)\) – in other words, the function \(\exp(-x^2)\) – is positive definite on \(\mathbb{R}^r\) for all \(r \geq 1\).

*Proof.* Observe that the case of \(\mathbb{R}^r\) for general \(r\) follows from the \(r = 1\) case, via the Schur product theorem. In turn, the \(r = 1\) case is a consequence of Pólya’s Lemma 6.8 above.

The following 1938 result of Schoenberg in *Trans. Amer. Math. Soc.* relates metric space embeddings with this positive definiteness of the Gaussian kernel:

**Theorem 15.15** (Schoenberg, 1938). A finite metric space \((X, d)\) with \(X = \{x_0, \ldots, x_n\}\) embeds isometrically into \(\mathbb{R}^r\) for some \(r > 0\) (which turns out to be at most \(n\)), if and only if the \((n + 1) \times (n + 1)\) matrix with \((j, k)\) entry

\[
\exp(-\lambda^2 d(x_j, x_k)^2), \quad 0 \leq j, k \leq n
\]

entries of \(A'\) are zero:

\[
(u')^T A' u' = \sum_{k=1}^n \left( \sum_{j=1}^n u_j \right) d(x_0, x_k)^2 u_k + \sum_{j=1}^n u_j d(x_j, x_0)^2 \left( \sum_{k=1}^n u_k \right) - \sum_{j,k=1}^n u_j d(x_j, x_k)^2 u_k
\]

\[
= \sum_{j,k=1}^n u_j u_k (d(x_0, x_k)^2 + d(x_j, x_0)^2 - d(x_j, x_k)^2)
\]

\[
= u^T A u, \quad \text{for all } u \in \mathbb{R}^n.
\]

This proves the result. □

This proves the result.

In arriving at Theorem 15.2 Schoenberg was motivated by metric geometry – as we just studied – as well as the study of positive definite functions. This latter was also of interest to other mathematicians in that era: Bochner, Pólya, and von Neumann, to name a few. In fact, positive definite functions are what originally led to the development of the entrywise calculus. Note that Bochner – and previously Carathéodory, Herglotz, Mathias, and others – studied functions on groups \(G\) that were positive definite in the ‘more standard’ sense – namely, where in the above definition \(f : G \to \mathbb{C}\), and one substitutes \(d(x_j, x_k)\) by \(x_j^{-1}x_k\).


We now present – from this paper – another characterization by Schoenberg of metric embeddings into a Euclidean space \(\mathbb{R}^r\), this time via positive definite functions. This requires a preliminary observation involving the positive definiteness of an even kernel:

**Lemma 15.14.** The Gaussian kernel \(\kappa(x, y) := \exp(-||x - y||^2)\) – in other words, the function \(\exp(-x^2)\) – is positive definite on \(\mathbb{R}^r\) for all \(r \geq 1\).

*Proof.* Observe that the case of \(\mathbb{R}^r\) for general \(r\) follows from the \(r = 1\) case, via the Schur product theorem. In turn, the \(r = 1\) case is a consequence of Pólya’s Lemma 6.8 above.

The following 1938 result of Schoenberg in *Trans. Amer. Math. Soc.* relates metric space embeddings with this positive definiteness of the Gaussian kernel:

**Theorem 15.15** (Schoenberg, 1938). A finite metric space \((X, d)\) with \(X = \{x_0, \ldots, x_n\}\) embeds isometrically into \(\mathbb{R}^r\) for some \(r > 0\) (which turns out to be at most \(n\)), if and only if the \((n + 1) \times (n + 1)\) matrix with \((j, k)\) entry

\[
\exp(-\lambda^2 d(x_j, x_k)^2), \quad 0 \leq j, k \leq n
\]
is positive semidefinite, along any sequence of nonzero scalars λ converging to 0 (equivalently, for all λ ∈ ℝ).

Proof. Clearly if (X, d) embeds isometrically into ℝⁿ, then identifying the xₖ with their images in ℝⁿ, it follows by Lemma 15.14 that the matrix with (j, k) entry

\[ \exp(-\lambda^2\|x_j - x_k\|^2) = \kappa(\lambda x_j, \lambda x_k) \]

is positive semidefinite for all λ ∈ ℝ.

Conversely, let λₘ → 0 with all λₘ nonzero. From the positivity of the exponentiated distance matrices for λₘ, it follows for any vector \( u := (u_0, \ldots, u_n)^T \in \mathbb{R}^{n+1} \) that

\[ \sum_{j,k=0}^n u_j u_k \exp(-\lambda_m^2 d(x_j, x_k)^2) \geq 0. \]

Expanding into Taylor series and interchanging the infinite sum with the two finite sums,

\[ \sum_{l=0}^\infty \frac{(-\lambda_m^2)^l}{l!} \sum_{j,k=0}^n u_j u_k d(x_j, x_k)^{2l} \geq 0, \quad \forall m \geq 1. \]

Suppose we restrict to the vectors \( u' \) satisfying: \( \sum_{j=0}^n u_j = 0 \). Then the \( l = 0 \) term vanishes.

Now dividing throughout by \( \lambda_m^2 \) and taking \( m \to \infty \), the “leading term” in \( \lambda_m \) must be non-negative. It follows that if \( A' := (-d(x_j, x_k)^2)_{j,k=0}^n \), then \( (u')^T A' u' \geq 0 \) whenever \( \sum_{j=0}^n u_j = 0 \).

By Lemma 15.11 and Theorem 15.8, \( (X, d) \) embeds isometrically into \( \mathbb{R}^r \), where \( r \leq n \) denotes the rank of the matrix \( A_{n \times n} \) as in [15.5].

15.5. Schoenberg: from spheres to correlation matrices, to positivity preservers.

The previous result, Theorem 15.15, says that Euclidean spaces \( \mathbb{R}^r \) – or their direct limit / union \( \mathbb{R}^\infty \) (which should more accurately be denoted \( \mathbb{R}^\mathbb{N} \)), or even its closure \( \ell^2 \) of square-summable real sequences (which Schoenberg and others called Hilbert space) – can be characterized by the property that the maps

\[ \exp(-\lambda^2 x^2), \quad \lambda \in (0, \rho) \quad (15.16) \]

are all positive definite on each (finite) metric subspace. As we saw, such a characterization holds for each \( \rho > 0 \).

Given this characterization, it is natural to seek out similar characterizations of distinguished submanifolds \( M \) in \( \mathbb{R}^r \) or \( \mathbb{R}^\infty \) or \( \ell^2 \). In fact in the aforementioned 1935 Ann. of Math. paper, Schoenberg showed the first such classification result, for \( M = S^{r-1} \) a unit sphere – as well as for the Hilbert sphere \( S^\infty \). Note here that the unit sphere \( S^{r-1} := \{ x \in \mathbb{R}^r : \|x\|^2 = 1 \} \), while the Hilbert sphere \( S^\infty \subset \ell^2 \) is the subset of all square-summable sequences with unit \( \ell^2 \)-norm. (This is the closure of the set of all real sequences with finitely many nonzero coordinates and unit \( \ell^2 \)-norm – which is the unit sphere \( \bigcup_{r \geq 1} S^{r-1} \) in \( \bigcup_{r \geq 1} \mathbb{R}^r \).)

One defines a rotationally invariant metric on \( S^\infty \) (whence on each \( S^{r-1} \)) as follows. The distance between \( x \) and \( -x \) is \( \pi \), and given points \( x \neq \pm y \) in \( S^\infty \), there exists a unique plane passing through \( x \) and \( y \), and the origin. This plane intersects the sphere \( S^\infty \) in a unit circle \( S^1 \):

\[ \{ \alpha x + \beta y : \alpha, \beta \in \mathbb{R}, \quad 1 = \|\alpha x + \beta y\|^2 = \alpha^2 + \beta^2 + 2\alpha\beta(x, y) \}, \]

and we let \( d(x, y) \) denote the angle – i.e. arclength – between \( x \) and \( y \):

\[ d(x, y) := \angle(x, y) = \arccos(\langle x, y \rangle) \in [0, \pi]. \]
Consider Proposition 15.17 (Schoenberg, 1935). Let \((X, d)\) be a finite metric space with \(X = \{x_1, \ldots, x_n\}\). Fix an integer \(r \geq 2\). Then \(X\) isometrically embeds into \(S^{r-1}\) but not \(S^{r-2}\), if and only if \(d(x_j, x_k) \leq \pi\) for all \(1 \leq j, k \leq n\) and the matrix \((\cos d(x_j, x_k))_{j,k=1}^n\) is positive semidefinite of rank \(r\).

In particular, \(X\) embeds isometrically into the Hilbert sphere \(S^\infty\) if and only if (a) \(
\text{diam}(X) \leq \pi\) and (b) \(\cos(\cdot)\) is positive definite on \(X\).

Thus matrix positivity is also intimately connected with spherical embeddings, which may not be surprising given Theorem 15.8.

**Proof.** If there exists an embedding as claimed, then letting \(x_0 := 0 \in \mathbb{R}^r\) and identifying \(x_j\) with its image in \(S^{r-1}\), we compute as in (15.10) for \(j, k > 0\):

\[
d'(x_0, x_j)^2 + d'(x_0, x_k)^2 - d'(x_j, x_k)^2 = 2\langle x_j, x_k \rangle = 2\cos d(x_j, x_k),
\]

where \(d'\) denotes Euclidean distance in \(\mathbb{R}^r\). Hence by Schoenberg’s theorem 15.8, the matrix with these entries must be positive semidefinite of rank \(r\). Moreover, the spherical distance between \(x_j, x_k\) (for \(j, k > 0\)) is at most \(\pi\), as desired.

Conversely, since \(A := (\cos d(x_j, x_k))_{j,k=1}^n\) is positive, it is a Gram matrix of rank \(r\), whence \(A = B^TB\) for some \(B_{r \times n}\) of rank \(r\) by Theorem 2.4. Let \(y_j \in \mathbb{R}^r\) denote the columns of \(B\); then clearly \(y_j \in S^{r-1}\) \(\forall j\); moreover,

\[
\cos \langle y_j, y_k \rangle = \langle y_j, y_k \rangle = a_{jk} = \cos d(x_j, x_k), \quad \forall 1 \leq j, k \leq n.
\]

Since \(d(x_j, x_k)\) lies in \([0, \pi]\) by assumption, as does \(\langle (y_j, y_k)\rangle\), we obtain an isometry \(\varphi : X \to S^{r-1}\), sending \(x_j \mapsto y_j\) for all \(j > 0\). Finally, \(\text{im}(\varphi)\) is not contained in \(S^{r-2}\), for otherwise \(A\) would have rank at most \(r - 1\). This shows the result for \(S^{r-1}\); the case of \(S^\infty\) is similar. \(\Box\)

The proof of Proposition 15.17 shows that \(\cos(-)\) is a positive definite function on unit spheres of all dimensions.

Note that Proposition 15.17 and the preceding two theorems by Schoenberg in the 1930s

- (i) characterize metric space embeddings into Euclidean spaces via matrix positivity;
- (ii) characterize metric space embeddings into Euclidean spaces via the positive definite functions \(\exp(-\lambda^2(\cdot)^2)\) on \(\mathbb{R}^r\) or \(\mathbb{R}^\infty\) (so this involves positive matrices); and
- (iii) characterize metric space embeddings into spheres \(S^{r-1}\) or \(S^\infty\) via a different positive definite function \(\cos(\cdot)\) on \(S^\infty\).

Around the same time (in the 1930s), S. Bochner had classified all of the positive definite functions on \(\mathbb{R}\). This result was extended in 1940 simultaneously by Weil, Povzner, and Raikov to classify the positive definite functions on any locally compact abelian group. Amidst this backdrop, in his *loc. cit.* 1942 paper Schoenberg was interested in understanding the positive definite functions on a unit sphere \(S^{r-1} \subset \mathbb{R}^r\), where \(r \geq 2\).

To present Schoenberg’s result, first consider the \(r = 2\) case. As mentioned above, distance matrices (i.e., matrices of angles) are not psd; but if one applies the cosine function entrywise, then we obtain the matrix with \((j, k)\) entry \(\cos(\theta_j - \theta_k)\), and this is psd by Lemma 2.16. But now \(f[-]\) preserves positivity on a set of Toeplitz matrices (among others), by Lemma 2.16 and the subsequent discussion. For general dimension \(r \geq 2\), we have \(\cos d(x_j, x_k) = \langle x_j, x_k \rangle\) (see also the proof of Proposition 15.17), whence \(\cos d(x_j, x_k)\) always yields Gram matrices.

Hence \( f[-] \) would once again preserve positivity on a set of positive matrices. It was this class of functions that Schoenberg characterized:

**Theorem 15.18** (Schoenberg, 1942). Suppose \( f : [-1,1] \to \mathbb{R} \) is continuous, and \( r \geq 2 \) is an integer. Then the following are equivalent:

1. \( (f \circ \cos) \) is positive definite on \( S^{r-1} \).
2. The function \( f(x) = \sum_{k=0}^{\infty} c_k C_k^{(r-2)}(x) \), where \( c_k \geq 0, \forall k \), and \( \{C_k^{(r-2)}(x) : k \geq 0\} \) comprise the first Chebyshev or Gegenbauer family of orthogonal polynomials.

**Remark 15.19.** Theorem [15.18] has an interesting reformulation in terms of entrywise positivity preservers on correlation matrices. Recall that on the unit sphere \( S^{r-1} \), applying \( \cos[-] \) entrywise to a distance matrix of points \( x_j \) yields precisely the Gram matrix with entries \( \langle x_j, x_k \rangle \), which is positive of rank at most \( r \). Moreover, as the vectors \( x_j \) lie on the unit sphere, the diagonal entries are all 1 and hence we obtain a correlation matrix. Putting these facts together, \( f \circ \cos \) is positive definite on \( S^{r-1} \) if and only if \( f[-] \) preserves positivity on all correlation matrices of arbitrary size but rank at most \( r \). Thus Schoenberg’s works in 1935 and 1942 already contained connections to entrywise preservers of correlation matrices, which brings us around to the modern-day motivations that arise from precisely this question (now arising in high-dimensional covariance estimation, and discussed in Section [13.1] above).

**Remark 15.20.** Schoenberg’s work on finite- (and infinite-)dimensional spheres has many other applications. One area of recent activity involves sphere packing, spherical codes, and configurations of points on spheres that maximize the minimal distance or some potential function. See e.g. the work of Cohn with coauthors in *J. Amer. Math. Soc.* 2007, 2012 and in *Duke Math. J.* 2014, 2018; and Musin in *Ann. of Math.* 2008.

Returning to the above discussion, if instead we let \( r = \infty \), then a corresponding result would classify positivity preservers on all correlation matrices (without rank constraints) by the preceding remark. And indeed, Schoenberg achieves this goal in the same paper:

**Theorem 15.21** (Schoenberg, 1942). Suppose \( f : [-1,1] \to \mathbb{R} \) is continuous. Then \( f \circ \cos \) is positive definite on \( S^\infty \) if and only if there exist scalars \( c_k \geq 0 \) such that

\[
 f(\cos \theta) = \sum_{k \geq 0} c_k \cos^k \theta, \quad \theta \in [0, \pi].
\]

Notice here that \( \cos^k \theta \) is positive definite on \( S^\infty \) for all integers \( k \geq 0 \), by Proposition [15.17] and the Schur product theorem. Hence so is \( \sum_{k \geq 0} c_k \cos^k \theta \) if all \( c_k \geq 0 \).

Freed from the sphere context, the preceding theorem says that a continuous function \( f : [-1,1] \to \mathbb{R} \) preserves positivity when applied entrywise to all correlation matrices, if and only if \( f(x) = \sum_{k \geq 0} c_k x^k \) on \([-1,1]\) with all \( c_k \geq 0 \). This finally explains how and why Schoenberg arrived at his celebrated converse to the Schur product theorem – namely, Theorem [15.2] on entrywise positivity preservers.

15.6. Digression on ultraspherical polynomials. Before proceeding further, we describe the orthogonal polynomials \( C_k^{(\alpha)}(x) \) for \( k \geq 0 \), where \( \alpha = \alpha(r) = (r-2)/2 \). Given \( r \geq 2 \), note that \( \alpha(r) \) ranges over the non-negative half-integers. Though not used below, here are several different (equivalent) definitions of the polynomials \( C_k^{(\alpha)} \) for general real \( \alpha \geq 0 \).

First, if \( \alpha = 0 \) then \( C_k^{(0)}(x) := T_k(x) \), the Chebyshev polynomials of the first kind:

\[
 T_0(x) = 1, \quad T_1(x) = x, \quad T_2(x) = 2x^2 - 1, \quad \ldots, \quad T_k(\cos(\theta)) = \cos(k\theta) \forall k \geq 0.
\]

A second way to compute the polynomials $C_k^{(0)}(x)$ is through their generating function:

$$\frac{1 - xt}{1 - 2xt + t^2} = \sum_{k=0}^{\infty} C_k^{(0)}(x) t^k.$$ 

For higher $\alpha$: setting $\alpha = \frac{1}{2}$ yields the family of Legendre polynomials. If $\alpha = 1$, we obtain the Chebyshev polynomials of the second kind. For general $\alpha > 0$, the functions $(C_k^{(\alpha)}(x))_{k \geq 0}$ are the Gegenbauer/ultraspherical polynomials, defined via their generating function:

$$(1 - 2xt + t^2)^{-\alpha} = \sum_{k=0}^{\infty} C_k^{(\alpha)}(x) t^k.$$ 

For all $\alpha \geq 0$, the polynomials $(C_k^{(\alpha)}(x))_{k \geq 0}$ form a complete orthogonal set in the Hilbert space $L^2([-1, 1], w_\alpha)$, where $w_\alpha$ is the weight function

$$w_\alpha(x) := (1 - x^2)^{\alpha - \frac{1}{2}}, \quad x \in (-1, 1).$$

Thus, another definition of $C_k^{(\alpha)}(x)$ is that it is a polynomial of degree $k$, with $C_0^{(\alpha)}(x) = 1$, and such that the $C_k^{(\alpha)}$ are orthogonal with respect to the bilinear form

$$\langle f, g \rangle := \int_{-1}^{1} f(x)g(x)w_\alpha(x) \, dx, \quad f, g \in L^2([-1, 1], w_\alpha),$$

and satisfy:

$$\langle C_k^{(\alpha)}, C_k^{(\alpha)} \rangle = \frac{\pi 2^{1-2\alpha} \Gamma(k + 2\alpha)}{k!(\Gamma(\alpha))^2}.$$ 

Yet another definition is that the Gegenbauer polynomials $C_k^{(\alpha)}(x)$ for $\alpha > 0$ satisfy the differential equation

$$(1 - x^2)y'' - (2\alpha + 1)xy' + k(k + 2\alpha)y = 0.$$ 

We also have a direct formula

$$C_k^{(\alpha)}(x) := \frac{|k/2|}{\Gamma(\alpha)j!(k - 2j)!} (2x)^{k-2j},$$

as well as a recursion:

$$C_0^{(\alpha)}(x) := 1, \quad C_1^{(\alpha)}(x) := 2\alpha x,$$

$$C_k^{(\alpha)}(x) := \frac{1}{k} (2x(k + \alpha - 1)C_k^{(\alpha)}(x) - (k + 2\alpha - 2)C_{k-2}^{(\alpha)}(x)) \forall k \geq 2.$$ 

15.7. **Sketch of proof of Theorem 15.18.** Schoenberg’s theorem 15.18 has subsequently been studied by many authors, and in a variety of settings over the years. This includes classifying the positive definite functions on different kinds of spaces: locally compact groups, spheres, and products of these. We next give a proof-sketch of this result. In what follows, we use without reference the observation that (akin to Lemma 3.1), the set of functions $f$ such that $f \circ \cos$ is positive definite on $S^{r-1}$, also forms a closed convex cone, which is moreover closed under taking entrywise products.

We first outline why $(2) \implies (1)$ in Theorem 15.18. By the above observation, it suffices to show that $C_k^{(\alpha)} \circ \cos$ is positive definite on $S^{r-1}$. The proof is by induction on $r$. For the
base case \( r = 2 \), let \( \theta_1, \theta_2, \ldots, \theta_n \in S^1 = [0, 2\pi) \). Up to sign, their distance matrix has \((i, j)\) entry \( d(\theta_i, \theta_j) = \theta_i - \theta_j \) (or a suitable translate modulo \( 2\pi \)). Now by Lemma \ref{lem:distance_matrix_positive_semidefinite} the matrix
\[
(\cos(k(\theta_i - \theta_j)))_{i,j=1}^n
\]
is positive semidefinite. But this is precisely the matrix obtained by applying \( C_k^{(0)} \circ \cos \) to the distance matrix above. This proves one implication for \( d = 2 \). The induction step (for general \( r \geq 2 \)) follows from addition formulas for \( C_k^{(\alpha)} \).

For the converse implication, set \( \alpha := (r - 2)/2 \) and note that \( f \in L^2([-1, 1], w_\alpha) \). Hence \( f \) has a series expansion \( \sum_{k=0}^{\infty} c_k C_k^{(\alpha)}(x) \), with \( c_k \in \mathbb{R} \). Now recover the \( c_k \) via:
\[
c_k = \int_{-1}^{1} f(x) C_k^{(\alpha)}(x) w_\alpha(x) \, dx,
\]
since the \( C_k^{(\alpha)} \) form an orthonormal family. Note that \( C_k^{(\alpha)} \) and \( f \) are both positive definite (upon pre-composing with the cosine function), whence so is their product by the Schur product theorem. A result of W.H. Young now shows that \( c_k \geq 0 \) for all \( k \geq 0 \).

15.8. **Entrywise preservers in fixed dimension.** We conclude by discussing a natural mathematical refinement of Schoenberg’s theorem:

"Which functions entrywise preserve positivity in fixed dimension?"

This turns out to be a challenging, yet important question from the point of view of applications (see Section \ref{section:entrywise} for more on this.) In particular, note that there exist functions which preserve positivity on \( \mathbb{P}_n \), but not on \( \mathbb{P}_{n+1} \): the power functions \( x^\alpha \) with \( \alpha \in (n-3, n-2) \) for \( n \geq 3 \), by Theorem \ref{thm:power_functions}. By Vasudeva’s theorem \ref{thm:vasudeva} it follows that these ‘non-integer’ power functions cannot be absolutely monotonic.

Surprisingly, though Schoenberg’s theorem is classical and provides a complete description in the dimension-free case, not much is known about the fixed-dimension case: namely, the classification of functions \( f : I \rightarrow \mathbb{R} \) such that \( f[-] : \mathbb{P}_n(I) \rightarrow \mathbb{P}_n(\mathbb{R}) \) for a fixed integer \( n \geq 1 \).

- If \( n = 1 \), then clearly, any function \( f : [0, \infty) \rightarrow [0, \infty) \) works.
- For \( n = 2 \) and \( I = (0, \infty) \), these are precisely the functions \( f : (0, \infty) \rightarrow \mathbb{R} \) that are non-negative, non-decreasing, and multiplicatively mid-convex. This was shown by Vasudeva (see Theorem \ref{thm:vasudeva}), and it implies similar results for \( I = [0, \infty) \) and \( I = \mathbb{R} \).
- For every integer \( n \geq 3 \), the question is *open* to date.

Given the scarcity of results in this direction, a promising line of attack has been to study refinements of the problem. These can involve restricting the test set of matrices in fixed dimension (say under rank or sparsity constraints) or the test set of functions (say to only the entrywise powers) as was studied in the previous sections; or to use both restrictions. See Section \ref{section:refinements} for more on this discussion, as well as the next and final chapter of these notes, where we study polynomial preservers in a fixed dimension.

To conclude: while the general problem in fixed dimension \( n \geq 3 \) is open to date, there is a known result: a necessary condition satisfied by positivity preservers on \( \mathbb{P}_n \), shown by R.A. Horn in his 1969 paper in *Trans. Amer. Math. Soc.* and attributed to his advisor, Loewner. The result is almost fifty years old; yet even today, it remains essentially the only known result for general preservers \( f \) on \( \mathbb{P}_n \). In the next two sections, we will state and prove this result – in fact, a stronger version. We will then show (stronger versions of) Vasudeva’s and Schoenberg’s theorems, via a different approach than the one by Schoenberg, Rudin, or others: we crucially use the fixed-dimension theory, via the result of Horn and Loewner.

As mentioned in the previous section, our goal in this chapter is to prove a stronger form of Schoenberg’s theorem \[15.2\] in the spirit of Rudin’s theorem \[15.3\] but replacing the word ‘Toeplitz’ by ‘Hankel’. In order to do so, we will first prove a stronger version of Vasudeva’s theorem \[15.4\] in which the test set is once again reduced to only low-rank Hankel matrices.

In turn, our proof of this version of Vasudeva’s theorem relies on a fixed-dimension result, alluded to at the end of the previous section. Namely, we will state and prove a stronger version of a 1969 theorem of Horn (attributed by him to Loewner), in this section and the next.

**Theorem 16.1** (Horn–Loewner theorem, stronger version). Let \( I = (0, \infty) \) and an integer \( n \geq 1 \). Define \( u := (1, u_0, \ldots, u_{n-1})^T \). Suppose \( f : I \to \mathbb{R} \) is such that \( f^{(-)} \) preserves positivity on \( P_2(I) \) and on the set \( \{ a_{1n \times n} + buu^T : a, b > 0 \} \). Then:

1. \( f \in C^{n-3}(I) \) and \( f, f', \ldots, f^{(n-3)} \) are non-negative on \( I \). Moreover, \( f^{(n-3)} \) is convex and non-decreasing on \( I \).
2. If moreover \( f \in C^{n-1}(I) \), then \( f^{(n-2)}, f^{(n-1)} \) are also non-negative on \( I \).

Notice that the test set consists entirely of Hankel matrices of rank at most 2 – and these are moreover totally non-negative by Corollary \[4.3\] since they arise as the truncated moment matrices of the measures \( a \delta_1 + b \delta_{u_0} \). This fits in with our future use of this result to prove stronger versions of Vasudeva’s and Schoenberg’s theorems, with similarly reduced test sets of low-rank Hankel matrices.

**Remark 16.2.** In the original result by Horn (and Loewner), \( f \) was assumed to be continuous and to preserve positivity on all of \( P_2((0, \infty)) \). In Theorem \[16.1\] we have removed the continuity hypothesis, in the spirit of Rudin’s work, and also greatly reduced the test set.

**Remark 16.3.** We also observe that Theorem \[16.1\] is ‘best possible’, in that the number of nonzero derivatives that must be positive is sharp. For example, let \( n \geq 2, I = (0, \infty) \), and \( f : I \to \mathbb{R} \) be given by: \( f(x) := x^{\alpha} \), where \( \alpha \in (n-2, n-1) \). Using Theorem \[9.2\], \( f^{(-)} \) preserves positivity on the given test set \( P_2(I) \) and \( \{ a_{1n \times n} + buu^T : a, b > 0 \} \). Moreover, \( f \in C^{n-1}(I) \) and \( f, f', \ldots, f^{(n-1)} \) are strictly positive on \( I \). However, \( f^{(n)} \) is negative on \( I \).

This low-rank Hankel example (and more generally, Theorem \[9.2\]) also shows that there exist (power) functions that preserve positivity on \( P_n \) but not on \( P_{n+1} \). In the final chapter, we will show that this also holds for polynomial preservers in the fixed dimension case.

We now proceed toward the proof of Theorem \[16.1\] for general functions. A major piece of the proof is the following preliminary calculation, which will essentially prove the result for smooth functions.

**Proposition 16.4.** Fix an integer \( n > 0 \) and define \( N := \binom{n}{2} \). Suppose \( a \in \mathbb{R} \) and let a function \( f : (a - \epsilon, a + \epsilon) \to \mathbb{R} \) be \( N \)-times differentiable for some fixed \( \epsilon > 0 \). Now fix vectors \( u, v \in \mathbb{R}^n \), and define \( \Delta : (-\epsilon', \epsilon') \to \mathbb{R} \) via:

\[
\Delta(t) := \det f[a1_{n \times n} + tuv^T],
\]

for a sufficiently small \( \epsilon' \in (0, \epsilon) \). Then \( \Delta(0) = \Delta'(0) = \cdots = \Delta^{(N-1)}(0) = 0 \), and

\[
\Delta^{(N)}(0) = \left( \begin{array}{c} N \\ 0, 1, \ldots, n - 1 \end{array} \right) V(u)V(v) \prod_{k=0}^{n-1} f^{(k)}(a),
\]

\( \Delta(t) := \det f[a1_{n \times n} + tuv^T], \)

for a sufficiently small \( \epsilon' \in (0, \epsilon) \). Then \( \Delta(0) = \Delta'(0) = \cdots = \Delta^{(N-1)}(0) = 0 \), and

\[
\Delta^{(N)}(0) = \left( \begin{array}{c} N \\ 0, 1, \ldots, n - 1 \end{array} \right) V(u)V(v) \prod_{k=0}^{n-1} f^{(k)}(a),
\]
where the first factor on the right is the multinomial coefficient, and given a vector \( u = (u_1, \ldots, u_n)^T \), we define its Vandermonde determinant to be 1 if \( n = 1 \), and

\[
V(u) := \prod_{1 \leq j < k \leq n} (u_k - u_j) = \det \begin{pmatrix}
1 & u_1 & \cdots & u_1^{n-1} \\
1 & u_2 & \cdots & u_2^{n-1} \\
\vdots & \vdots & \ddots & \vdots \\
1 & u_n & \cdots & u_n^{n-1}
\end{pmatrix}, \quad \text{if } n > 1. \tag{16.5}
\]

While the result seemingly involves (higher) derivatives, it is in fact a completely algebraic phenomenon, valid over any ground ring. For the interested reader, we isolate this phenomenon in Proposition 16.8 below; its proof is more or less the same as the one now provided for Proposition 16.4. To gain some feel for the computations, the reader may wish to work out the \( N = 3 \) case first.

**Proof.** Let \( w_k \) denote the \( k \)th column of \( a 1_{n \times n} + t u v^T \); thus \( w_k \) has \( j \)th entry \( a + t u_j v_k \). To differentiate \( \Delta(t) \), we will use the multilinearity of the determinant and the Laplace expansion of \( \Delta(t) \) into a linear combination of \( n! \) ‘monomials’, each of which is a product of \( n \) terms \( f(\cdot) \). By the product rule, taking the derivative yields \( n \) terms from each monomial, and we may rearrange all of these terms into \( n \) ‘clusters’ of terms (grouping by the column which gets differentiated), and regroup back using the Laplace expansion to obtain:

\[
\Delta'(t) = \sum_{k=1}^{n} \det(f[w_1] \mid \cdots \mid f[w_{k-1}] \mid v_k u \circ f'[w_k] \mid f[w_{k+1}] \mid \cdots \mid f[w_n]).
\]

Now apply the derivative repeatedly, using this principle. By the Chain Rule, for \( m \geq 0 \) the derivative \( \Delta^{(m)}(t) \) – evaluated at \( t = 0 \) – is an integer linear combination of terms of the form

\[
\det(v_1^{m_1} u_1^{m_1} \circ f^{(m_1)}[a 1] \mid \cdots \mid v_n^{m_n} u_n^{m_n} \circ f^{(m_n)}[a 1])
= \det(f^{(m_1)}(a) v_1^{m_1} u_1^{m_1} \mid \cdots \mid f^{(m_n)}(a) v_n^{m_n} u_n^{m_n}),
\]

where \( 1 = (1, \ldots, 1)^T \in \mathbb{R}^n \) and all \( m_j \geq 0 \). Notice that if any \( m_j = m_k \) for \( j \neq k \) then the corresponding determinant (16.6) vanishes. Thus, the lowest degree derivative \( \Delta^{(m)}(0) \) whose expansion contains a non-vanishing determinant is when \( m = 0 + 1 + \cdots + (n-1) = N \). This proves the first part of the result.

To show the second part, consider \( \Delta^{(N)}(0) \). Once again, the only determinant terms that do not vanish in its expansion correspond to applying \( 0, 1, \ldots, n-1 \) derivatives to the columns in some order. We first compute the integer multiplicity of each such determinant, noting by symmetry that these multiplicities are all equal. As we are applying \( N \) derivatives to \( \Delta \) (before evaluating at \( 0 \)), the derivative applied to get \( f' \) in some column can be any of \( \binom{N}{1} \); now the two derivatives applied to get \( f'' \) in a (different) column can be chosen in \( \binom{N-1}{2} \) ways; and so on. Thus, the multiplicity is precisely

\[
\binom{N}{1} \binom{N-1}{2} \cdots \binom{N-3}{n-2} = \prod_{k=0}^{n-1} \binom{N-k}{k} = \frac{N!}{\prod_{k=0}^{n-1} k!} = \binom{N}{0, 1, \ldots, n-1}.
\]

We next compute the sum of all determinant terms. Each term corresponds to a unique permutation of the columns \( \sigma \in S_n \), with say \( \sigma_k - 1 \) the order of the derivative applied to
the $k$th column $f[w_k]$. Using $[16.6]$, the determinant corresponding to $\sigma$ equals
\[
\prod_{k=0}^{n-1} f^{(k)}(a)v_k^{\sigma_k-1} \cdot (-1)^\sigma \cdot \det(u^{0} | u^{1} | \cdots | u^{(n-1)})
\]
\[
= V(u) \prod_{k=0}^{n-1} f^{(k)}(a) \cdot (-1)^\sigma \prod_{k=0}^{n-1} v_k^{\sigma_k-1}.
\]
Summing this term over all $\sigma \in S_n$ yields precisely:
\[
V(u) \prod_{k=0}^{n-1} f^{(k)}(a) \sum_{\sigma \in S_n} (-1)^\sigma \prod_{k=0}^{n-1} v_k^{\sigma_k-1} = V(u) \prod_{k=0}^{n-1} f^{(k)}(a) \cdot V(v).
\]
Now multiplying by the (common) integer multiplicity computed above, the proof is complete. □

We next present the promised algebraic formulation of Proposition $[16.4]$ For this, some notation is required. Fix a commutative (unital) ring $R$ and an $R$-algebra $S$. The first step is to formalize the notion of the derivative, on a sub-class of $S$-valued functions. This involves more structure than the more common notion of a derivation, so we give it a different name.

**Definition 16.7.** Given a commutative ring $R$, a commutative $R$-algebra $S$ (with $R \subset S$), and an $R$-module $X$, a differential calculus is a pair $(A, \partial)$, where $A$ is an $R$-subalgebra of functions $X \to S$ (under pointwise addition and multiplication and $R$-action) which contains the constant functions, and $\partial : A \to A$ satisfies the following properties:

1. $\partial$ is $R$-linear:
   \[
   \partial \sum_j r_j f_j = \sum_j r_j \partial f_j, \quad \forall r_j \in R, \ f_j \in A, \ \forall j.
   \]
2. $\partial$ is a derivation (product rule):
   \[
   \partial(fg) = f \cdot (\partial g) + (\partial f) \cdot g, \quad \forall f, g \in A.
   \]
3. $\partial$ satisfies a variant of the ‘Chain Rule’ for composing with linear functions. Namely, if $x' \in X, r \in R$, and $f \in A$, then the function $g : X \to S, g(x) := f(x' + rx)$ also lies in $A$, and moreover,
   \[
   (\partial g)(x) = r \cdot (\partial f)(x' + rx).
   \]

With this definition in hand, we can now state our desired algebraic generalization of Proposition $[16.4]$ the proof is essentially the same.

**Proposition 16.8.** Suppose $R, S, X$ are as in Definition $[16.7]$ with an associated differential calculus $(A, \partial)$. Now fix an integer $n > 0$, two vectors $u, v \in R^n$, a vector $a \in X$, and a function $f \in A$. Define $N \in \mathbb{N}$ and $\Delta : X \to R$ via:
\[
N := \binom{n}{2}, \quad \Delta(t) := \det f[a1_{n \times n} + tuv^T], \ t \in X.
\]
Then $\Delta(0_X) = (\partial \Delta)(0_X) = \cdots = (\partial^{N-1} \Delta)(0_X) = 0_R$, and
\[
(\partial^N \Delta)(0_X) = \left(\begin{array}{c} N \\ 0,1,\ldots,n-1 \end{array}\right) V(u)V(v) \prod_{k=0}^{n-1} (\partial^k f)(a).
\]
Notice that the algebra $A$ is supposed to remind the reader of ‘smooth functions’, and is used here for ease of exposition. One can instead work with an appropriate algebraic notion of ‘$N$-times differentiable functions’ in order to “truly” generalize Proposition 16.4; we leave the details to the interested reader.

Remark 16.9. Note that Proposition 16.4 is slightly more general than the original argument of Horn and Loewner, which involved the special case $u = v$. As the above proof (as well as the preceding proposition 16.8) shows, the argument is essentially ‘algebraic’, and hence holds for arbitrary $u, v$.

Finally, we use Proposition 16.4 to prove the Horn–Loewner theorem 16.1 for smooth functions. The remainder of the proof will be discussed in the next section.

Proof of Theorem 16.1 for smooth functions. Suppose $f$ is smooth – or more generally, $C^N$ where $N = \binom{n}{2}$. We prove the result by induction on $n$. For $n = 1$ the result says that $f$ is non-negative if it preserves positivity on the given test set, which is obvious. For the induction step, we know that $f, f', \ldots, f^{(n-2)} \geq 0$ on $I$, since the given test set of $(n-1) \times (n-1)$ matrices can be embedded into the test set of $n \times n$ matrices. Now define $f_\epsilon(x) := f(x) + \epsilon x^n$ for each $\epsilon > 0$, and note by the Schur product theorem 3.8 (or Lemma 15.1) that $f_\epsilon$ also satisfies the hypotheses.

Given $a, t > 0$ and the vector $u$ as in the theorem, define $\Delta(t) := \det f_\epsilon[a_{1:n \times n} + tuu^T]$ as in Proposition 16.4 (but replacing $f, v$ by $f_\epsilon, u$ respectively). Then $\Delta(t) \geq 0$ for $t > 0$ by assumption, whence

$$0 \leq \lim_{t \to 0^+} \frac{\Delta(t)}{t^N}, \quad \text{where } N = \binom{n}{2}.$$ On the other hand, by Proposition 16.4 and L’Hôpital’s rule,

$$\lim_{t \to 0^+} \frac{\Delta(t)}{t^N} = \frac{\Delta^{(N)}(0)}{N!} = \frac{1}{N!} \binom{N}{0, 1, \ldots, n-1} V(u)^2 \prod_{k=0}^{n-1} f_\epsilon^{(k)}(a) = V(u)^2 \prod_{k=0}^{n-1} f_\epsilon^{(k)}(a) \frac{1}{k!}.$$ Thus, the right-hand side here is non-negative. Since $u$ has distinct coordinates, we can cancel all positive factors to conclude that

$$\prod_{k=0}^{n-1} f_\epsilon^{(k)}(a) \geq 0, \quad \forall \epsilon, a > 0.$$ But $f_\epsilon^{(k)}(a) = f^{(k)}(a) + \epsilon n(n-1) \cdots (n-k+1)a^{n-k}$, and this is positive for $k = 0, \ldots, n-2$ by the induction hypothesis. Hence,

$$f_\epsilon^{(n-1)}(a) = f^{(n-1)}(a) + \epsilon n! \geq 0, \quad \forall \epsilon, a > 0.$$ It follows that $f^{(n-1)}(a) \geq 0$, whence $f^{(n-1)}$ is non-negative on $(0, \infty)$, as desired. 

We conclude by mentioning that the Horn–Loewner theorem, as well as Proposition 16.4 and its algebraic avatar in Proposition 16.8 afford generalizations; the latter results reveal a surprising and novel application to Schur polynomials and to symmetric function identities. For more details, the reader is referred to the 2018 preprint Smooth entrywise positivity preservers, a Horn–Loewner master theorem, and Schur polynomial identities by this author.
17. The stronger Horn–Loewner theorem. Mollifiers.

We continue with the proof of the Horn–Loewner theorem. This is in three steps:

1. Theorem 16.1 holds for smooth functions. This was proved in the previous section.

2. If Theorem 16.1 holds for smooth functions, then it holds for continuous functions. Here, we need to assume \( n \geq 3 \).

3. If \( f \) satisfies the hypotheses in Theorem 16.1, then it is continuous. This essentially follows from Vasudeva’s 2 \( \times \) 2 result discussed earlier – see (the proof of) Theorem 12.6.

To carry out the second step – as well as a similar step in proving Schoenberg’s theorem later in these notes – we will use a standard tool in analysis called mollifiers.

17.1. An introduction to (one-variable) mollifiers. In this part we examine some basic properties of mollifiers of one variable; the theory extends to \( \mathbb{R}^n \) for all \( n > 1 \), but that is not required in what follows.

First recall that one can construct smooth functions \( g : \mathbb{R} \to \mathbb{R} \) such that \( g \) and all its derivatives vanish on \( (-\infty, 0] \); for instance, \( g(x) = \exp(-1/x) \cdot 1(x > 0) \). Indeed, one shows that \( g^{(n)}(x) = p_n(1/x)g(x) \) for some polynomial \( p_n \); hence \( g^{(n)}(x) \to 0 \) as \( x \to 0 \). Hence:

**Lemma 17.1.** Given a closed and bounded interval \([a, b] \subset \mathbb{R} \) with \(-1 < a < b < 0 \), there exists a smooth function \( \phi \) that vanishes outside \([a, b] \), is positive on \((a, b)\), and is a probability distribution on \( \mathbb{R} \).

Of course, the assumption \([a, b] \subset (-1, 0) \) is completely unused in the proof of the lemma, but is included for easy reference since we will require it in what follows.

**Proof.** The function \( \varphi(x) := g(x-a)g(b-x) \) is non-negative, smooth, and supported precisely on \((a, b)\). In particular, \( \int_{\mathbb{R}} \varphi > 0 \), so the normalization \( \phi := \varphi/\int_{\mathbb{R}} \varphi \) has the desired properties.

We now introduce mollifiers.

**Definition 17.2.** A mollifier is a one-parameter family of functions (in fact probability distributions)

\[
\{ \phi_\delta(x) := \frac{1}{\delta} \phi(x/\delta) : \delta > 0 \},
\]

with real domain and range, corresponding to any function \( \phi \) satisfying Lemma 17.1.

A continuous, real-valued function \( f \) (with suitable domain inside \( \mathbb{R} \)) is said to be mollified by convolving with the family \( \phi_\delta \). In this case, we define

\[
f_\delta(x) := \int_{\mathbb{R}} f(t) \phi_\delta(x - t) \, dt,
\]

where one extends \( f \) outside its domain by zero. (This operation is called convolution: \( f_\delta = f \ast \phi_\delta \).

Mollifiers, or Friedrichs mollifiers, were used by Horn and Loewner in the late 1960s, as well as previously by Rudin in his 1959 proof of Schoenberg’s theorem; they were a relatively modern tool at the time, having been introduced by Friedrichs in his seminal 1944 paper on PDEs in *Trans. Amer. Math. Soc.*, as well as slightly earlier by Sobolev in his famous 1938 paper in *Mat. Sbornik* (which contained the proof of the Sobolev embedding theorem).

Returning to the definition of a mollifier, notice by the change of variables \( u = x - t \) and Lemma 17.1 that

\[
f_\delta(x) = \frac{1}{\delta} \int_{\mathbb{R}} f(x - u) \phi(u/\delta) \, du = \int_{-\delta}^{0} f(x - u) \phi_\delta(u) \, du.
\]

(17.3)
In particular, \( f_\delta \) is a ‘weighted average’ of the image set \( f([x, x+\delta]) \), since \( \phi \) is a probability distribution. Now it is not hard to see that \( f_\delta \) is continuous, and converges to \( f \) pointwise as \( \delta \to 0^+ \). In fact, more is true:

**Proposition 17.4.** If \( I \subset \mathbb{R} \) is a right-open interval and \( f : I \to \mathbb{R} \) is continuous, then for all \( \delta > 0 \), the mollified functions \( f_\delta \) are smooth on \( \mathbb{R} \) (where we extend \( f \) outside \( I \) by zero), and converge uniformly to \( f \) on compact subsets of \( I \) as \( \delta \to 0^+ \).

To prove this result, we show two lemmas in somewhat greater generality. First, some notation: a (Lebesgue measurable) function \( f : \mathbb{R} \to \mathbb{R} \) is said to be **locally \( L^1 \)** if it is \( L^1 \) on each compact subset of \( \mathbb{R} \).

**Lemma 17.5.** If \( f : \mathbb{R} \to \mathbb{R} \) is locally \( L^1 \), and \( \psi : \mathbb{R} \to \mathbb{R} \) is continuous with compact support, then \( f * \psi : \mathbb{R} \to \mathbb{R} \) is also continuous.

*Proof.* Suppose \( x_n \to x \) in \( \mathbb{R} \); without loss of generality \( |x_n - x| < 1 \) for all \( n > 0 \). Also choose \( r, M > 0 \) such that \( \psi \) is supported on \([-r, r]\) and \( M = \|\psi\|_{L^\infty(\mathbb{R})} = \max_{\mathbb{R}} |\psi(x)| \). Then for each \( t \in I \), we have:

\[
(f(t) \psi(x_n - t)) \to (f(t) \psi(x - t)), \quad |f(t)\psi(x_n - t)| \leq M |f(t)| \cdot 1(|x - t| \leq r + 1)
\]

(the second inequality follows by considering separately the cases \(|x - t| \leq r + 1\) and \(|x - t| > r + 1\)). Since the right-hand side is integrable, the Dominated Convergence Theorem applies:

\[
\lim_{n \to \infty} (f * \psi)(x_n) = \lim_{n \to \infty} \int_{\mathbb{R}} f(t)\psi(x_n - t) \, dt = \int_{\mathbb{R}} \lim_{n \to \infty} f(t)\psi(x_n - t) \, dt
\]

whence \( f * \psi \) is continuous on \( \mathbb{R} \). \( \square \)

**Lemma 17.6.** If \( f : \mathbb{R} \to \mathbb{R} \) is locally \( L^1 \), and \( \psi : \mathbb{R} \to \mathbb{R} \) is \( C^1 \) with compact support, then \( f * \psi : \mathbb{R} \to \mathbb{R} \) is also \( C^1 \), and \((f * \psi)' = f * \psi'\) on \( \mathbb{R} \).

*Proof.* We compute:

\[
(f * \psi)'(x) = \lim_{h \to 0} \frac{1}{h} \int_{\mathbb{R}} f(y)\psi(x + h - y) \, dy - \frac{1}{h} \int_{\mathbb{R}} f(y)\psi(x - y) \, dy
\]

\[
= \lim_{h \to 0} \int_{\mathbb{R}} f(y) \left( \frac{\psi(x + h - y) - \psi(x - y)}{h} \right) \, dy
\]

\[
= \lim_{h \to 0} \int_{\mathbb{R}} f(y)\psi'(x - y + c(h, y)) \, dy,
\]

where for each \( y \in \mathbb{R} \), \( c(h, y) \in [0, h] \) is chosen using the Mean Value Theorem. While \( c(h, y) \to 0 \) as \( h \to 0 \), the problem is that \( y \) is not fixed inside the integral. Thus, to proceed, we argue as follows: suppose \( \psi \) is supported inside \([-r, r]\) as above, whence so is \( \psi' \). Choose any sequence \( h_n \to 0 \). We now **claim** that the last integral above, evaluated at \( h_n \), converges to \((f * \psi')(x)\) as \( n \to \infty \) — whence so does the limit of the last integral above. Indeed, we may first assume all \( h_n \in (-1, 1) \); and then check for each \( n \) that the above integral equals

\[
\int_{\mathbb{R}} f(y)\psi'(x - y + c(h_n, y)) \, dy = \int_{x-(r+1)}^{x+(r+1)} f(y)\psi'(x - y + c(h_n, y)) \, dy
\]
by choice of \( r \). Now the integrand on the right-hand side is bounded above by \( M_1 |f(y)| \) in absolute value, where \( M_1 := \|\psi''\|_{L^\infty(\mathbb{R})} \). Hence by the Dominated Convergence Theorem,

\[
\lim_{n \to \infty} \int_{x-(r+1)}^{x+(r+1)} f(y)\psi'(x-y + c(h_n,y)) \, dy = \int_{x-(r+1)}^{x+(r+1)} f(y)\psi'(x-y) \, dy = (f * \psi')(x).
\]

where the first equality also uses that \( \psi \) is \( C^1 \), and the second is by choice of \( r \). Since this happens for every sequence \( h_n \to 0 \), it follows that \( (f * \psi')(x) = (f * \psi')(x) \). Moreover, \( f * \psi' \) is continuous by Lemma \[17.5\] since \( \psi \) is \( C^1 \). This shows \( f * \psi \) is \( C^1 \) as claimed. \( \square \)

Finally, we show the claimed properties of mollified functions.

**Proof of Proposition 17.4.** Extending \( f \) by zero outside \( I \), it follows that \( f \) is locally \( L^1 \) on \( \mathbb{R} \). Repeatedly applying Lemma \[17.6\] to \( \psi = \phi_\delta, \phi_\delta', \phi_\delta'', \ldots \), we conclude that \( f_\delta \in C^\infty(\mathbb{R}) \).

To prove local uniform convergence, let \( K \) be a compact subset of \( I \) and \( \epsilon > 0 \). Denote \( b := \sup K \) and \( a := \inf K \). Since \( I \) is right-open, there is a number \( l > 0 \) such that \( J := [a, b + l] \subset I \). Since \( f \) is uniformly continuous on \( J \), given \( \epsilon > 0 \) there exists \( \delta \in (0, l) \) such that \( |x - y| < \delta \), \( x, y \in J \Rightarrow |f(x) - f(y)| < \epsilon \).

We now claim that if \( 0 < \xi < \delta \) then \( \|f_\xi - f\|_{L^\infty(K)} \leq \epsilon \); note this proves the uniform convergence of the family \( f_\delta \) to \( f \) on \( K \). To show the claim, for \( x \in K \) we compute using (17.3):

\[
|f_\xi(x) - f(x)| = \left| \int_{-\xi}^{0} (f(x-u) - f(x))\phi_\xi(u) \, du \right| 
\leq \int_{-\xi}^{0} |f(x-u) - f(x)|\phi_\xi(u) \, du 
\leq \epsilon \int_{-\xi}^{0} \phi_\xi(u) \, du = \epsilon.
\]

This is true for all \( x \in K \) by the choice of \( \xi < \delta \), and hence proves the claim. \( \square \)

### 17.2. Completing the proof of the Horn–Loewner theorem.

With mollifiers in hand, we finish the proof of Theorem 16.1. As mentioned above, the proof can be divided into three steps, and two of them are already worked out. It remains to show that if \( n \geq 3 \) and if the result holds for smooth functions, then it holds for continuous functions.

Thus, suppose \( I = (0, \infty) \) and \( f : I \to \mathbb{R} \) is continuous. Define the mollified functions \( f_\delta, \delta > 0 \) as above; thus \( f_\delta \) are smooth. Moreover, given \( a, b > 0 \), by (17.3) the function \( f_\delta \) satisfies:

\[
f_\delta[a1_{n \times n} + buu^T] = \int_{-\delta}^{0} \phi_\delta(y) \cdot f([a + |y|]1_{n \times n} + buu^T) \, dy,
\]

and this is positive semidefinite by the assumptions for \( f \). Thus \( f_\delta[-] \) preserves positivity on the given test set in \( \mathbb{P}_n(I) \); a similar argument shows that \( f_\delta[-] \) preserves positivity on \( \mathbb{P}_2(I) \). Hence by the proof in the previous section, \( f_\delta, f_\delta', \ldots, f_\delta^{(n-1)} \) are non-negative on \( I \).

Observe that the theorem amounts to deducing a similar statement for \( f \); however, as \( f \) is a priori known only to be continuous, we can only deduce non-negativity for a ‘discrete’ version of the derivatives – namely, divided differences:

**Definition 17.8.** Suppose \( I \) is a real interval and a function \( f : I \to \mathbb{R} \). Given \( h > 0 \) and an integer \( k \geq 0 \), the \( k \)th order forward differences with step size \( h > 0 \) are defined as follows:

\[
(\Delta_h^0 f)(x) := f(x), \quad (\Delta_h^k f)(x) := (\Delta_h^{k-1} f)(x+h) - (\Delta_h^{k-1} f)(x) = \sum_{j=0}^{k} \binom{k}{j} (-1)^{k-j} f(x+jh),
\]
whenever \( k > 0 \) and \( x, x + kh \in I \). Similarly, the \( k \)th order divided differences with step size \( h > 0 \) are

\[
(D_h^k f)(x) := \frac{1}{h^k} (\Delta_h^k f)(x), \quad \forall k \geq 0, \ x, x + kh \in I.
\]

The key point is that if a function is differentiable to some order, and its derivatives of that order are non-negative on an open interval, then using the mean-value theorem for divided differences, one shows the corresponding divided differences are also non-negative, whence so are the corresponding forward differences. Remarkably, the converse also holds, including differentiability! This is a classical result by Boas and Widder:

**Theorem 17.9.** Suppose \( I \subset \mathbb{R} \) is an open interval, bounded or not, and \( f : I \to \mathbb{R} \).

1. (Cauchy’s mean-value theorem for divided differences: special case.) If \( f \) is \( k \)-times differentiable in \( I \) for some integer \( k > 0 \), and \( x, x + kh \in I \) for \( h > 0 \), then there exists \( y \in (x, x + kh) \) such that \((D_h^k f)(x) = f^{(k)}(y)\).

2. (Boas–Widder, Duke Math. J., 1940.) Suppose \( k \geq 2 \) is an integer, and \( f : I \to \mathbb{R} \) is continuous and has all forward differences of order \( k \) non-negative on \( I \):

\[
(\Delta_h^k f)(x) \geq 0, \quad \text{whenever } h > 0 \text{ and } x, x + kh \in I.
\]

Then on all of \( I \), the function \( f^{(k-2)} \) exists, is continuous and convex, and has non-decreasing left and right hand derivatives.

We make a few remarks on Boas and Widder’s result. First, for \( k = 2 \) the result seems similar to Ostrowski’s theorem [12.2, except for the local boundedness being strengthened to continuity. Second, note that while \( f_{\pm}^{(k-1)} \) is non-decreasing by the theorem, one can not claim here that the lower-order derivatives \( f, \ldots, f^{(k-2)} \) are non-decreasing on \( I \). Indeed, a counterexample for such an assertion for \( f^{(0)} \), where \( 0 \leq l \leq k-2 \), is \( f(x) = -x^{l+1} \) on \( I \subset \mathbb{R} \). Finally, we refer the reader to Section 22.1 for additional related observations and results.

**Proof.** The second part will be proved in detail in Section 22. For the first, consider the Newton form of the Lagrange interpolation polynomial \( P(X) \) for \( f(X) \) at \( X = x, x+h, \ldots, x+kh \). The highest term of \( P(X) \) is

\[
(D_h^k f)(x) \cdot (X - x_{k-1}) \cdots (X - x_1)(X - x_0), \quad \text{where } x_j = x_0 + jh \ \forall j \geq 0.
\]

Writing \( g(X) := f(X) - P(X) \) to be the remainder function, note that \( g \) vanishes at \( x, x + h, \ldots, x + kh \). By successively applying Rolle’s theorem to \( g, g', \ldots, g^{(k-1)} \), it follows that \( g^{(k)} \) has a root in \( (x, x+kh) \), say \( y \). But then,

\[
0 = g^{(k)}(y) = f^{(k)}(y) - (D_h^k f)(x)k! ,
\]

which concludes the proof.

Returning to our proof of the stronger Horn–Loewner theorem [16.1, since \( f_\delta, f'_\delta, \ldots, f^{(n-1)}_\delta \geq 0 \) on \( I \), by the above theorem the divided differences of \( f_\delta \) up to order \( n - 1 \) are non-negative on \( I \), whence the same holds for the forward differences of \( f_\delta \). Applying Proposition 17.4, the forward differences of \( f \) of orders \( k = 0, \ldots, n - 1 \) are also non-negative on \( I \). Finally, invoke the Boas–Widder theorem for \( k = 2, \ldots, n - 1 \) to conclude the proof of the (stronger) Horn–Loewner theorem – noting for ‘low orders’ that \( f \) is non-negative and non-decreasing on \( I \) by using forward differences of orders \( k = 0, 1 \) respectively, whence \( f, f' \geq 0 \) on \( I \) as well. 

\[\square\]
18. The stronger Vasudeva and Schoenberg theorems. Bernstein’s theorem. 
Moment-sequence transforms.

18. The stronger Vasudeva and Schoenberg theorems. Bernstein’s theorem. 
Moment-sequence transforms.

18.1. The theorems of Vasudeva and Bernstein. Having shown (the stronger form of) 
the Horn–Loewner theorem [16.1], we use it to prove the following strengthening of Vasudeva’s 
theorem [15.4]. In it, recall from Definition [12.10] that $\mathbb{HTN}_n$ denotes the set of $n \times n$ Hankel 
totally non-negative matrices. (These are automatically positive semidefinite.)

**Theorem 18.1** (Vasudeva’s theorem, stronger version – also see Remark 18.17). Suppose $I = (0, \infty)$ and $f : I \to \mathbb{R}$. The following are equivalent:

1. The entrywise map $f[-]$ preserves positivity on $\mathbb{P}_n(I)$ for all $n \geq 1$.
2. The entrywise map $f[-]$ preserves positivity on all matrices in $\mathbb{HTN}_n$ with positive 
entries and rank at most 2, for all $n \geq 1$.
3. The function $f$ equals a convergent power series $\sum_{k=0}^{\infty} c_k x^k$ for all $x \in I$, with the 
Maclaurin coefficients $c_k \geq 0$ for all $k \geq 0$.

To show the theorem, we require the following well-known classical result by Bernstein:

**Definition 18.2.** If $I \subset \mathbb{R}$ is open, we say that $f : I \to \mathbb{R}$ is 
absolutely monotonic if $f$ is 
smooth on $I$ and $f^{(k)} \geq 0$ on $I$ for all $k \geq 0$.

**Theorem 18.3** (Bernstein). Suppose $-\infty < a < b \leq \infty$. If $f : [a,b) \to \mathbb{R}$ is continuous at 
a and absolutely monotonic on $(a,b)$, then $f$ can be extended analytically to the complex disc 
$D(a,b-a)$.

With Bernstein’s theorem in hand, the ‘stronger Vasudeva theorem’ follows easily:

**Proof of Theorem 18.1.** By the Schur product theorem or Lemma 15.1, (3) $\implies$ (1); and 
clearly (1) $\implies$ (2). Now suppose (2) holds. By the stronger Horn–Loewner theorem 16.1, 
f$^{(k)} \geq 0$ on $I$ for all $k \geq 0$, i.e., $f$ is absolutely monotonic on $I$. In particular, $f$ is non-
negative and non-decreasing on $I = (0, \infty)$, so it can be continuously extended to the origin 
via: $f(0) := \lim_{x \to 0^+} f(x) \geq 0$. Now apply Bernstein’s theorem with $a = 0$ and $b = \infty$ to 
deduce that $f$ agrees on $[0, \infty)$ with an entire function $\sum_{k=0}^{\infty} c_k x^k$. Moreover, since $f^{(k)} \geq 0$ 
on $I$ for all $k$, it follows that $f^{(k)}(0) \geq 0$, i.e., $c_k \geq 0 \forall k \geq 0$. Restricting to $I$, we obtain (3), 
as desired. □

On a related note, recall Theorems 9.2 and 12.11, which showed that when studying 
entrywise powers preserving the two closed convex cones $\mathbb{P}_n([0, \infty))$ and $\mathbb{HTN}_n$, the answers 
were identical. This is perhaps not very surprising, given Theorem 4.1. In this vein, we 
observ that such an equality of preserver sets also holds when classifying the entrywise 
maps preserving Hankel TN matrices with positive entries:

**Corollary 18.4.** With $I = (0, \infty)$ and $f : I \to \mathbb{R}$, the three assertions in Theorem 18.1 are 
further equivalent to:

4. The entrywise map $f[-]$ preserves total non-negativity on all matrices in $\mathbb{HTN}_n$ with 
positive entries, for all $n \geq 1$.
5. The entrywise map $f[-]$ preserves total non-negativity on the matrices in $\mathbb{HTN}_n$ with 
pol positive entries and rank at most 2, for all $n \geq 1$.

**Proof.** Clearly (4) $\implies$ (5) $\implies$ (2), where (1)–(3) are as in Theorem 18.1. That (3) $\implies$ 
(4) follows from Lemma 15.1 and Theorem 4.1 □
Remark 18.5. Here is one situation where the two sets of preservers – of positivity on \( \mathbb{P}_n \) for all \( n \), and of total non-negativity on \( \text{HTN}_n \) for all \( n \) – differ: if we also allow zero entries, as opposed to only positive entries as in the preceding corollary and Theorem 18.1. In this case, one shows that the preservers of \( \text{HTN}_n \) for all \( n \) are the functions \( f \) such that \( f|_{(0,\infty)} \) is absolutely monotonic, whence a power series with non-negative Maclaurin coefficients; and such that \( 0 \leq f(0) \leq \lim_{x \to 0^+} f(x) \), since the only Hankel TN matrices with a zero entry arise as truncated moment matrices of measures \( a\delta_0 \). On the other hand, by considering the rank-two Hankel positive semidefinite matrix \( A := \begin{pmatrix} 2 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \), and considering the inequality

\[
\lim_{x \to 0^+} \det f[xA] \geq 0,
\]

it follows that \( f \) is continuous at \( 0^+ \). (Note, \( A \) is not TN.) In particular, the entrywise preservers of positivity on \( \bigcup_{n \geq 1} \mathbb{P}_n([0,\infty)) \) are precisely the functions \( \sum_{k \geq 0} c_k x^k \), with all \( c_k \geq 0 \).

Remark 18.6. As a reminder, we recall that if one instead tries to classify the entrywise preservers of total non-negativity on all (possibly symmetric) TN matrices, then one obtains only the constant or linear functions \( f(x) = c, cx \) for \( c, x \geq 0 \). See Theorem 12.7 above.

To complete the proof of the stronger Vasudeva theorem 18.1 as well as its corollary above, it remains to show Bernstein’s theorem.

Proof of Bernstein’s theorem 18.3. First we claim that \( f^{(k)}(a^+) \) exists and equals \( \lim_{x \to a^+} f^{(k)}(x) \) for all \( k \geq 0 \). The latter limit here exists because \( f^{(k+1)} \geq 0 \) on \((a,b)\), so \( f^{(k)}(x) \) is non-negative and non-decreasing on \([a,b]\).

It suffices to show the claim for \( k = 1 \). But here we compute:

\[
f'(a^+) = \lim_{h \to 0^+} \frac{f(a+h) - f(a)}{h} = \lim_{h \to 0^+} f'(a + c(h)),
\]

where \( c(h) \in [0,h] \) exists and goes to zero as \( h \to 0^+ \), by the Mean Value Theorem. The claim now follows from the previous paragraph. In particular, \( f^{(k)} \) exists and is continuous, non-negative, and non-decreasing on \([a,b]\).

Applying Taylor’s theorem, we have

\[
f(x) = f(a^+) + f'(a^+)(x-a) + \cdots + f^{(n)}(a^+) \frac{(x-a)^n}{n!} + R_n(x),
\]

where \( R_n \) is the Taylor remainder:

\[
R_n(x) = \int_a^x \frac{(x-t)^n}{n!} f^{(n+1)}(t) \, dt. \tag{18.7}
\]

By the assumption on \( f \), we see that \( R_n(x) \geq 0 \). Changing variables to \( t = a + y(x-a) \), the limits for \( y \) change to \( 0,1 \), and we have:

\[
R_n(x) = \frac{(x-a)^{n+1}}{n!} \int_0^1 (1-y)^n f^{(n+1)}(a+y(x-a)) \, dy.
\]

Since \( f^{(n+2)} \geq 0 \) on \([a,b]\), if \( a \leq x \leq c \) for some \( c < b \), then uniformly in \([a,c]\) we have:

\[0 \leq f^{(n+1)}(a+y(x-a)) \leq f^{(n+1)}(a+y(c-a)).\]

Therefore using Taylor’s remainder formula
18. The stronger Vasudeva and Schoenberg theorems. Bernstein’s theorem.

Moment-sequence transforms.

Once again, we obtain:

\[ 0 \leq R_n(x) \leq \frac{(x-a)^{n+1}}{n!} \int_0^1 (1-y)^n f^{(n+1)}(a+y(c-a)) \, dy \]

\[ = R_n(c) \frac{(x-a)^{n+1}}{(c-a)^{n+1}} \]

\[ = \frac{(x-a)^{n+1}}{(c-a)^{n+1}} \left( f(c) - \sum_{k=0}^n \frac{f^{(k)}(a^+)}{k!} (c-a)^k \right) \]

\[ \leq f(c) \frac{(x-a)^{n+1}}{(c-a)^{n+1}}. \]

From this it follows that \( \lim_{n \to \infty} R_n(x) = 0 \) for all \( x \in [a, c] \). Since this holds for all \( c \in (a, b) \), the Taylor series of \( f \) converges to \( f \) on \( [a, b] \). In other words,

\[ f(x) = \sum_{k=0}^\infty \frac{f^{(k)}(a^+)}{k!} (x-a)^k, \quad x \in [a, b]. \]

Now if \( z \in D(a, b-a) \), then clearly \( a^+|z-a| < a+(b-a) = b \). Choosing any \( c \in (a+|z-a|, b) \), we check that the Taylor series converges (absolutely) at \( z \):

\[ \left| \sum_{k=0}^\infty \frac{f^{(k)}(a^+)}{k!} (z-a)^k \right| \leq \sum_{k=0}^\infty \frac{f^{(k)}(a^+)}{k!} |z-a|^k \]

\[ \leq \sum_{k=0}^\infty \frac{f^{(k)}(a^+)}{k!} |c-a|^k \]

\[ = f(c) < \infty. \]

This completes the proof of Bernstein’s theorem – and with it, the stronger form of Vasudeva’s theorem. \( \square \)

As a concluding remark for this subsection, we mention for completeness that Bernstein’s theorem was extended by Bernstein in 1926, to show that even if \( f^{(k)} \geq 0 \) in \( (a, b) \) only for \( k \geq 0 \) even, then \( f \) is necessarily analytic in \( (a, b) \). This was further extended by Boas in Duke Math. J.:

**Theorem 18.8** (Boas, 1941). Let \( \{n_p : p \geq 1\} \) be an increasing sequence of positive integers such that \( n_{p+1}/n_p \) is uniformly bounded. Let \( (a, b) \subset \mathbb{R} \) and \( f : (a, b) \to \mathbb{R} \) be smooth. If for each \( p \geq 1 \), the derivative \( f^{(n_p)} \) does not change sign in \( (a, b) \), then \( f \) is analytic in \( (a, b) \).

18.2. **The stronger version of Schoenberg’s theorem.** We now come to the main result of this part of the notes: the promised strengthening of Schoenberg’s theorem.

**Theorem 18.9** (Schoenberg’s theorem, stronger version). Given \( f : \mathbb{R} \to \mathbb{R} \), the following are equivalent:

1. The entrywise map \( f[-] \) preserves positivity on \( \mathbb{P}_n(\mathbb{R}) \) for all \( n \geq 1 \).
2. The entrywise map \( f[-] \) preserves positivity on the Hankel matrices in \( \mathbb{P}_n(\mathbb{R}) \) of rank at most 3, for all \( n \geq 1 \).
3. The function \( f \) equals a convergent power series \( \sum_{k=0}^\infty c_k x^k \) for all \( x \in \mathbb{R} \), with the Maclaurin coefficients \( c_k \geq 0 \) for all \( k \geq 0 \).

See also Theorem [18.14] for two a priori weaker, yet equivalent, assertions.
Remark 18.10. Recall the two definitions of positive definite functions, from Definition \[15.13\] and the discussion preceding Lemma \[15.14\]. The metric-space version for Euclidean and Hilbert spheres was connected by Schoenberg, for functions of the form \(f \circ \cos\), to kernels \(f: S^{r-1} \times S^{r-1}\) satisfying: \((f((x_j, x_k)))_{j,k \geq 0}\) is positive semidefinite. A third notion of positive definite kernels on Hilbert spaces \(H\) arises from here, and is important in machine learning among other areas: one says \(f: \mathbb{R} \to \mathbb{R}\) is positive definite on \(H\) if, for any choice of finitely many points \(x_j, j \geq 0\), the matrix \((f((x_j, x_k)))_{j,k \geq 0}\) is positive semidefinite. Since Gram matrix ranks are bounded above by \(\dim H\), this shows that Rudin’s 1959 theorem \[15.3\] classifies the positive definite functions on \(H\) for any real Hilbert space of dimension \(\geq 3\). The stronger Schoenberg theorem \[18.9\] above, provides a second proof.

Returning to the stronger Schoenberg theorem \[18.9\], clearly \((1) \implies (2)\), and \((3) \implies (1)\) by the Pólya–Szegő observation \[15.1\]. Thus, our goal over the next few sections is to prove \((2) \implies (3)\). The proof is simplified when some of the arguments below are formulated in the language of moment-sequences and their preservers. We begin by defining these and explaining the dictionary between moment-sequences and positive-semidefinite Hankel matrices, due to Hamburger (among others).

Definition 18.11. Recall that given an integer \(k \geq 0\) and a real measure \(\mu\) supported on a subset of \(\mathbb{R}\), \(\mu\) has \(k\)th moment equal to the following (if it converges):

\[s_k(\mu) := \int_{\mathbb{R}} x^k \, d\mu.\]

Henceforth we only work with admissible measures, i.e. such that \(\mu\) is non-negative on \(\mathbb{R}\) and \(s_k(\mu)\) converges for all \(\mu\). The moment-sequence of such a measure \(\mu\) is the sequence \(s(\mu) := (s_0(\mu), s_1(\mu), \ldots).\)

We next define transforms of moment-sequences: a function \(f: \mathbb{R} \to \mathbb{R}\) acts entrywise, to take moment sequences to real sequences:

\[f[s(\mu)] := (f(s_k(\mu)))_{k \geq 0}.\] (18.12)

We are interested in examining when the transformed sequence (18.12) is also the moment-sequence of an admissible measure supported on \(\mathbb{R}\). This connects to the question of positivity preservers via the following classical result.

Theorem 18.13 (Hamburger). A real sequence \((s_k)_{k \geq 0}\) is the moment-sequence of an admissible measure, if and only if the semi-infinite Hankel matrix \(H := (s_{j+k})_{j,k \geq 0}\) is positive semidefinite.

Recall that the easy half of this result was proved long ago, in Lemma \[2.21\].

Thus, entrywise functions preserving positivity on Hankel matrices are intimately related to moment-sequence preservers. Also note that if a measure \(\mu\) has finite support in the real line, then by examining e.g. \((2.20)\), the Hankel moment matrix \(H_\mu\) (i.e., every submatrix) has rank at most the size of the support. From this and Hamburger’s theorem, we deduce all but the last sentence of the following result:

Theorem 18.14. Theorem \[18.9\](2) implies the following a priori weaker statement:

\(4)\) For each measure

\[\mu = a\delta_1 + b\delta_{u_0} + c\delta_{-1}, \quad \text{with } u_0 \in (0, 1), \ a, b, c \geq 0,\]

there exists an admissible (non-negative) measure \(\sigma = \sigma_\mu\) on \(\mathbb{R}\) such that \(f(s_k(\mu)) = s_k(\sigma) \forall k \geq 0.\)
18. The stronger Vasudeva and Schoenberg theorems. Bernstein’s theorem.
Moment-sequence transforms.

In turn, this implies the still-weaker statement:

(5) For each measure \( \mu \) as in (18.15), with semi-infinite Hankel moment matrix \( H_\mu \), the matrix \( f[H_\mu] \) is positive semidefinite.

In fact, these statements are equivalent to the assertions in Theorem 18.9.

Remark 18.16. In these notes, we do not prove Hamburger’s theorem; but we have used it to state Theorem 18.14(4) – i.e., in working with the admissible measure \( \sigma = \sigma_\mu \). A closer look reveals that the use of Hamburger’s theorem and moment-sequences is not required to prove Schoenberg’s theorem, or even its stronger form in Theorem 18.14, which is explained in (proving) Theorem 18.14(5). Our workaround is explained in the next section, via a ‘positivity-certificate trick’ involving limiting sum-of-squares representations of polynomials. That said, moment-sequences help simplify the presentation of the proof, and hence we will continue to use them in the proof, in later sections.

The next few sections are devoted to proving (4) \( \implies \) (3) (in Theorems 18.9 and 18.14). Here is an outline of the steps in the proof:

1. All matrices \( A = \begin{pmatrix} a & b \\ b & c \end{pmatrix} \in \mathbb{P}_2((0, \infty)) \) with \( a \geq c \) occur as leading principal submatrices of the Hankel moment matrices \( H_\mu \), where \( \mu \) is as in (18.15).
2. Apply the stronger Horn–Loewner theorem and Bernstein’s theorem to deduce that \( f|_{(0, \infty)} = \sum_{k=0}^\infty c_k x^k \) for some \( c_k \geq 0 \).
3. If \( f \) satisfies assertion (5) in Theorem 18.14, then \( f \) is continuous on \( \mathbb{R} \).
4. If \( f \) is smooth and satisfies assertion (5) in Theorem 18.14, then \( f \) is real analytic.
5. Real analytic functions satisfy the desired implication above: (5) \( \implies \) (3).
6. Using mollifiers and complex analysis, one can go from smooth functions to continuous functions.

Notice that Steps 3, 4–5, and 6 resemble the three steps in the proof of the stronger Horn–Loewner theorem 16.1.

In this section, we complete the first two steps in the proof.

**Step 1:** For the first step, suppose \( A = \begin{pmatrix} a & b \\ b & c \end{pmatrix} \in \mathbb{P}_2((0, \infty)) \) with \( a \geq c \). We analyze three cases. First, if \( b = \sqrt{ac} \) then use \( \mu = a \delta_{b/a} \), since \( 0 < b/a \leq 1 \).

Henceforth, assume \( 0 < b < \sqrt{ac} \leq a \). (In particular, \( 2b < 2\sqrt{ac} \leq a + c \).) The second case is if \( b > c \); we then find \( t > 0 \) such that \( A - t \mathbf{1}_{2 \times 2} \) is singular. This condition amounts to a linear equation in \( t \), with solution (to be verified by the reader):

\[
t = \frac{ac-b^2}{a+c-2b} > 0.
\]

Then \( c - t = \frac{(b-c)^2}{a+c-2b} > 0 \), whence \( a - t, b - t > 0 \) and \( A = \begin{pmatrix} s_0(\mu) & s_1(\mu) \\ s_1(\mu) & s_2(\mu) \end{pmatrix} \), where

\[
\mu = \frac{ac-b^2}{a+c-2b} \delta_1 + \frac{(a-b)^2}{a+c-2b} \delta_{b-c + \varepsilon}, \quad \text{with} \quad \frac{b-c}{a-b} \in (0, 1),
\]

The third case is when \( 0 < b \leq c \leq \sqrt{ac} \leq a \), with \( b < \sqrt{ac} \). We now find \( t > 0 \) such that the matrix \( \begin{pmatrix} a-t & b+t \\ b+t & c-t \end{pmatrix} \in \mathbb{P}_2((0, \infty)) \) and is singular. Once again we need to solve a linear
18. The stronger Vasudeva and Schoenberg theorems. Bernstein’s theorem.  

Moment-sequence transforms.

equation, and find that 

\[
\begin{align*}
t &= \frac{ac - b^2}{a + c + 2b}, \\
c - t &= \frac{(b + c)^2}{a + c + 2b}, \\
a - t &= \frac{(a + b)^2}{a + c + 2b}, \\
b + t &= \frac{(a + b)(b + c)}{a + c + 2b},
\end{align*}
\]

and all of these are strictly positive. So \(a, b, c > 0\) are the first three moments of 

\[
\mu = \frac{ac - b^2}{a + c + 2b} \delta_{-1} + \frac{(a + b)^2}{a + c + 2b} \delta_{\frac{b+c}{a+b}}, \quad \text{with } \frac{b + c}{a + b} \in (0, 1].
\]

**Step 2:** Observe that the hypotheses of the stronger Horn–Loewner theorem [16.1] can be rephrased as saying that \(f[-]\) sends \(P_2((0, \infty))\) to \(P_2(\mathbb{R})\), and that assertion (4) in Theorem [18.14] holds for all measures \(a \delta_1 + b \delta_{u_0}\), where \(u_0 \in (0, 1)\) is fixed and \(a, b > 0\). By Step 1 and the hypotheses, we can apply the stronger Horn–Loewner theorem in our setting for each \(n \geq 3\), whence \(f|_{(0, \infty)}\) is smooth and absolutely monotonic. As in the proof of the stronger Vasudeva theorem [18.1], extend \(f\) continuously to the origin, say to a function \(\tilde{f}\), and apply Bernstein’s theorem [18.3]. It follows that \(\tilde{f}|_{[0, \infty)}\) is a power series with non-negative Maclaurin coefficients, and Step 2 follows by restricting to \(\tilde{f}|_{(0, \infty)} = f|_{(0, \infty)}\). □

**Remark 18.17.** From Step 2 above, it follows that assertions (1), (2) in the stronger Vasudeva theorem [18.1] can be further weakened, to deal only with \(2 \times 2\) matrices and with (Hankel TN moment matrices of) measures \(a \delta_1 + b \delta_{u_0}\), for a single fixed \(u_0 \in (0, 1)\) and all \(a, b > 0\).

We continue with the proof of the stronger Schoenberg theorem \[18.9\]. Previously, we have shown the first two of the six steps in the proof (these are listed following Theorem \[18.14\]).

**Step 3:** The next step is to show that if assertion (4) (or (5)) in Theorem \[18.14\] holds, then \( f \) is continuous on \( \mathbb{R} \). Notice from Steps 1, 2 of the proof that \( f \) is absolutely monotonic, whence continuous, on \((0,\infty)\).

19.1. Integration trick and proof of continuity. At this stage, we transition to moment-sequence preservers, via Hamburger’s theorem \[18.13\]. The following ‘integration trick’ will be used repeatedly in what follows: Suppose \( p(t) \) is a real polynomial that takes non-negative values for \( t \in [-1,1] \). Write \( p(t) = \sum_{k=0}^{\infty} a_k t^k \) (with only finitely many \( a_k \) nonzero, but not necessarily all positive, note). If \( \mu \geq 0 \) is an admissible measure – in particular, non-negative by Definition \[18.11\] – then by assumption and Hamburger’s theorem we have \( f(s_k(\mu)) = s_k(\sigma_\mu) \forall k \geq 0 \), for some admissible measure \( \sigma_\mu \geq 0 \) on \( \mathbb{R} \). Now assuming \( \sigma_\mu \) is supported on \([-1,1]\) (which is not \textit{a priori} clear from the hypotheses), we have:

\[
0 \leq \int_{-1}^{1} p(t) \, d\sigma_\mu = \sum_{k=0}^{\infty} \int_{-1}^{1} a_k t^k \, d\sigma_\mu = \sum_{k=0}^{\infty} a_k s_k(\sigma_\mu) = \sum_{k=0}^{\infty} a_k f(s_k(\mu)).
\]  

(19.1)

**Example 19.2.** Suppose \( p(t) = 1 - t^d \) on \([-1,1]\), for some integer \( d \geq 1 \). Then \( f(s_0(\mu)) - f(s_d(\mu)) \geq 0 \). As a further special case, if \( \mu = a \delta_1 + b \delta_{-1} + c \delta_{-1} \) as in Theorem \[18.14\](4), if \( \sigma_\mu \) is supported on \([-1,1]\) then this would imply:

\[
f(a+b+c) \geq f(a+bu_0^d + c(-1)^d), \quad \forall u_0 \in (0,1), \ a, b, c \geq 0.
\]

It is not immediately clear how the preceding inequalities can be obtained by considering only the preservation of matrix positivity by \( f[-] \) (or more involved such assertions). As we will explain shortly, this has connections to real algebraic geometry; in particular, to a well-known program of Hilbert.

Returning to the proof of continuity in Schoenberg’s theorem, we suppose without further mention that \( f \) satisfies only Theorem \[18.14\](4) above – and hence is absolutely monotonic on \((0,\infty)\). We begin by showing two preliminary lemmas, which are used in the proof of continuity.

**Lemma 19.3.** \( f \) is bounded on compact subsets of \( \mathbb{R} \).

**Proof.** If \( K \subseteq \mathbb{R} \) is compact, say \( K \subseteq [-M,M] \) for some \( M > 0 \), then note that \( f|_{(0,\infty)} \) is non-decreasing, whence \( 0 \leq |f(x)| \leq f(M), \forall x \in (0,M] \). Now apply \( f[-] \) to the matrix

\[
\begin{pmatrix}
x & -x \\
-x & x
\end{pmatrix},
\]

arising from \( \mu = x \delta_{-1}, \) with \( x > 0 \). This implies \( |f(-x)| \leq f(x) \leq f(M) \).

Similarly considering \( \mu = \frac{M}{2} \delta_1 + \frac{M}{2} \delta_{-1} \) shows that \( |f(0)| \leq f(M) \). \( \square \)

Now say \( \mu = a \delta_1 + b \delta_{-1} + c \delta_{-1} \) as above, or more generally, \( \mu \) is any non-negative measure supported in \([-1,1]\). It is easily seen that its moments \( s_k(\mu), \ k \geq 0 \) are all uniformly bounded in absolute value – in fact, by the mass \( s_0(\mu) \). Our next lemma shows that the converse is also true.

**Lemma 19.4.** Given an admissible measure \( \sigma \) on \( \mathbb{R} \), the following are equivalent:

1. The moments of \( \sigma \) are all uniformly bounded in absolute value.
2. The measure \( \sigma \) is supported on \([-1,1]\).

The positivity-certificate trick.

Proof. As discussed above, (2) \(\implies\) (1). To show the converse, suppose (1) holds but (2) fails. Then \(\sigma\) has positive mass in \((1, \infty) \cup (-\infty, -1)\). We obtain a contradiction in the first case; the proof is similar in the other case. Thus, suppose \(\sigma\) has positive mass on \((1, \infty) = [1 + \frac{1}{t}, 1 + \frac{1}{t}) \cup [1 + \frac{1}{t}, 1 + \frac{1}{t}) \cup \cdots \cup [1 + \frac{1}{n}, 1 + \frac{1}{n}] \cup \cdots\), where \(1/0 := \infty\). Then \(\sigma(I_n) > 0\) for some \(n \geq 0\), where we denote \(I_n := [1 + \frac{1}{n}, 1 + \frac{1}{n}]\) for convenience. But now we obtain the desired contradiction:

\[
s_{2k}(\sigma) = \int_{\mathbb{R}} x^{2k} \sigma = \int_{1 + \frac{1}{n+1}}^{1 + \frac{1}{n}} x^{2k} \sigma \geq \int_{1 + \frac{1}{n+1}}^{1 + \frac{1}{n}} (1 + \frac{1}{n+1})^k \sigma \geq \sigma(I_n)(1 + \frac{1}{n+1})^k,
\]

and this is not uniformly bounded over all \(k \geq 0\).

With these basic lemmas in hand, we have:

Proof of Step 3 for the stronger Schoenberg theorem: continuity. Given a function \(f\) satisfying Theorem [18.14](4), and any measure \(\mu = a\delta_1 + b\delta_{u_0} + c\delta_{-1}\) for \(u_0 > 0\) and \(a, b, c \geq 0\), note that \(|s_k(\mu)| \leq s_0(\mu) = a + b + c\). Hence by Lemma [19.3], the moments \(s_k(\sigma)\) are uniformly bounded over all \(k\). By Lemma [19.4], it follows that \(\sigma\) must be supported in \([-1, 1]\). In particular, we can apply the integration trick (19.1) above.

We use this trick to prove continuity at \(-\beta\) for \(\beta \geq 0\). (By Step 2, this proves the continuity of \(f\) on \(\mathbb{R}\).) Thus, fix \(\beta \geq 0\), \(u_0 \in (0, 1)\), and \(b > 0\), and define

\[
\mu := (\beta + bu_0)\delta_1 + b\delta_{u_0}.
\]

Let \(p_{\pm, 1}(t) := (1 \pm t)(1 - t^2)\); note that these polynomials are non-negative on \([-1, 1]\). By the integration trick (19.1),

\[
\int_{-1}^{1} p_{\pm, 1}(t) \, d\sigma_\mu(t) \geq 0
\]

\[
\implies s_0(\sigma_\mu) - s_2(\sigma_\mu) \geq \pm (s_1(\sigma_\mu) - s_3(\sigma_\mu))
\]

\[
\implies f(s_0(\mu)) - f(s_2(\mu)) \geq \pm |f(s_1(\mu)) - f(s_3(\mu))|
\]

\[
\implies f(\beta + b(1 + u_0)) - f(\beta + b(u_0 + u_0^2)) \geq \pm |f(-\beta) - f(-\beta - bu_0(1 + u_0^2))|.
\]

Now let \(b \to 0^+\). Then the left-hand side goes to zero by Step 2 (in the previous section), whence so does the right-hand side. This implies \(f\) is left-continuous at \(-\beta\) for all \(\beta \geq 0\). To show \(f\) is right-continuous at \(-\beta\), use \(\mu' := (\beta + bu_0^2)\delta_1 + b\delta_{u_0}\) instead of \(\mu\).

\[\square\]

Remark 19.5. Akin to its use in proving the continuity of \(f\), the integration trick (19.1) can also be used to prove the boundedness of \(f\) on compact sets \([-M, M]\), as in Lemma [19.3]. To do so, work with the polynomials \(p_{\pm, 0}(t) := 1 \pm t\), which are also non-negative on \([-1, 1]\). Given \(0 \leq x < M\), applying (19.1) to \(\mu = M\delta_{-x/M}\) and \(\mu' = x\delta_{-1}\) shows Lemma [19.3].

19.2. The integration trick explained: semi-algebraic geometry. Earlier in this section, we used the following 'integration trick': if \(\sigma \geq 0\) is a real measure supported in \([-1, 1]\) with all moments finite, i.e. the Hankel moment matrix \(H_\sigma := (s_{j+k}(\sigma))_{j,k=0}^\infty\) is positive semidefinite; if a polynomial \(p(t) = \sum_{k=0}^\infty a_k t^k\) is non-negative on \([-1, 1]\), then

\[
0 \leq \int_{-1}^{1} p(t) \, d\sigma = \sum_{k=0}^\infty \int_{-1}^{1} a_k t^k \, d\sigma = \sum_{k=0}^\infty a_k s_k(\sigma).
\]

This integration trick is at the heart of the link between moment problems and (Hankel) matrix positivity. We now explain the trick; namely, how this integral inequality can be
understood purely in terms of the positive semidefiniteness of $H_\sigma$. This also has connections to real algebraic geometry and Hilbert’s seventeenth problem.

The basic point is as follows: if a $d$-variate polynomial (in one or several variables) is a sum of squares of real polynomials – also called a s.o.s. polynomial – then it is automatically non-negative on $\mathbb{R}^d$. However, there exist other polynomials that are not sums of squares, but are non-negative on $\mathbb{R}^d$ – for instance, the well-known Motzkin polynomial $x^4y^2+x^2y^4-3x^2y^2+1$. Such phenomena are also studied on polytopes (results of Farkas, Pólya, and Handelman), and on more general ‘semi-algebraic sets’ including compact ones (results of Stengle, Schmüdgen, Putinar, and Vasilescu, among others).

Now given say a one-variable polynomial that is non-negative on a semi-algebraic set such as $[-1,1]$, one would like a positivity certificate for it, meaning a sum-of-squares ("s.o.s.") representation mentioned above, or more generally, a limiting s.o.s. representation. To make this precise, define the $L^1$-norm, or the Wiener norm, of a polynomial $p(t) = \sum_{k=0}^n a_k t^k$ as:

$$\|p(t)\|_{1, +} := \sum_{k \geq 0} |a_k|. \quad (19.6)$$

One would thus like to find a sequence $p_n$ of s.o.s. polynomials such that $\|p_n(t) - p(t)\|_{1, +} \to 0$ as $n \to \infty$. Two simple cases are if there exist polynomials $q_n(t)$ such that (i) $p_n(t) = q_n(t)^2$ $\forall n$, or (ii) $p_n(t) = \sum_{k=1}^n q_k(t)^2$ $\forall n$.

How does this connect to matrix positivity? It turns out that in our given situation, what is required is precisely a positivity certificate. For example, say $p(t) = (3-t)^2 = 9-6t+t^2 \geq 0$ on $\mathbb{R}$. Then

$$\int_{-1}^1 p \, d\sigma = 9s_0(\sigma) - 6s_1(\sigma) + s_2(\sigma) = (3, -1, \begin{pmatrix} s_0(\sigma) & s_1(\sigma) \\ s_1(\sigma) & s_2(\sigma) \end{pmatrix}) (3, -1)^T \quad (19.7)$$

where $e_0 = (1,0,0,\ldots)^T, e_1 = (0,1,0,0,\ldots)^T, \ldots$ comprise the standard basis for $\mathbb{R}^{N \cup \{0\}}$, and $H_\sigma$ is the semi-infinit, positive semidefinite Hankel moment matrix for $\sigma$. From this calculation, it follows that $\int_{-1}^1 p \, d\sigma$ is non-negative – and this holds more generally, whenever there exists a (limiting) s.o.s. representation for $p$.

We now prove the existence of such a limiting s.o.s. representation in two different ways for general $p(t)$, and in a constructive third way for the special family of polynomials

$$p_{\pm,n}(t) := (1 \pm t)(1 - t^2)^n, \quad n \geq 0.$$  

(Note, we used $p_{\pm,0}$ and $p_{\pm,1}$ to prove the local boundedness and continuity of $f$ on $\mathbb{R}$, respectively; and next time we will use $p_{\pm,n}$ to prove that smoothness implies real analyticity.)

**Proof 1:** We claim more generally that for any dimension $d \geq 1$, every polynomial that is non-negative on $[-1,1]^d$ has a limiting s.o.s. representation. This is proved at the end of the 1976 paper of Berg, Christensen, and Ressel in *Math. Ann.*

**Proof 2:** Here is a constructive proof of a positivity certificate for the polynomials $p_{\pm,n}(t) = (1 \pm t)(1 - t^2)^n, \quad n \geq 0$. (It turns out, we only need to work with these in order to show the stronger Schoenberg theorem.) First notice that

$$p_{+,0}(t) = (1 + t), \quad p_{-,0}(t) = (1 - t),$$

$$p_{+,1}(t) = (1 - t)(1 + t)^2, \quad p_{-,1}(t) = (1 + t)(1 - t)^2,$$

$$p_{+,2}(t) = (1 + t)(1 - t^2)^2, \quad p_{-,2}(t) = (1 - t)(1 - t^2)^2. \quad (19.8)$$
and so on. Thus, if we show that \( p_{\pm,0}(t) = 1 \pm t \) are limits of s.o.s polynomials, then so are \( p_{\pm,n}(t) \) for all \( n \geq 0 \) (where limits are taken in the Wiener norm). But we have:

\[
\begin{align*}
\frac{1}{2} (1 \pm t)^2 &= \frac{1}{2} \pm t^2, \\
\frac{1}{4} (1 - t^2)^2 &= \frac{1}{4} - \frac{t^2}{2} + \frac{t^4}{4}, \\
\frac{1}{8} (1 - t^4)^2 &= \frac{1}{8} - \frac{t^4}{4} + \frac{t^8}{8},
\end{align*}
\]

and so on. Adding the first \( k \) of these equations shows that the partial sum

\[
p_k^\pm(t) := (1 - \frac{1}{2^k}) \pm t + \frac{t^{2k}}{2^k} = (1 \pm t) + \frac{t^{2k} - 1}{2^k}
\]

is a s.o.s. polynomial, for every \( k \geq 1 \). This provides a positivity certificate for \( 1 \pm t \), as desired. It also implies the sought-for interpretation of the integration trick in Step 3 above:

\[
\left| \int_{-1}^{1} p_k^\pm(t) - (1 \pm t) \, d\sigma \right| \leq \int_{-1}^{1} |p_k^\pm(t) - (1 \pm t)| \, d\sigma \leq \int_{-1}^{1} \frac{1}{2^k} \, d\sigma + \int_{-1}^{1} \frac{t^{2k}}{2^k} \, d\sigma \leq \frac{1}{2^k} \cdot 2s_0(\sigma),
\]

which goes to 0 as \( k \to \infty \). Hence, using the notation following (19.7),

\[
\int_{-1}^{1} (1 \pm t) \, d\sigma = \lim_{k \to \infty} \int_{-1}^{1} p_k^\pm(t) \, d\sigma = \frac{1}{2} (e_0 \pm e_1)^T H_\sigma (e_0 \pm e_1) + \sum_{j=2}^{\infty} \frac{1}{2} (e_0 - e_{2j-1})^T H_\sigma (e_0 - e_{2j-1}),
\]

and this is non-negative because \( H_\sigma \) is positive semidefinite.

**Proof 3:** If we only want to interpret the integration trick (19.1) in terms of the positivity of the Hankel moment matrix \( H_\sigma \), then the restriction of using the Wiener norm \( \| \cdot \|_{1,+} \) can be relaxed, and one can work instead with the weaker notion of the uniform norm. With this metric, we claim more generally that every continuous function \( f(t_1, \ldots, t_d) \) that is non-negative on a compact subset \( K \subset \mathbb{R}^d \) has a limiting s.o.s. representation on \( K \). (Specialized to \( d = 1 \) and \( K = [-1, 1] \), this proves the integration trick.)

To see the claim, observe that \( \sqrt{f(t_1, \ldots, t_d)} : K \to [0, \infty) \) is continuous, so by the Stone–Weierstrass theorem, there exists a polynomial sequence \( q_n \) converging uniformly to \( \sqrt{f} \) in \( L^\infty(K) \). Thus \( q_n^2 \to f \) in \( L^\infty(K) \), as desired. Explicitly, if \( d = 1 \) and \( q_n(t) = \sum_{k=0}^{\infty} c_{n,k} t^k \), then define the semi-infinite vectors

\[
u_n := (c_{n,0}, c_{n,1}, \ldots)^T, \quad n \geq 1.
\]

Now consider any admissible measure \( \sigma \) supported in \( K \). We compute:

\[
\int_K f \, d\sigma = \lim_{n \to \infty} \int_K q_n^2(t) \, d\sigma = \lim_{n \to \infty} \nu_n^T H_\sigma \nu_n \geq 0,
\]

which is a positivity certificate for all continuous, non-negative functions on compact \( K \subset \mathbb{R} \).

This reasoning extends to all dimensions \( d \geq 1 \) and compact \( K \subset \mathbb{R}^d \), by Lemma 2.23.

**19.3. From the integration trick to the positivity-certificate trick.** The above ‘Proof 2’ is the key to understanding why Hamburger’s theorem is not required to prove the stronger Schoenberg theorem (namely, (5) \( \Rightarrow \) (3)). Specifically, in this part and the next, we only need to use the following fact:

*For each fixed \( n \geq 0 \), if \( \sum k a_k t^k \) is the expansion of \( p_{\pm,n}(t) = (1 \pm t)(1 - t^2)^n \geq 0 \) on \([-1, 1]\), then \( \sum k a_k f(s_k(\mu)) \geq 0 \).*

The positivity-certificate trick. 105

This was derived above using the integration trick (19.1) via the auxiliary admissible measure $\sigma_\mu$, which exists by Theorem 18.14(4). We now explain a workaround via a related ‘positivity-certificate trick’ that requires using only that $f[H_\mu]$ is positive semidefinite, hence allowing us to work with the weaker hypothesis, Theorem 18.14(5) instead. In particular, one can avoid using Hamburger’s theorem and requiring the existence of $\sigma_\mu$.

The positivity-certificate trick is as follows:

**Theorem 19.11.** Fix a semi-infinite real Hankel matrix $H = (f_{j+k})_{j,k \geq 0}$ that is positive semidefinite (i.e., its principal minors are), with all entries $f_j$ uniformly bounded. If a polynomial $p(t) = \sum_{j \geq 0} a_j t^j$ has a positivity certificate – i.e. a Wiener-limiting s.o.s. representation – then $\sum_{j \geq 0} a_j f_j \geq 0$.

By the ‘Proof 2’ above (see the discussion around (19.9)), Theorem 19.11 applies to $p = p_{\pm,n}$ for all $n \geq 0$ and $H = f[H_\mu]$, where $f, \mu$ are as in Theorem 18.14(5). This implies the continuity of the entrywise positivity preserver $f$ in the above discussion, and also suffices to complete the proof below, of the stronger Schoenberg theorem.

**Proof.** As an illustrative special case, if $p(t)$ is the square of a polynomial $q(t) = \sum_{j \geq 0} c_j t^j$, then as in (19.7),

$$\sum_{j \geq 0} a_j f_j = \sum_{j,k \geq 0} c_j c_k f_{j+k} = u^T H u \geq 0,$$

where $u = (c_0, c_1, \ldots)^T$.

By additivity, the result therefore also holds for a sum of squares of polynomials. The subtlety in working with a limiting s.o.s. is that the degrees of each s.o.s. polynomial in the limiting sequence need not be uniformly bounded. Nevertheless, suppose in the Wiener norm (19.6) that

$$p(t) = \lim_{n \to \infty} q_n(t), \text{ where } q_n(t) = \sum_{k=0}^{K_n} q_{n,k}(t)^2$$

is a sum of squares of polynomials for each $n$.

Define the linear functional $\Psi_H$ (given the Hankel matrix $H$) that sends a polynomial $p(t) = \sum_{j \geq 0} a_j t^j$ to the scalar $\Psi_H(p) := \sum_{j \geq 0} a_j f_j$. Now define the vectors $u_{n,k}$ via:

$$q_{n,k}(t) = \sum_{j \geq 0} q_{n,k}^{[j]} t^j \in \mathbb{R}[t], \quad u_{n,k} := (q_{n,k}^{[0]}, q_{n,k}^{[1]}, \ldots)^T.$$

Similarly, define $q_n(t) = \sum_{j \geq 0} q_n^{[j]} t^j$. Then for all $n \geq 1$,

$$\sum_{j \geq 0} q_n^{[j]} f_j = \Psi_H(q_n) = \sum_{k=0}^{K_n} \Psi_H(q_{n,k}^2) = \sum_{k=0}^{K_n} u_{n,k}^T H u_{n,k} \geq 0.$$

Finally, taking the limit as $n \to \infty$, and writing $p(t) = \sum_{j \geq 0} a_j t^j$, we claim that

$$\sum_{j \geq 0} a_j f_j = \lim_{n \to \infty} \sum_{j \geq 0} q_n^{[j]} f_j \geq 0.$$

Indeed, the (first) equality holds because if $M \geq \sup_j |f_j|$ is a uniform (and finite) upper bound, then

$$\left| \sum_{j \geq 0} q_n^{[j]} f_j - \sum_{j \geq 0} a_j f_j \right| \leq \sum_{j \geq 0} |q_n^{[j]} - a_j| |f_j| \leq M \|q_n - p\|_{1,+},$$

and this goes to zero as $n \to \infty$.  \qed
20. Proof of stronger Schoenberg Theorem: II. Smoothness implies real analyticity.

Having explained the positivity-certificate trick, we return to the proof of the stronger Schoenberg theorem. The present goal is to prove that if a smooth function \( f : \mathbb{R} \to \mathbb{R} \) satisfies assertion (5) in Theorem 18.14, then \( f \) is real analytic and hence satisfies assertion (3) in Theorem 18.9. (See Steps (4), (5) in the discussion following Theorem 18.14.) To show these results, we first discuss the basic properties of real analytic functions that are required in the proofs.

20.1. Preliminaries on real analytic functions.

**Definition 20.1.** Suppose \( I \subset \mathbb{R} \) is an open interval, and \( f : I \to \mathbb{R} \) is smooth, denoted \( f \in C^\infty(I) \). Recall that the Taylor series of \( f \) at a point \( x \in I \) is

\[
(Tf)_x(y) := \sum_{j=0}^{\infty} \frac{f^{(j)}(x)}{j!} (y-x)^j, \quad y \in I,
\]

if this sum converges at \( y \). Notice that this sum is not equal to \( f(y) \) in general.

Next, we say that \( f \) is real analytic on \( I \), denoted \( f \in C^\infty(I) \), if \( f \in C^\infty(I) \) and for all \( x \in I \) there exists \( \delta_x > 0 \) such that the Taylor series of \( f \) at \( x \) converges to \( f \) on \((x - \delta_x, x + \delta_x)\).

Clearly, real analytic functions on \( I \) form a real vector space. Less obvious is the following useful property, which is stated without proof:

**Proposition 20.2.** Real analytic functions are closed under composition. More precisely, if \( I \rightarrow^f J \rightarrow^g \mathbb{R} \), and \( f, g \) are real analytic on their domains, then so is \( g \circ f \) on \( I \).

We also develop a few preliminary results on real analytic functions, which are needed to prove the stronger Schoenberg theorem. We begin with an example of real analytic functions, which depicts what happens in our setting.

**Lemma 20.3.** Suppose \( I = (0, R) \) for \( 0 < R \leq \infty \), and \( f(x) = \sum_{k=0}^{\infty} c_k x^k \) on \( I \), where \( c_k \geq 0 \) \( \forall k \). Then \( f \in C^\infty(I) \) and \((Tf)_0(x)\) converges at all \( x \in I \) such that \(|x-a| < R-a\).

In particular, if \( R = \infty \) and \( a > 0 \), then \((Tf)_0(x) \rightarrow f(x)\) on the domain of \( f \).

**Proof.** Note that \( \sum_{k=0}^{\infty} c_k x^k \) converges on \((-R, R)\). Thus, we show more generally that \((Tf)_0(x)\) converges to \( f(x) \) for \(|x-a| < R-a\), \(|x| < R\), \( a \geq 0 \) (whenever \( f \) is defined at \( x \)). Indeed,

\[
f(x) = \sum_{k=0}^{\infty} c_k ((x-a) + a)^k = \sum_{k=0}^{\infty} \sum_{j=0}^{k} \binom{k}{j} c_k (x-a)^j a^{k-j}.
\]

Notice that this double sum is absolutely convergent, since

\[
\sum_{k=0}^{\infty} \sum_{j=0}^{k} \binom{k}{j} c_k |x-a|^j a^{k-j} = f(a + |x-a|) < \infty.
\]

Hence we can rearrange the double sum (e.g. by Fubini’s theorem), to obtain

\[
f(x) = \sum_{j=0}^{\infty} \left( \sum_{m=0}^{\infty} \binom{m+j}{j} c_{m+j} a^m \right) (x-a)^j = \sum_{j=0}^{\infty} \frac{f^{(j)}(a)}{j!} (x-a)^j = (Tf)_0(x)
\]

using standard properties of power series. In particular, \( f \) is real analytic on \( I \). \( \square \)
We also require the following well-known result on zeros of real analytic functions.

**Theorem 20.4** (Identity theorem). Suppose $I \subset \mathbb{R}$ is an open interval and $f, g : I \to \mathbb{R}$ are real analytic. If the subset of $I$ where $f = g$ has an accumulation point in $I$, then $f \equiv g$ on $I$.

In other words, the zeros of a non-zero (real) analytic function form a discrete set.

**Proof.** Without loss of generality, we may suppose $g \equiv 0$. Suppose $c \in I$ is an accumulation point of the zero set of $f$. Expand $f$ locally at $c$ into its Taylor series. We claim that $f^{(k)}(c) = 0$ for all $k \geq 0$. Indeed, suppose for contradiction that $f^{(0)}(c) = \cdots = f^{(k-1)}(c) = 0 \neq f^{(k)}(c)$ for some $k \geq 0$. Then, $f(x) = \frac{f^{(k)}(c)}{k!} + o(x-c)$, whence $f$ is nonzero close to $c$, and this contradicts the hypotheses. Thus, $f^{(k)}(c) = 0 \forall k \geq 0$, which in turn implies that $f \equiv 0$ on an open interval around $c$. □

20.2. **Proof of the stronger Schoenberg theorem for smooth functions.** We continue with the proof of the stronger Schoenberg theorem ((5) $\implies$ (2) in Theorems 18.9 and 18.14).

Akin to the proof of the stronger Horn–Loewner theorem 16.1, we have shown that any function satisfying the hypotheses in Theorem 18.14(5) must be continuous. Hence by the first two steps in the proof – listed after Theorem 18.14 – we have that $f(x) = \sum_{k=0}^{\infty} c_k x^k$ on $[0, \infty)$, with all $c_k \geq 0$.

Again similar to the proof of the stronger Horn–Loewner theorem 16.1, we next prove the stronger Schoenberg theorem for smooth functions. The key step here is:

**Theorem 20.5.** Let $f \in C^\infty(\mathbb{R})$ be as in the preceding discussion, and define the family of smooth functions

$$H_a(x) := f(a + e^x), \quad a, x \in \mathbb{R}.$$  

Then $H_a$ is real analytic on $\mathbb{R}$, for all $a \in \mathbb{R}$.

For ease of exposition, we break up the proof into several steps.

**Lemma 20.6.** For all $n \geq 1$, we have

$$H^{(n)}(x) = a_{n,1} f'(a + e^x) e^x + a_{n,2} f''(a + e^x) e^{2x} + \cdots + a_{n,n} f^{(n)}(a + e^x) e^{nx},$$

where $a_{n,j}$ is a positive integer for all $1 \leq j \leq n$.

**Proof and remarks.** One shows by induction on $n \geq 1$ (with the base case of $n = 1$ immediate) that the array $a_{n,j}$ forms a ‘weighted variant’ of Pascal’s triangle, in that:

$$a_{n,j} = \begin{cases} 1, & \text{if } j = 1, n, \\ a_{n-1,j-1} + ja_{n-1,j}, & \text{otherwise.} \end{cases}$$
This concludes the proof. Notice that some of the entries of the array \( a_{n,j} \) are easy to compute:

\[
a_{n,1} = 1, \quad a_{n,2} = 2^{n-1} - 1, \quad a_{n,n-1} = \binom{n}{2}, \quad a_{n,n} = 1.
\]

An interesting combinatorial exercise may be to seek a closed-form expression and a combinatorial interpretation for the other entries. \( \square \)

**Lemma 20.7.** We have the following bound:

\[
|H^{(n)}_a(x)| \leq H^{(n)}_{|a|}(x), \quad \forall a, x \in \mathbb{R}, \ n \in \mathbb{Z}_{\geq 0}.
\]  

(20.8)

**Proof.** By Lemma 20.6 we have that \( H^{(n)}_{|a|}(x) \geq 0 \) for all \( a, x, n \) as in (20.8), so it remains to show the inequality. For this, we assume \( a < 0 \), and use the ‘positivity-certificate trick’ from the previous section – i.e. Theorem 19.11 applied to the polynomials

\[
p_{\pm,n}(t) := (1 \pm t)(1 - t^2)^n, \quad n \geq 0
\]

and the admissible measure

\[
\mu := |a|\delta_{-1} + e^x\delta_{e^{-h}}, \quad a, h > 0, \ x \in \mathbb{R}.
\]

Notice that \( p_{\pm} \geq 0 \) on \([-1,1]\). Hence by Theorem 19.11 and akin to the calculation in the previous section to prove continuity – we get:

\[
\sum_{k=0}^{n} \binom{n}{k} (-1)^k f(|a| + e^{-2kh}) \geq \sum_{k=0}^{n} \binom{n}{k} (-1)^{n-k} f(a + e^{-(2k+1)h}).
\]

Dividing both sides by \((2h)^n\) and sending \( h \to 0^+\), we obtain:

\[
|H^{(n)}_{a}(x)| \geq |H^{(n)}_{|a|}(x)|,
\]

which proves (20.8), as desired. \( \square \)

**Remark 20.9.** In this computation, we do not need to use the measures \( \mu = |a|\delta_{-1} + e^x\delta_{e^{-h}} \) for all \( h > 0 \). It suffices to fix a single \( u_0 \in (0,1) \) and consider the sequence \( h_n := -\log(u_0)/n, \) whence we work with \( \mu = |a|\delta_{-1} + e^x\delta_{u_0^{1/n}} \) (supported at 1, \( u_0^{1/n} \)) for \( a > 0, x \in \mathbb{R}, n \geq 1 \).

**Lemma 20.10.** For all integers \( n \geq 0 \), the assignment \((a, x) \mapsto H^{(n)}_{a}(x)\) is non-decreasing in both \( a \geq 0 \) and \( x \in \mathbb{R} \). In particular if \( a \geq 0 \), then \( H_a \) is absolutely monotonic on \( \mathbb{R} \), and its Taylor series at \( b \in \mathbb{R} \) converges absolutely at all \( x \in \mathbb{R} \).

**Proof.** The monotonicity in \( a \geq 0 \) follows from the absolute monotonicity of \( f|_{[0,\infty)} \) mentioned above. The monotonicity in \( x \) for a fixed \( a \geq 0 \) follows because \( H^{(n+1)}_{a}(x) \geq 0 \) by Lemma 20.6.

To prove the (absolute) convergence of \((TH_a)_b\) at \( x \in \mathbb{R} \), notice that

\[
|(TH_a)_b(x)| = \left| \sum_{n=0}^{\infty} H^{(n)}_{a}(b) \frac{(x-b)^n}{n!} \right| \leq \sum_{n=0}^{\infty} H^{(n)}_{a}(b) \frac{|x-b|^n}{n!} = (TH_a)_b(b + |x-b|).
\]

We claim that this final sum is bounded above by \( H_{a}(b + |x-b|) \), which would complete the proof. Indeed, by Taylor’s theorem, the \( n \)th Taylor remainder term for \( H_{a}(b + |x-b|) \) can be written as (see e.g. (18.7))

\[
\int_{b}^{b+|x-b|} \frac{(b + |x-b| - t)^n}{n!} H^{(n+1)}_{a}(t) \ dt,
\]

which is non-negative from above. This shows the claim and completes the proof. \( \square \)
Now we can prove the real-analyticity of $H_a$:

**Proof of Theorem 20.5.** Fix scalars $a, \delta > 0$. We show that for all $b \in [-a,a]$ and $x \in \mathbb{R}$, the $n$th remainder term for the Taylor series $TH_b$ around the point $x$ converges to zero as $n \to \infty$, uniformly near $x$. More precisely, define

$$
\Psi_n(x) := \sup_{y \in [x-\delta,x+\delta]} |R_n((TH_b)_x)(y)|.
$$

We then claim $\Psi_n(x) \to 0$ as $n \to \infty$, for all $x$. This will imply that at all $x \in \mathbb{R}$, $(TH_b)_x$ converges to $H_b$ on a neighborhood of radius $\delta$. Moreover, this holds for all $\delta > 0$ and at all $b \in [-a,a]$ for all $a > 0$.

Thus, it remains to prove for each $x \in \mathbb{R}$ that $\Psi_n(x) \to 0$ as $n \to \infty$. By the above results, we have:

$$
|H_b^{(n)}(y)| \leq H_{|b|}^{(n)}(y) \leq H_a^{(n)}(x+\delta), \quad \forall b \in [-a,a], \ y \in [x-\delta,x+\delta], \ n \in \mathbb{Z} > 0.
$$

Using a standard estimate for the Taylor remainder, for all $b, y, n$ as above, it follows that

$$
|R_n((TH_b)_x)(y)| \leq H_a^{(n+1)}(x+\delta) \frac{|y-x|^{n+1}}{(n+1)!} \leq H_a^{(n+1)}(x+\delta) \frac{\delta^{n+1}}{(n+1)!}.
$$

But the right-hand term goes to zero by the calculation in Lemma 20.10 since

$$
0 \leq \sum_{n=-1}^{\infty} H_a^{(n+1)}(x+\delta) \frac{\delta^{n+1}}{(n+1)!} \leq H_a(x+\delta+\delta) = f(a+e^{x+2\delta}) < \infty.
$$

Hence we obtain:

$$
\lim_{n \to \infty} \sup_{y \in [x-\delta,x+\delta]} |R_n((TH_b)_x)(y)| \to 0, \quad \forall x \in \mathbb{R}, \ \delta > 0, \ b \in [-a,a], \ a > 0.
$$

From above, this shows that the Taylor series of $H_b$ converges locally to $H_b$ at all $x \in \mathbb{R}$, for all $b$ as desired. (In fact, the ‘local’ neighborhood of convergence around $x$ is all of $\mathbb{R}$.)

With the above analysis in hand, we can prove **Steps 4, 5** of the proof of the stronger Schoenberg theorem:

**Suppose** $f : \mathbb{R} \to \mathbb{R}$ **satisfies assertion (4) of Theorem 18.14.**

(4) If $f$ is smooth on $\mathbb{R}$, then $f$ is real analytic on $\mathbb{R}$.

(5) If $f$ is real analytic on $\mathbb{R}$, then $f(x) = \sum_{k=0}^{\infty} c_k x^k$ on $\mathbb{R}$, with $c_k \geq 0 \ \forall k$.

**Proof of Step 4 for the stronger Schoenberg theorem.** Given $x \in \mathbb{R}$, we want to show that the Taylor series $(Tf)_x$ converges to $f$ locally around $x$. Choose $a > |x|$ and define

$$
L_a(y) := \log(a+y) = \log(a) + \log(1+y/a), \quad y \in (-a,a).
$$

This is real analytic on $(-a,a)$ (e.g. akin to Lemma 20.3). Hence by Proposition 20.2 and Theorem 20.5, the composite

$$
y \xrightarrow{L_a} \log(a+y) \xrightarrow{H_{-a}} H_{-a}(L_a(y)) = f(-a + \exp(\log(a+y))) = f(y)
$$

is also real analytic on $(-a,a)$, whence around $x \in \mathbb{R}$. □

**Proof of Step 5 for the stronger Schoenberg theorem.** By Step 4, $f$ is real analytic on $\mathbb{R}$; and as observed above, by Steps 1 and 2 $f(x) = \sum_{k=0}^{\infty} c_k x^k$ on $(0,\infty)$, with $c_k \geq 0 \ \forall k$. Let $g(x) := \sum_{k=0}^{\infty} c_k x^k \in C^\omega(\mathbb{R})$. Since $f \equiv g$ on $(0,\infty)$, it follows by the Identity Theorem 20.4 that $f \equiv g$ on $\mathbb{R}$. □

We can now complete the proof of the final Step 6 of the stronger Schoenberg theorem. Namely, suppose \( f : \mathbb{R} \to \mathbb{R} \) is such that for each measure
\[
\mu = a\delta_1 + b\delta_{u_0} + c\delta_{-1}, \quad \text{with } u_0 \in (0,1), \ a, b, c \geq 0,
\]
with semi-infinite Hankel moment matrix \( H_\mu \), the matrix \( f[H_\mu] \) is positive semidefinite.

Under these assumptions, we have previously shown (in Steps 1, 2; 3; 4, 5 respectively):
- There exist real scalars \( c_0, c_1, \cdots \geq 0 \) such that \( f(x) = \sum_{k=0}^{\infty} c_k x^k \) for all \( x \in (0, \infty) \).
- \( f \) is continuous on \( \mathbb{R} \).
- If \( f \) is smooth, then \( f(x) = \sum_{k=0}^{\infty} c_k x^k \) on \( \mathbb{R} \).

We now complete the proof by showing that one can pass from smooth functions to continuous functions. The tools we will use are the “three M’s”: Montel, Morera, and Mollifiers. We first discuss some basic results in complex analysis that are required below.

21.1. Tools from complex analysis. Throughout this section, \( D \subset \mathbb{C} \) is an open set.

Definition 21.1. Suppose \( D \subset \mathbb{C} \) is open and \( f : D \to \mathbb{C} \) is a continuous function.

1. (Holomorphic.) We say \( f \) is holomorphic at a point \( x \in D \) if the limit \( \lim_{y \to x} \frac{f(y) - f(x)}{y - x} \) exists. We say \( f \) is holomorphic on \( D \) if it is holomorphic at every point of \( D \).
2. (Complex analytic.) \( f \) is said to be complex analytic around \( c \in D \) if \( f \) can be expressed as a power series locally around \( c \), which converges to \( f(x) \) for every \( x \) sufficiently close to \( c \). Similarly, \( f \) is analytic on \( D \) if it is so at every point of \( D \).
3. Let \( F \) be a family of holomorphic functions : \( D \to \mathbb{C} \). We say \( F \) is normal if given any compact \( K \subset D \) and a sequence \( \{f_n : n \geq 1\} \subset F \), there exists a subsequence \( f_{n_k} \) and a function \( f : K \to \mathbb{C} \) such that \( f_{n_k} \to f \) uniformly on \( K \).

Remark 21.2. Note that it is not specified that the limit function \( f \) be holomorphic. However, this will turn out to be the case, as we shall see later.

We use without proof the following results (and Cauchy’s theorem, which we do not state).

Theorem 21.3. Let \( D \subset \mathbb{C} \) be an open set.

1. A function \( f : D \to \mathbb{C} \) is holomorphic if and only if \( f \) is complex analytic.
2. (Montel.) Let \( F \) be a family of holomorphic functions on \( D \). If \( F \) is uniformly bounded on \( D \), then \( F \) is normal on \( D \).
3. (Morera.) Suppose that for every closed oriented piecewise \( C^1 \) curve \( \gamma \) in \( D \), we have \( \oint_\gamma f \, dz = 0 \). Then \( f \) is holomorphic on \( D \).

21.2. Proof of the stronger Schoenberg theorem: conclusion. Let \( f : \mathbb{R} \to \mathbb{R} \) be as described above; in particular, \( f \) is continuous on \( \mathbb{R} \) and absolutely monotonic on \( (0, \infty) \).

As discussed in the proof of the stronger Horn–Loewner theorem [16.1], we mollify \( f \) with the family \( \phi_\delta(u) = \phi(u/\delta) \) for \( \delta > 0 \) as in Proposition [17.4]. As shown in [17.7], \( f_\delta \) satisfies assertion (5) in Theorem [18.14] so (e.g. by the last bulleted point above, and Steps 4, 5,)
\[
f_\delta(x) = \sum_{k=0}^{\infty} c_{k,\delta} x^k \quad \forall x \in \mathbb{R}, \quad \text{with } c_{k,\delta} \geq 0 \forall k \geq 0, \ \delta > 0.
\]

Since \( f_\delta \) is a power series with infinite radius of convergence, it extends analytically to an entire function on \( \mathbb{C} \) (see e.g. Lemma [20.3]). Let us call this \( f_\delta \) as well; now define
\[
\mathcal{F} := \{f_{1/n} : n \geq 1\}.
\]
We claim that for any $0 < r < \infty$, the family $F$ is uniformly bounded on the complex disc $D(0, r)$. Indeed, since $f_\delta \to f$ uniformly on $[0, r]$ by Proposition [17.4], we have that $|f_{1/n} - f|$ is uniformly bounded over all $n$ and on $[0, r]$, say by $M_r > 0$. Now if $z \in D(0, r)$, then

$$|f_{1/n}(z)| \leq \sum_{k=0}^{\infty} c_{k, 1/n} |z|^k = f_{1/n}(|z|) \leq M_r + f(|z|) \leq M_r + f(r) < \infty,$$

and this bound (uniform over $z \in D(0, r))$ does not depend on $n$.

By Montel’s theorem, the previous claim implies that $F$ is a normal family on $D(0, r)$ for each $r > 0$. Hence on the closed disc $\overline{D}(0, r)$, there is a subsequence $f_{1/n_l}$ with $n_l$ increasing, which converges uniformly to some (continuous) $g = g_r$. Since $f_{1/n_l}$ is holomorphic for all $l \geq 1$, by Cauchy’s theorem we obtain for every closed oriented piecewise $C^1$ curve $\gamma \subset D(0, r)$:

$$\oint_\gamma g_r \, dz = \lim_{l \to \infty} \oint_\gamma f_{1/n_l} \, dz = \lim_{l \to \infty} \oint_\gamma f_{1/n_l} \, dz = 0.$$

It follows by Morera’s theorem that $g_r$ is holomorphic, whence analytic, on $D(0, r)$. Moreover, $g_r \equiv f$ on $(-r, r)$ by the properties of mollifiers; thus, $f$ is real analytic on $(-r, r)$ for every $r > 0$. We are now done by the previous step of the proof – i.e., the Identity Theorem [20.4].

21.3. Concluding remarks and variations. We conclude with several generalizations of the above results. First, the results by Horn–Loewner, Vasudeva, and Schoenberg (more precisely, their stronger versions) that were shown in this part, together with the proofs given above, can be adapted to versions with bounded domains $(0, \rho)$ or $(-\rho, \rho)$ for $0 < \rho < \infty$. The small change is to use admissible measures with bounded mass:

$$\mu = a\delta_1 + b\delta_{u_1} + c\delta_{-1},$$

where $u_0 \in (0, 1)$, $a, b, c \geq 0$; but moreover, one now imposes the condition that $s_0(\mu) = a + b + c < \rho$.

Second, all of these results, including for bounded domains (i.e., masses of the underlying measures), can be extended to studying functions of several variables. In this case, given a domain $I \subset \mathbb{R}$ and integers $m, n \geq 1$, a function $f : I^m \to \mathbb{R}$ acts entrywise on an $m$-tuple of $n \times n$ matrices $A_1 = (a_{1j}^{(1)}), \ldots, A_m = (a_{jk}^{(m)})$ in $I^{m \times m}$, via:

$$f[A_1, \ldots, A_m] := (f(a_{jk}^{(1)}), \ldots, f(a_{jk}^{(m)}))_{j,k=1}^n.$$

One can now ask the multivariable version of the same question as above:

"Which functions, when applied entrywise to $m$-tuples of matrices, preserve positivity?"

Observe that the coordinate functions $f(x_1, \ldots, x_m) := x_l$ work for all $1 \leq l \leq m$. Hence by the Schur product theorem and the Pólya–Szegö observation (Lemma [15.1], since $\mathbb{P}_n$ is a closed convex cone for all $n \geq 1$), every convergent multi-power series of the form

$$f(x) := \sum_{\mathbf{n} \geq 0} c_\mathbf{n} x^n,$$

with $c_\mathbf{n} \geq 0 \forall \mathbf{n} \geq 0$ (21.5)

preserves positivity in all dimensions (where $x^n := x_1^{n_1} \cdots x_m^{n_m}$, etc.). Akin to the Schoenberg–Rudin theorem in the one-variable case, it was shown by FitzGerald, Michelli, and Pinkus in Linear Algebra Appl. (1995) that the functions (21.5) are the only such preservers.

One can ask if the same result holds when one restricts the test set to $m$-tuples of Hankel matrices of rank at most 3, as in the treatment above. While this does turn out to yield the same classification, the proofs get more involved and now require multivariable machinery. For these stronger multivariate results, we refer the reader to the 2016 preprint “Moment-sequence transforms” by Belton, Guillot, Khare, and Putinar.

In this section we reproduce the complete proof of the theorem by Boas and Widder on functions with non-negative forward differences (Duke Math. J., 1940). This result was stated as Theorem 17.9(2) above, and we again write down its statement here for convenience. In it (and below), recall from just before Theorem 17.9 that given an interval $I \subset \mathbb{R}$ and a function $f : I \to \mathbb{R}$, the $k$th order forward differences of $f$ with step size $h > 0$ are defined as follows:

$$(\Delta_0^k f)(x) := f(x), \quad (\Delta_h^k f)(x) := (\Delta_{h}^{k-1} f)(x+h) - (\Delta_h^{k-1} f)(x) = \sum_{j=0}^{k} \binom{k}{j} (-1)^{k-j} f(x+jh),$$

whenever $k > 0$ and $x, x + kh \in I$. It is easily seen that these difference operators commute:

$$\Delta_n^m (\Delta_n^m f(x)) = \Delta_n^m (\Delta_n^m f(x)), \quad \text{whenever } x, x + m\delta + n e \in I,$$

and so we will omit parentheses and possibly permute these operators below, without further reference. Now we (re)state the theorem of interest:

**Theorem 22.1** (Boas–Widder). Suppose $k \geq 2$ is an integer, $I \subset \mathbb{R}$ is an open interval, bounded or not, and $f : I \to \mathbb{R}$ is a function that satisfies the following condition:

$$(\Delta_h^k f)(x) \geq 0 \text{ whenever } h > 0 \text{ and } x, x + kh \in I, \quad \text{and } f \text{ is continuous on } I. \quad (H_k)$$

(In other words, $f$ is continuous and has all forward differences of order $k$ non-negative on $I$.) Then on all of $I$, the function $f^{(k-2)}$ exists, is continuous and convex, and has non-decreasing left and right hand derivatives.

22.1. Further remarks and results. Before writing down Boas and Widder’s proof of Theorem 22.1, we make several additional observations beyond the result and its proof. The first observation (which was previously mentioned following Theorem 17.9(2)) is that while $f_{\pm}^{(k-1)}$ is non-decreasing by the above theorem, it is not always true that any other lower-order derivatives $f_1, \ldots, f_{k-2}^{(k-2)}$ are non-decreasing on $I$. For example, let $0 \leq l \leq k-2$ and consider $f(x) := -x^{l+1}$ on $I \subset \mathbb{R}$; then $f^{(l)}$ is strictly decreasing on $I$.

Second, it is natural to seek examples of non-smooth functions satisfying the differentiability conditions of Theorem 22.1, but no more – in other words, to explore if Theorem 22.1 is indeed “sharp”. This is now verified to be true:

**Example 22.2.** Let $I = (a, b) \subset \mathbb{R}$ be an open interval, where $-\infty \leq a < b \leq \infty$. Consider any function $g : I \to \mathbb{R}$ that is non-decreasing, whence Lebesgue integrable. For any interior point $c \in I$, the function $f_2(x) := \int_c^x g(t) \, dt$ satisfies $(H_2)$, but clearly not every monotone $g$ gives rise to an anti-derivative that is differentiable on all of $I$. Also, for completeness we write down the proof of $(H_2)$:

$$\Delta_h^2 f_2(x) = \int_c^x g(t) \, dt \, h - \int_c^{x+h} g(t) \, dt + \int_c^{x+2h} g(t) \, dt$$

$$= \int_{x+h}^{x+2h} g(t) \, dt - \int_x^{x+h} g(t) \, dt$$

$$= \int_x^{x+h} (\Delta_h g)(t) \, dt \geq 0.$$

Finally, to see that the condition $(H_k)$ is sharp for all $k > 2$ as well, define $f$ to be the $(k-1)$-fold indefinite integral of $g$. We claim that $f$ satisfies $(H_k)$. Continuity is obvious;
and to study the \( k \)th order divided differences of \( f \), first note by the fundamental theorem of calculus that \( f \) is \((k - 2)\)-times differentiable, with \( f^{(k-2)}(x) \equiv f_2(x) = \int_c^x g(t) \, dt \). In particular, \( \Delta_h^2 f \in C^{k-2}(a, b - kh) \) whenever \( a < x < x + kh < b \) as in \((H_2)\).

Now given such \( x, h \), we compute using the Cauchy mean-value theorem \(17.9(1)\) for divided differences (and its notation):

\[
\Delta_h^k f(x) = \Delta_h^{k-2}(\Delta_h^2 f)(x) = h^{k-2}D_h^{k-2}(\Delta_h^2 f)(x) = \frac{h^{k-2}}{(k-2)!}(\Delta_h^2 f)^{(k-2)}(y),
\]

for some \( y \in (a, b - 2h) \). But this is easily seen to equal

\[
\frac{h^{k-2}}{(k-2)!}(\Delta_h^2 f)^{(k-2)}(y) = \frac{h^{k-2}}{(k-2)!} \Delta_h^2 f_2(y),
\]

and we showed above that this is non-negative. \( \square \)

Our final observation in this part is that there are natural analogues for \( k = 0, 1 \) of the Boas–Widder theorem (which is stated for \( k \geq 2 \)). For this, we make the natural definition: \( f^{(k)} \) will denote the \(|k|\)-fold anti-derivative of \( f \). Since \( f \) is assumed to be continuous, this is just the iterated indefinite Riemann integral starting at any interior point of \( I \). With this notation at hand, we have:

**Proposition 22.3.** The Boas–Widder theorem \(22.1\) also holds for \( k = 0, 1 \).

**Proof.** In both cases, the continuity of \( f^{(k-2)} \) is immediate by the fundamental theorem of calculus. Next, suppose \( k = 1 \) and choose \( c \in I \). We verify that if \( f \) is continuous and non-decreasing (i.e., \((H_1)\)), then \( f^{(-1)}(x) := \int_c^x f(t) \, dt \) is convex on \( I \). Indeed, given \( x_0 < x_1 \in I \), define \( x_{\lambda} := (1 - \lambda)x_0 + \lambda x_1 \) for \( \lambda \in [0, 1] \), and compute:

\[
(1 - \lambda)f^{(-1)}(x_0) + \lambda f^{(-1)}(x_1) - f^{(-1)}(x_{\lambda})
= (1 - \lambda)\int_c^{x_1} 1(t \leq x_0)f(t) \, dt + \lambda \int_c^{x_1} 1(t \leq x_1)f(t) \, dt - \int_c^{x_1} 1(t \leq x_{\lambda})f(t) \, dt
= - (1 - \lambda)\int_{x_0}^{x_{\lambda}} f(t) \, dt + \lambda \int_{x_{\lambda}}^{x_1} f(t) \, dt.
\]

But since \( f \) is non-decreasing, each integral – together with the accompanying sign – is bounded below by the corresponding expression where \( f(t) \) is replaced by \( f(x_{\lambda}) \). An easy computation now yields:

\[
(1 - \lambda)f^{(-1)}(x_0) + \lambda f^{(-1)}(x_1) - f^{(-1)}(x_{\lambda}) \geq f(x_{\lambda})(\lambda(x_1 - x_{\lambda}) - (1 - \lambda)(x_{\lambda} - x_0)) = 0;
\]

therefore \( f^{(-1)} \) is convex, as desired.

This shows the result for \( k = 1 \). Next, if \( k = 0 \) then \( f \) is continuous and non-negative on \( I \), whence \( f^{(-1)} \) is non-decreasing on \( I \). Now the above computation shows that \( f^{(-2)} \) is convex; the remaining assertions are obvious. \( \square \)

22.2. **Proof of the main result.** In this part, we reproduce Boas and Widder’s proof of Theorem \(22.1\). We first make a few clarifying remarks about this proof.

1. As Boas and Widder mention, Theorem \(22.1\) was shown earlier by T. Popoviciu (Mathematica, 1934) via an alternate argument using divided differences involving unequally spaced points. Here we will only explain Boas and Widder’s proof.
(2) There is a minor error in the arguments of Boas and Widder, which is resolved by adding one word. See Remark [22.11] and the proof of Lemma [22.13] for more details. (There are other minor typos in the writing of Lemmas [22.6] and [22.10] and in some of the proofs; these are corrected in the exposition below.)

(3) Boas and Widder do not explicitly write out a proof of the convexity of $f$ (in the case $k = 2$). This is addressed below as well.

Notice that Theorem [22.1] follows for the case of unbounded domain $I$ from that for bounded domains, so we assume henceforth that

$$I = (a, b), \quad \text{with } -\infty < a < b < \infty.$$ 

We now reproduce a sequence of fourteen lemmas shown by Boas and Widder, which culminate in the above theorem. These lemmas are numbered “Lemma [22.1]”, . . . , “Lemma [22.14]” and will be referred to only in this Appendix. The rest of the results, equations, and remarks – starting from Theorem [22.1] and ending with Proposition [22.12] – are numbered using a common (separate) counter. None of the results in this Appendix are cited elsewhere in the text.

The first of the fourteen lemmas by Boas and Widder says that if the $k$th order “equi-spaced” forward differences are non-negative, then so are the $k$th order “possibly non-equi-spaced” differences (the converse is immediate):

**Lemma 22.1.** If $f(x)$ satisfies $(H_k)$ in $(a, b)$ for some $k \geq 2$, then for any $k$ positive numbers $\delta_1, \ldots, \delta_k > 0$,

\[
\Delta_{\delta_1} \Delta_{\delta_2} \cdots \Delta_{\delta_k} f(x) \geq 0, \quad \text{whenever } a < x < x + \delta_1 + \delta_2 + \cdots + \delta_k < b.
\]

**Proof.** The key step is to prove using $(H_k)$ that

\[
\Delta_{\delta_1}^{-1} f(x) \geq 0, \quad \text{whenever } a < x < x + (k - 1)h + \delta_1 < b. \quad (22.4)
\]

After this, the lemma is proved using induction on $k \geq 2$. Indeed, (22.4) is precisely the assertion in the base case $k = 2$; and using (22.4) we can show the induction step as follows: for a fixed $\delta_1 \in (0, b - a)$, it follows that $\Delta_{\delta_1} f$ satisfies $(H_{k-1})$ in the interval $(a, b - \delta_1)$.

Therefore,

\[
\Delta_{\delta_2} \cdots \Delta_{\delta_k} (\Delta_{\delta_1} f(x)) \geq 0, \quad \text{whenever } a < x < x + \delta_1 + \cdots + \delta_k < b.
\]

Since the $\Delta_{\delta_j}$ commute, and since $\delta_1$ was arbitrary, the induction step follows.

Thus, it remains to show (22.4). Let $h > 0$ and $n \in \mathbb{N}$ be such that $a < x < x + h/n + (k - 1)h < b$. One checks using an easy telescoping computation that

\[
\Delta_h f(x) = \sum_{i=0}^{n-1} \Delta_{h/n} f(x + ih/n);
\]

iterating this procedure, we obtain:

\[
\Delta_{h}^{k-1} f(x) = \sum_{i_1=0}^{n-1} \cdots \sum_{i_{k-1}=0}^{n-1} \Delta_{h/n}^{k-1} f(x + [i_1 + \cdots + i_{k-1}]h/n). \quad (22.5)
\]

(This works by induction on $k \geq 2$: the previous telescoping identity is the base case for $k = 2$, and for the induction step we evaluate the innermost sum using the base case.)
From the above computations, it further follows that
\[ \Delta_{h/n} \Delta_h^{k-1} f(x) = \sum_{i_1=0}^{n-1} \ldots \sum_{i_{k-1}=0}^{n-1} \Delta_{h/n}^k f(x + [i_1 + \ldots + i_{k-1}]h/n) \geq 0, \]
where the final inequality uses the assumption \((H_k)\). From this it follows that \( \Delta_h^{k-1} f(x) \leq \Delta_h^{k-1} f(x + h/n). \)

Now suppose \( x \) is such that \( a < x < x + mh/n + (k-1)h < b \). Applying the preceding inequality to \( x, x + h/n, \ldots, x + (m-1)h/n \), we obtain:
\[ \Delta_h^{k-1} f(x) \leq \Delta_h^{k-1} f(x + h/n) \leq \cdots \leq \Delta_h^{k-1} f(x + mh/n). \quad (22.6) \]

We can now prove \((22.4)\). As in it, choose \( \delta_1 > 0 \) such that \( a < x < x + \delta_1 + (k-1)h < b \); and choose sequences \( m_j, n_j \) of positive integers such that \( m_j/n_j \to \delta_1/h \) and \( x + m_jh/n_j + (k-1)h < b \) for all \( j \geq 1 \).

Since \( f(x) \) is continuous, \( f(x + m_jh/n_j) \) converges to \( f(x + \delta_1) \), and \( \Delta_h^{k-1} f(x + m_jh/n_j) \) to \( \Delta_h^{k-1} f(x + \delta_1) \), as \( j \to \infty \). Hence using \((22.6)\) with \( m_j, n_j \) in place of \( m, n \) respectively, we obtain by taking limits:
\[ \Delta_h^{k-1} f(x) \leq \Delta_h^{k-1} f(x + \delta_1). \]

But this is equivalent to \((22.4)\), as desired. \( \square \)

**Lemma 22.2.** If \( f(x) \) satisfies \((H_k)\) in \( (a, b) \) for some \( k \geq 2 \), then \( \Delta_h^{k-1} f(x) \) and \( \Delta_h^{k-1} f(x-\epsilon) \) are non-decreasing functions of \( x \) in \( (a, b - (k-1)\epsilon) \) and \( (a + \epsilon, b - (k-2)\epsilon) \) respectively.

**Proof.** For the first part, suppose \( y < z \) are points in \( (a, b - (k-1)\epsilon) \), and set
\[ \delta_1 := z - y, \quad \delta_2 = \cdots = \delta_k := \epsilon. \]

Then by Lemma 22.1 or simply \((22.4)\) it follows that
\[ \Delta_h^{k-1} f(z) - \Delta_h^{k-1} f(y) = \Delta_{\delta_1} \Delta_h^{k-1} f(y) \geq 0, \]
which is what was asserted.

Similarly, for the second part we suppose \( y < z \) are points in \( (a + \epsilon, b - (k-2)\epsilon) \). Then \( y - \epsilon < z - \epsilon \) are points in \( (a, b - (k-1)\epsilon) \), so we are done by the first part. (Remark: Boas and Widder repeat the computations of the first part in this second part; but this is not required.) \( \square \)

We assume for the next four lemmas that \( f \) satisfies \((H_2)\) in the interval \( x \in (a, b) \).

**Lemma 22.3.** Suppose \( f \) satisfies \((H_2)\) in \( (a, b) \), and \( x \in (a, b) \). Then \( h^{-1}\Delta_h f(x) \) is a non-decreasing function of \( h \) in \( (a-x, b-x) \).

**Remark 22.7.** Notice that \( h = 0 \) lies in \( (a-x, b-x) \), and at this point the expression \( h^{-1}\Delta_h f(x) \) is not defined. Hence the statement of Lemma 22.3 actually says that \( h \mapsto h^{-1}\Delta_h f(x) \) is non-decreasing for \( h \) in \( (0, b-x) \) and separately for \( h \) in \( (a-x, 0) \). The latter can be reformulated as follows: since \( \Delta_{-h} f(x) = -\Delta_h f(x-h) \), Lemma 22.3 asserts that the map \( h \mapsto h^{-1}\Delta_h f(x-h) \) is a non-increasing function of \( h \) in \( (0, x-a) \).

**Proof of Lemma 22.3.** We first prove the result for \( h \in (0, b-x) \). Thus, suppose \( 0 < \epsilon < \delta < b-x \). By condition \((H_2)\), for all integers \( n \geq 2 \) we have:
\[ \Delta_{\delta/n}^2 f(x) \geq 0, \quad \Delta_{\delta/n}^2 f(x + \delta/n) \geq 0, \quad \cdots, \quad \Delta_{\delta/n}^2 f(x + (n-2)\delta/n) \geq 0 \]
\[ \implies \Delta_{\delta/n} f(x) \leq \Delta_{\delta/n} f(x + \delta/n) \leq \cdots \leq \Delta_{\delta/n} f(x + (n-1)\delta/n). \]
If $0 < m < n$, then the average of the first $m$ terms here cannot exceed the average of all $n$ terms. Therefore,

$$\frac{f(x + m\delta/n) - f(x)}{m\delta/n} \leq \frac{f(x + \delta) - f(x)}{\delta}.$$

Now since $\epsilon \in (0, \delta)$, choose integer sequences $0 < m_j < n_j$ such that $m_j/n_j \to \epsilon/\delta$ as $j \to \infty$. Applying the preceding inequality (with $m, n$ replaced respectively by $m_j, n_j$) and taking limits, it follows that $\epsilon^{-1}\Delta_\epsilon f(x) \leq \delta^{-1}\Delta_\delta f(x)$, since $f$ is continuous. This proves the first part of the lemma, for positive $h$.

The proof for negative $h \in (a - x, 0)$ is similar, and is shown using the reformulation of the assertion in Remark 22.7. Given $0 < \epsilon < \delta < x - a$, by condition (H2) it follows for all integers $0 < m < n$ that

$$\Delta_{\delta/n} f(x - \delta) \leq \Delta_{\delta/n} f(x - (n - 1)\delta/n) \leq \cdots \leq \Delta_{\delta/n} f(x - \delta/n)$$

$$\implies \frac{f(x) - f(x - \delta)}{\delta} \leq \frac{f(x) - f(x - m\delta/n)}{m\delta/n},$$

this time using the last $m$ terms instead of the first. Now work as above: using integer sequences $0 < m_j < n_j$ such that $m_j/n_j \to \epsilon/\delta$, it follows from the continuity of $f$ that $\delta^{-1}\Delta_\delta f(x - \delta) \leq \epsilon^{-1}\Delta_\epsilon f(x - \epsilon)$, as desired. \hfill \Box

We next define the one-sided derivatives of functions.

**Definition 22.8.** Let $f$ be a real-valued function on $(a, b)$. Define:

$$f'_+(x) := \lim_{\delta \to 0^+} \frac{\Delta_\delta f(x)}{\delta}, \quad f'_-(x) := \lim_{\delta \to 0^+} \frac{\Delta_\delta f(x)}{\delta} = \lim_{\delta \to 0^+} \frac{\Delta_\delta f(x - \delta)}{\delta}.$$

**Lemma 22.4.** Suppose $f$ satisfies (H2) in $(a, b)$. Then $f'_+, f'_-$ exist and are finite and non-decreasing on all of $(a, b)$.

**Proof.** That $f'_+$ exist on $(a, b)$ follows from Lemma 22.3, though the limits may possibly be infinite. Now fix scalars $\delta, \epsilon, x, y, z$ satisfying:

$$0 < \delta < \epsilon \quad \text{and} \quad a < z - \epsilon < x - \epsilon < x < x + \epsilon < y + \epsilon < b,$$

which implies that $a < z < x < y < b$. Then we have:

$$\frac{\Delta_\epsilon f(z - \epsilon)}{\epsilon} \leq \frac{\Delta_\epsilon f(x - \epsilon)}{\epsilon} \leq \frac{\Delta_\epsilon f(x - \delta)}{\epsilon} \leq \frac{\Delta_\epsilon f(x)}{\epsilon} \leq \frac{\Delta_\epsilon f(y)}{\epsilon},$$

where the five inequalities follow respectively using Lemma 22.2, Remark 22.7, Lemma 22.2, Lemma 22.3, and Lemma 22.2.

Now let $\delta \to 0^+$ keeping $x, y, z$ fixed; this yields:

$$\frac{\Delta_\epsilon f(z - \epsilon)}{\epsilon} \leq f'_-(x) \leq f'_+(x) \leq \frac{\Delta_\epsilon f(y)}{\epsilon},$$

which implies that $f'_\pm(x)$ are finite on $(a, b)$. In turn, letting $\epsilon \to 0^+$ yields:

$$f'_-(z) \leq f'_-(x) \leq f'_+(x) \leq f'_+(y),$$

which shows that $f'_\pm$ are non-decreasing on $(a, b)$. \hfill \Box

**Lemma 22.5.** If $f$ satisfies (H2) in $(a, b)$ then $f$ approaches a limit in $(-\infty, +\infty]$ as $x$ goes to $a^+$ and $x$ goes to $b^-$. 

Proof. Note by Lemma 22.2 that $\Delta_\delta f(x)$ is non-decreasing in $x \in (a, b - \delta)$. Hence $\lim_{x \to a^+} \Delta_\delta f(x)$ exists or equals $-\infty$. (The key point is that it is not $+\infty$.) Therefore, since $f$ is continuous,

$$ +\infty > \lim_{x \to a^+} \Delta_\delta f(x) = \lim_{x \to a^+} (f(x + \delta) - f(x)) = f(a + \delta) - f(a^+) . $$

It follows that $f(a^+)$ exists and cannot equal $-\infty$.

By the same reasoning, the limit $\lim_{x \to (b - \delta)^-} \Delta_\delta f(x)$ exists or equals $+\infty$, whence

$$ -\infty < \lim_{x \to (b - \delta)^-} \Delta_\delta f(x) = f(b^-) - f(b - \delta) . $$

It follows that $f(b^-)$ exists and cannot equal $-\infty$. \qed

Lemma 22.6. Suppose $f(x)$ satisfies $\{H_k\}$ in $(a, b)$.

(1) If $f(a^+) < +\infty$, define $f(a) := f(a^+)$. Then $f'_+(a)$ exists and is finite or $-\infty$.

(2) If $f(b^-) < +\infty$, define $f(b) := f(b^-)$. Then $f'_-(b)$ exists and is finite or $+\infty$.

Proof. First, if $f(a^+)$ or $f(b^-)$ are not $+\infty$ then they are finite by Lemma 22.5. We now prove (1). By Lemma 22.3, for $h \in (0, b - a)$ the map $h \mapsto h^{-1}\Delta_h f(a)$ is non-decreasing. Therefore $h \mapsto h^{-1}\Delta_h f(a)$ is the limit of a set of non-decreasing functions in $h$, whence it too is non-decreasing in $h$. This proves (1).

The second part is proved similarly, using that $h \mapsto h^{-1}\Delta_h f(b - h)$ is a non-increasing function in $h$. \qed

Common hypothesis for Lemmas 7–14: $f$ satisfies $\{H_k\}$ in $(a, b)$, for some $k > 2$. (We use this hypothesis below without mention or any other reference.)

Lemma 22.7. For any $a < x < b$, the map $h \mapsto h^{-k+1}\Delta_h^{k-1} f(x)$ is a non-decreasing function of $h$ in $(0, (b - x)/(k - 1))$.

Proof. First note that the given map is indeed well-defined. Now we prove the result by induction on $k \ge 2$; the following argument is similar in spirit to (for instance) computing by induction the derivative of $x^{k-1}$.

For $k = 2$ the result follows from Lemma 22.3. To show the induction step, given fixed $0 < h < (b - a)/(k - 2)$ and $\delta \in (0, b - a)$, it is clear by Lemma 22.1 that if $f$ satisfies $\{H_k\}$ in $(a, b)$, then we have, respectively:

$$ \Delta_h^{k-2} f \text{ satisfies } \{H_2\} \text{ in } (a, b - (k - 2)h), $$

$$ \Delta_\delta f \text{ satisfies } \{H_{k-1}\} \text{ in } (a, b - \delta). \quad (22.9) $$

In particular, if $0 < \delta < \epsilon < (b - x)/(k - 1)$, then we have:

$$ \frac{\Delta_\delta \Delta_\epsilon^{k-2} f(x)}{\epsilon^{k-2} \delta} \ge \frac{\Delta_\delta \epsilon^{k-2} f(x)}{\epsilon^{k-2} \delta} = \frac{\Delta_\delta \epsilon^{k-2} \Delta_\delta f(x)}{\epsilon^{k-2} \delta} \ge \frac{\Delta_\delta^{k-2} \Delta_\delta f(x)}{\delta^{k-2} \delta} . $$

Indeed, the first inequality is by the assertion for $k = 2$, which follows via Lemma 22.3 from the first condition in (22.9); and the second inequality is by the induction hypothesis (i.e., the assertion for $k - 1$) applied using the second condition in (22.9).

We saw in the preceding calculation that $\epsilon^{k-1} \Delta_\epsilon^{k-1} f(x) \ge \delta^{k-1} \Delta_\delta^{k-1} f(x)$. But this is precisely the induction step. \qed

Lemma 22.8. There is a point $c \in [a, b]$, such that $f(x)$ satisfies $\{H_{k-1}\}$ in $(c, b)$ and $-f(x)$ satisfies $\{H_{k-1}\}$ in $(a, c)$. 
Proof. Define subsets $A, B \subset (a, b)$ via:

$$A := \{x \in (a, b) : \Delta^{k-1}_n f(x) \geq 0 \text{ for all } \delta \in (0, (b - x)/(k - 1))\},$$

$$B := (a, b) \setminus A.$$

If both $A, B$ are nonempty, and $z \in A, y \in B$, then we claim that $y < z$. Indeed, since $y \not\in A$, there exists $0 < \epsilon < (b - y)/(k - 1)$ such that $\Delta^{k-1}_\epsilon f(y) < 0$. By Lemma 22.2 if $z' \in (a, y)$ then $\Delta^{k-1}_\epsilon f(z') < 0$, whence $z' \not\in A$. We conclude that $z > y$.

The above analysis implies the existence of $c \in [a, b]$ such that $(a, c) \subset B \subset (a, c)$ and $(c, b) \subset A \subset [c, b)$. It is also clear that $f$ satisfies $(H_{k-1})$ in $(c, b)$.

It remains to show that if $a < c$ then $-f$ satisfies $(H_{k-1})$ in $(a, c)$. We begin by defining a map $\varepsilon : (a, c) \to (0, \infty)$ as follows: for $x \in (a, c)$, there exists $\varepsilon \in (0, (c - x)/(k - 1))$ such that $\Delta^{k-1}_\varepsilon f(x) < 0$. By Lemmas 22.2 and 22.7 this implies that

$$\Delta^{k-1}_\delta f(y) < 0, \quad \forall a < y \leq x, \quad 0 < \delta \leq \epsilon.$$

Now define $\varepsilon : (a, c) \to (0, \infty)$ by setting

$$\varepsilon(x) := \sup\{\delta \in (0, \frac{c-x}{k-1}) : \Delta^{k-1}_\delta f(x) < 0\}.$$

By the reasoning just described, $\varepsilon$ is a non-increasing function on $(a, c)$.

With the function $\varepsilon$ in hand, we now complete the proof by showing that $-f(x)$ satisfies $(H_{k-1})$ in $(a, c)$. Let $x \in (a, c)$ and let $h > 0$ be such that $x + (k - 1)h < c$. Choose any $y \in (x + (k - 1)h, c)$ as well as an integer $n > h/\varepsilon(y)$. It follows that $\Delta^{k-1}_h f(y) < 0$.

Now recall from Equation (22.5) that

$$\Delta^{k-1}_h f(x) = \sum_{i_1=0}^{n-1} \cdots \sum_{i_{k-1}=0}^{n-1} \Delta^{k-1}_{h/n} f(x + [i_1 + \cdots + i_{k-1}] h/n).$$

But in each summand, the argument $x + [i_1 + \cdots + i_{k-1}] h/n < y$, whence by Lemmas 22.2 and 22.7 the previous paragraph implies that each summand is negative. It follows that $\Delta^{k-1}_h f(x) < 0$. This shows that $-f(x)$ satisfies $(H_{k-1})$ in $(a, c)$, as desired, and concludes the proof. \qed

Lemma 22.9. There are points

$$a = x_0 < x_1 < \cdots < x_p = b, \quad \text{with } 1 \leq p \leq 2^{k-1},$$

such that in each interval $x_j < x < x_{j+1}$, either $f(x)$ or $-f(x)$ satisfies $(H_2)$.

This follows immediately from Lemma 22.8 by induction on $k \geq 2$.

Lemma 22.10. The derivatives $f'_\pm$ both exist and are finite on all of $(a, b)$.

We remark here that $f'_\pm$ are both needed in what follows; yet Boas and Widder completely avoid discussing $f''$ in this lemma or its proof (or in the sequel). For completeness, the proof for $f'_{-}$ is also now described.

Proof. By Lemmas 22.9 [22.4] and 22.6 the functions $f'_\pm$ exist on all of $(a, b)$, and are finite, possibly except at the points $x_1, \ldots, x_{p-1}$ in Lemma 22.9. We now show that $f'_\pm$ are finite at each of these points $x_j$.

First suppose $f'_+(x_j)$ or $f'_-(x_j)$ equals $+\infty$. Choose $\delta > 0$ small enough such that

$$x_{j-1} < x_j - (k-2)\delta < x_j < x_j + \delta < x_{j+1}.$$
Now if \( f'_+(x_j) = +\infty \), then
\[
\Delta_{\delta}^{k-1} f'_+(x_j - (k-2)\delta) = -\infty
\]
\[
\implies \lim_{h \to 0^+} \frac{1}{h} \Delta_{\delta}^{k-1} \Delta_h f(x_j - (k-2)\delta) = -\infty
\]
\[
\implies \Delta_{\delta}^{k-1} \Delta_h f(x_j - (k-2)\delta) < 0 \quad \text{for all small positive } h.
\]
But this contradicts Lemma \[22.1\]. Similarly, if \( f'_-(x_j) = +\infty \), then
\[
\Delta_{\delta}^{k-1} f'_-(x_j - (k-2)\delta) = -\infty
\]
\[
\implies \lim_{h \to 0^+} \frac{1}{h} \Delta_{\delta}^{k-1} \Delta_h f(x_j - (k-2)\delta - h) = -\infty
\]
\[
\implies \Delta_{\delta}^{k-1} \Delta_h f(x_j - (k-2)\delta - h) < 0 \quad \text{for all small positive } h,
\]
which again contradicts Lemma \[22.1\].

The other case is if \( f'_-(x_j) \) or \( f'_+(x_j) \) equals \(-\infty\). We treat the first of these sub-cases; the sub-case \( f'_-(x_j) = -\infty \) is similar. Begin as above by choosing \( \delta > 0 \) such that
\[
x_{j-1} < x_j - (k-1)\delta < x_j < x_{j+1}.
\]
Now if \( f'_+(x_j) = +\infty \), then a similar computation to above yields:
\[
\Delta_{\delta}^{k-1} f'_+(x_j - (k-1)\delta) = -\infty
\]
\[
\implies \lim_{h \to 0^+} \frac{1}{h} \Delta_{\delta}^{k-1} \Delta_h f(x_j - (k-1)\delta - h) = -\infty
\]
\[
\implies \Delta_{\delta}^{k-1} \Delta_h f(x_j - (k-1)\delta - h) < 0 \quad \text{for all small positive } h,
\]
which contradicts Lemma \[22.1\]. □

The above trick of studying \( \Delta_{\delta}^n g(y - p\delta) \) where \( p = k-1 \) or \( k-2 \) (and \( n = k-1 \), \( g = f'_\pm \) so that we deal with the \( k \)th order divided differences / derivatives of \( f \)) is a powerful one. Boas and Widder now use the same trick to further study the derivative of \( f \), and show its existence, finiteness, and continuity in Lemmas \[22.11\] and \[22.13\].

**Lemma 22.11.** \( f' \) exists and is finite on \((a,b)\).

*Proof.* We fix \( x \in (a,b) \), and work with \( \delta > 0 \) small such that \( a < a + k\delta < x < b - 2\delta < b \). Let \( p \in \{0,1,\ldots,k\} \); then
\[
0 \leq \frac{1}{\delta} \Delta_{\delta}^k f(x - p\delta) = \frac{1}{\delta} \sum_{i=0}^{k} (-1)^{k-i} \binom{k}{i} f(x + (i-p)\delta).
\]
Subtract from this the identity \( 0 = \delta^{-1} f(x)(1-1)^k = \delta^{-1} f(x) \sum_{i=0}^{k} \binom{k}{i} (-1)^{k-i} \), so that the \( i = p \) term cancels; and multiply and divide the remaining terms by \( (i-p) \) to obtain:
\[
0 \leq \frac{1}{\delta} \Delta_{\delta}^k f(x - p\delta) = \sum_{i=0, \text{ } i \neq p}^{i=k} (-1)^{k-i} \binom{k}{i} \frac{f(x + (i-p)\delta) - f(x)}{(i-p)\delta} (i-p).
\]
Letting \( \delta \to 0^+ \), it follows that
\[
A_p f'_-(x) + B_p f'_+(x) \geq 0,
\]
where
\[
A_p := \sum_{i=0}^{p-1} (-1)^{k-i} \binom{k}{i} (i - p),
\]
and
\[
B_p := \sum_{i=p+1}^{k} (-1)^{k-i} \binom{k}{i} (i - p);
\]

note here that
\[
A_p + B_p = \sum_{i=0}^{k} (-1)^{k-i} \binom{k}{i} (i - p) = k \sum_{i=1}^{k} (-1)^{k-i} \binom{k-1}{i-1} - p \sum_{i=0}^{k} (-1)^{k-i} \binom{k}{i} = 0.
\]

Now specialize \( p \) to be \( k - 1 \) and \( k - 2 \). In the former case \( B_p = 1 \), whence \( A_p = -1 \), and by (22.10) we obtain: \( f'_+(x) \geq f'_-(x) \). In the latter case \( p = k - 2 \) (with \( k > 2 \)), we have \( B_p = 2 - k < 0 \). Thus \( A_p = k - 2 > 0 \), and by (22.10) we obtain: \( f'_-(x) \geq f'_+(x) \). Therefore \( f'(x) \) exists, and by Lemma 22.10 it is finite. \( \square \)

**Lemma 22.12.** If \( a < x < x + (k - 1)h < b \), then \( \Delta_h^{k-1} f'(x) \geq 0 \).

**Proof.**
\[
\Delta_h^{k-1} f'(x) = \lim_{\delta \to 0^+} \frac{\Delta_h^{k-1} f(x)}{\delta},
\]
and this is non-negative by Lemma 22.1. \( \square \)

**Lemma 22.13.** \( f' \) is continuous on \((a, b)\).

**Remark 22.11.** We record here a minor typo in the Boas–Widder paper. Namely, the authors begin the proof of Lemma 22.13 by claiming that \( f' \) is monotonic. However, this is not true as stated: for any \( k > 2 \), the function \( f(x) = x^3 \) satisfies \((H_k)\) on \( I = (-1, 1) \) but \( f' \) is not monotone on \( I \). The first paragraph of the following proof addresses this issue, using that \( f' \) is piecewise monotone on \((a, b)\).

**Proof of Lemma 22.13.** By Lemmas 22.9 and 22.14, there are finitely many points \( x_j \), \( 0 \leq j \leq p \leq 2^{k-1} \), such that on each \((x_j, x_{j+1})\), \( f'_\pm = f' \) is monotone (where this last equality follows from Lemma 22.11). Thus \( f' \) is piecewise monotone on \((a, b)\).

Now define the limits
\[
f'(x\pm) := \lim_{h \to 0^+} f'(x \pm h), \quad x \in (a, b).
\]
It is clear that \( f'(x\pm) \) exists on \((a, b)\), including at each \( x_j \neq a, b \). Note that \( f'(x_j\pm) \in [-\infty, +\infty] \), while \( f'(x\pm) \in \mathbb{R} \) for all other points \( x \neq x_j \). We show first that \( f'(x+) = f'(x-) \) – where this common limit is possibly infinite – and then that \( f'(x+) = f'(x) \), which will rule out the infinitude using Lemma 22.11, and complete the proof.

For each of the two steps, we proceed as in the proof of Lemma 22.11. Begin by fixing \( x \in (a, b) \), and let \( \delta > 0 \) be such that \( a < x - k\delta < x < x + 2\delta < b \). Let \( p \in \{0, 1, \ldots, k\} \); then by Lemma 22.12
\[
0 \leq \Delta_{\delta}^{k-1} f'(x - (p - \frac{1}{2})\delta) = \sum_{i=0}^{k-1} (-1)^{k-1-i} \binom{k-1}{i} f'(x + (i - p + \frac{1}{2})\delta).
\]

Let \( \delta \to 0^+ \); then,
\[
A_p f'(x-) - A_p f'(x+) \geq 0,
\]
where
\[
A_p := \sum_{i=0}^{p-1} (-1)^{k-1-i} \binom{k-1}{i} = -\sum_{i=p}^{k-1} (-1)^{k-1-i} \binom{k-1}{i}.
\]
Now specialize \( p \) to be \( k - 1 \) and \( k - 2 \). In the former case \( A_p = -1 \), whence \( f'(x-) \leq f'(x+) \); whereas if \( p = k - 2 \) then \( A_p = k - 2 > 0 \), whence \( f'(x-) \geq f'(x+) \). These inequalities and the trichotomy of the extended real line \([-\infty, +\infty]\) imply that \( f'(x-) = f'(x+) \).

Using the same \( \delta \in ((x-a)/k, (b-x)/2) \), and \( p \in \{0, 1, \ldots, k\} \), Lemma 22.12 also implies:

\[
0 \leq \Delta_{\delta}^{k-1} f'(x - p\delta).
\]

Taking \( \delta \to 0^+ \) and using that \( f'(x-) = f'(x+) \) yields:

\[
B_p f'(x) - B_p f'(x+) \geq 0, \quad \text{where } B_p := (-1)^{k-1-p} \binom{k-1}{p} = -\sum_{i=0, i \neq p}^{k-1} (-1)^{k-1-i} \binom{k-1}{i}.
\]

Now specialize \( p \) to be \( k - 1 \) and \( k - 2 \). In the former case \( B_p = 1 \), whence \( f'(x) \geq f'(x+) \); whereas if \( p = k - 2 \) then \( B_p = 1 - k < 0 \), whence \( f'(x) \leq f'(x+) \). These inequalities imply that \( f'(x+) = f'(x-) \) equals \( f(x) \), and in particular is finite, for all \( x \in (a,b) \).

The final lemma simply combines the preceding two:

**Lemma 22.14.** \( f' \) satisfies the condition \((H_{k-1})\) in \((a,b)\).

**Proof.** This follows immediately from Lemmas 22.12 and 22.13.

Having shown the fourteen lemmas above, we conclude with:

**Proof of the Boas–Widder Theorem 22.1.** The proof is by induction on \( k \geq 2 \). The induction step is clear: use Lemma 22.14. We now show the base case of \( k = 2 \). By Lemma 22.4, the functions \( f'_+ \) exist and are non-decreasing on \((a,b)\). Moreover, \( f \) is continuous by assumption. To prove its convexity, we make use of the following basic result from one-variable calculus:

**Proposition 22.12.** Let \( f : [p, q] \to \mathbb{R} \) be a continuous function whose right-hand derivative \( f'_+ \) exists on \([p, q)\) and is Lebesgue integrable. Then,

\[
f(y) = f(p) + \int_p^y f'_+(t) \, dt, \quad \forall y \in [p, q].
\]

Proposition 22.12 applies to our function \( f \) satisfying \((H_2)\), since \( f'_+ \) is non-decreasing by Lemma 22.4, whence Lebesgue integrable. Therefore \( f(y) - f(x) = \int_x^y f'_+(t) \, dt \) for \( a < x < y < b \). Now repeat the proof of Proposition 22.3 to show that \( f \) is convex on \((a,b)\). This completes the base case of \( k = 2 \), and concludes the proof.
Part 4:
Entrywise polynomials preserving positivity in fixed dimension
23. Entrywise polynomial preservers. Horn-type necessary conditions.
Classification of sign patterns.

Part 4: Entrywise polynomials preserving positivity in fixed dimension

In the previous part, we classified the entrywise functions preserving positivity in all dimensions; these are precisely the power series with non-negative coefficients. In the previous part to that, we had classified the entrywise powers preserving positivity (as well as total positivity and total non-negativity) in fixed dimension. In this final part of the notes, we study polynomials that entrywise preserve positive semidefiniteness in fixed dimension.

Recall from the Schur product theorem 3.8 and its converse, the Schoenberg–Rudin theorem 15.2, that the only polynomials that entrywise preserve positivity in all dimensions are the ones with all non-negative coefficients. Thus, if one fixes the dimension $N \geq 3$ of the test set of positive matrices, then it is reasonable to expect that there should exist more polynomial preservers – in other words, polynomial preservers with negative coefficients. However, this problem remained completely open until very recently ($\sim 2016$): not a single polynomial preserver was known with a negative coefficient, nor was a non-existence result proved!

In this final part, we will answer this existence question, as well as stronger variants of it. Namely, not only do we produce such polynomial preservers, we also fully resolve the more challenging question: which coefficients of polynomial preservers on $N \times N$ matrices can be negative? Looking ahead in this chapter:

- We classify the sign patterns of entrywise polynomial preservers on $P_N$, for fixed $N$.
- We extend this to all power series; but also countable sums of real powers, such as $\sum_{\alpha \in \mathbb{Q}} c_\alpha x^\alpha$. This case is more subtle than that of polynomial preservers.
- We will also completely classify the sign patterns of polynomials that entrywise preserve Hankel totally non-negative matrices of a fixed dimension. Recall from the discussions around Theorems 12.11 and 18.1 that this is expected to be similar to (if not exactly the same as) the classification for positivity preservers.

In what follows, we work with $P_N((0, \rho))$ for $N > 0$ fixed and $0 < \rho < \infty$. Since we work with polynomials and power series, this is equivalent to working over $P_N((0, \rho))$, by density and continuity. If $\rho = +\infty$, one can prove results that are similar to the ones shown below; but for a ‘first look’ at the proofs and techniques used, we restrict ourselves to $P_N((0, \rho))$.

For full details of the $\rho = +\infty$ case, as well as ramifications and applications of the results below, we refer the reader to the 2017 preprint “On the sign patterns of entrywise positivity preservers in fixed dimension” by Khare and Tao.

23.1. Horn-type necessary condition; matrices with negative entries. In this section and beyond, we work with polynomials or power series

$$f(x) = c_{n_0} x^{n_0} + c_{n_1} x^{n_1} + \cdots, \quad \text{with } n_0 < n_1 < \cdots,$$

(23.1)

and $c_{n_j} \in \mathbb{R}$ typically non-zero. Recall the (stronger) Horn–Loewner theorem 16.1 which shows that if $f \in C^{(N-1)}(I)$ for $I = (0, \infty)$, and $f[-]$ preserves positivity on (rank two Hankel TN matrices in) $P_N(I)$, then $f, f', \ldots, f^{(N-1)} \geq 0$ on $I$. In the special case that $f$ is a polynomial or a power series, one can say more, and under weaker assumptions:

**Lemma 23.2** (Horn-type necessary condition). Fix an integer $N > 0$. Let $\rho > 0$ and $f : (0, \rho) \to \mathbb{R}$ be of the form (23.1), and assume that $f[-]$ preserves positivity on rank-one Hankel TN matrices in $P_N((0, \rho))$. If $c_m < 0$ for some $m$, then $c_n > 0$ for at least $N$ values $n < m$. 

Remark 23.3. In both (23.1) as well as Lemma 23.2 (and its proof below), we have deliberately not insisted on the exponents $n_j$ being non-negative integers. In fact, one can choose \{\{n_j : j \geq 0\} to be an arbitrary increasing sequence of real numbers.

Proof of Lemma 23.2. Suppose the result is false. Then \( f(x) = \sum_{j=0}^{m} c_{n_j} x^{n_j} \), with \( m = n_k \) for some \( 0 \leq k < N \) and \( c_{n_0}, \ldots, c_{n_k-1} > 0 \). Choose any \( u_0 \in (0,1) \) and define \( u := (1, u_0, \ldots, u_0^{N-1})^T \in \mathbb{R}^N \). Then \( u^{n_0}, \ldots, u^{n_k} \) are linearly independent, forming (some of) the columns of a generalized Vandermonde matrix. Hence there exists \( v \in \mathbb{R}^N \) such that

\[
\begin{align*}
    v^T u^{n_0}, \ldots, v^T u^{n_k-1} & \quad \text{and} \quad v^T u^{n_0} = v^T u^{n_k} = 1.
\end{align*}
\]

For \( 0 \leq \epsilon < \rho \), we let \( A_{\epsilon} := cuu^T \), which is a rank-one Hankel moment matrix in \( \mathbb{P}_N((0,\rho)) \) (and hence TN). Now compute using the hypotheses:

\[
0 \leq v^T f[A_{\epsilon}] v = v^T \left( \sum_{j=0}^{\infty} c_{n_j} \epsilon^{n_j} (u^{n_j}(u^{n_0}))^T \right) v = c_{n_k} \epsilon^{n_k} + o(\epsilon^{n_k}).
\]

Thus \( 0 \leq \lim_{\epsilon \to 0^+} \frac{v^T f[A_{\epsilon}] v}{\epsilon^{n_k}} < 0 \), which is a contradiction. Hence \( k \geq N \), proving the claim. \( \square \)

Thus, by Lemma 23.2, any entrywise polynomial preserver of positivity on \( \mathbb{P}_N((0,\rho)) \) must have its \( N \) non-zero Maclaurin coefficients of ‘lowest degree’ to be positive. The obvious question is if any of the other terms can be negative, e.g. the immediate next coefficient.

We tackle this question in the remainder of these notes, and show that in fact every other coefficient can indeed be negative. For now, we point out that working with positive matrices with other entries can not provide such a structured answer (in the flavor of Lemma 23.2).

As a simple example, consider the family of polynomials

\[
p_{k,t}(x) := t(1 + x^2 + \cdots + x^{2k}) - x^{2k+1}, \quad t > 0,
\]

where \( k \geq 0 \) is an integer. We claim that \( p_{k,t}[-] \) can never preserve positivity on \( \mathbb{P}_N((-\rho,\rho)) \) for \( N \geq 2 \). Indeed, if \( u := (1, -1, 0, \ldots, 0)^T \) and \( A := (\rho/2) uu^T \in \mathbb{P}_N((-\rho,\rho)) \), then

\[
u^T p_{k,t}[A] u = -4(\rho/2)^{2k+1} < 0.
\]

Therefore \( p_{k,t}[A] \) is not positive semidefinite for any \( k \geq 0 \). If one allows complex entries, similar examples with higher-order roots of unity can be constructed, in which such negative results (compared to Lemma 23.2) can be obtained.

23.2. Classification of sign patterns for polynomials. In light of the above discussion, henceforth we restrict ourselves to working with matrices in \( \mathbb{P}_N((0,\rho)) \) for \( 0 < \rho < \infty \). By Lemma 23.2, every polynomial preserver on \( \mathbb{P}_N((0,\rho)) \) must have its \( N \) lowest-degree Maclaurin coefficients (which are nonzero) to be positive.

We are interested in understanding if any (or every) other coefficient can be negative. If say the next lowest-degree coefficient could be negative, this would achieve two goals:

- It would provide (the first example of) a polynomial preserver in fixed dimension, which has a negative Maclaurin coefficient.
- It would provide (the first example of) a polynomial that preserves positivity on \( \mathbb{P}_N((0,\rho)) \), but necessarily not on \( \mathbb{P}_{N+1}((0,\rho)) \). In particular, this would show that the Horn-type necessary condition in Lemma 23.2 is “best possible”. (See Remark 16.3 in the parallel setting of entrywise power preservers, for the original Horn condition.)

We show in this part of the notes that these goals are indeed achieved:
23. Entrywise polynomial preservers. Horn-type necessary conditions.
Classification of sign patterns. 127

Theorem 23.4 (Classification of sign patterns, fixed dimension). Fix integers $N > 0$ and $0 \leq n_0 < n_1 < \cdots < n_{N-1}$, as well as a sign $\varepsilon_M \in \{-1,0,1\}$ for each integer $M > n_{N-1}$. Given reals $\rho, c_{n_0}, c_{n_1}, \ldots, c_{n_{N-1}} > 0$, there exists a power series

$$f(x) = c_{n_0}x^{n_0} + \cdots + c_{n_{N-1}}x^{n_{N-1}} + \sum_{M>n_{N-1}} c_M x^M,$$

satisfying the following properties:

1. $f$ is convergent on $(0, \rho)$.
2. $f[-]: \mathbb{P}_N((0, \rho)) \to \mathbb{P}_N$.
3. $\text{sgn}(c_M) = \varepsilon_M$ for each $M > n_{N-1}$.

This is slightly stronger than classifying the sign patterns, in that the ‘initial coefficients’ are also specified. In fact, this result can be strengthened in two different ways; see (i) Theorem 23.7 in which the set of powers allowed is vastly more general; and (ii) Theorem 26.14 and the discussion preceding it, in which the coefficients for $M > n_{N-1}$ are also ‘specified’.

Proof. Suppose we can prove the theorem in the special case when exactly one $\varepsilon_M$ is negative. Then for each $M > n_{N-1}$, there exists $0 < \delta_M < \frac{1}{M}$ such that

$$f_M(x) := \sum_{j=0}^{N_1} c_{n_j} x^{n_j} + c_M x^M$$

preserves positivity on $\mathbb{P}_N((0, \rho))$ whenever $|c_M| \leq \delta_M$. Set $c_M := \varepsilon_M \delta_M$ for each $M > n_{N-1}$ and define $f(x) := \sum_{M>n_{N-1}} 2^{n_{N-1} - M} f_M(x)$. If $x \in (0, \rho)$, then we have

$$|f(x)| \leq \sum_{M>n_{N-1}} 2^{n_{N-1} - M} |f_M(x)| \leq \sum_{M>n_{N-1}} 2^{n_{N-1} - M} \left( \sum_{j=0}^{N-1} c_{n_j} x^{n_j} + \delta_M x^M \right)$$

$$\leq \sum_{j=0}^{N-1} c_{n_j} x^{n_j} + e^x < \infty.$$

Hence $f$ converges on $(0, \rho)$. As each $f_M[-]$ preserves positivity and $\mathbb{P}_N$ is a closed convex cone, $f[-]$ also preserves positivity. It therefore remains to show that the result holds when one coefficient is negative. But this follows from Theorem 23.5 below. \qed

Thus, it remains to show the following ‘qualitative’ result:

Theorem 23.5. Let $N > 0$ and $0 \leq n_0 < n_1 < \cdots < n_{N-1} < M$ be integers, and $\rho, c_{n_0}, c_{n_1}, \ldots, c_{n_{N-1}} > 0$ be real. Then the function $f(x) = \sum_{j=0}^{N-1} c_{n_j} x^{n_j} + c_M x^M$ entrywise preserves positivity on $\mathbb{P}_N((0, \rho))$, for some $c_M < 0$.

We will show this result in the next two sections.

23.3. Classification of sign patterns for sums of real powers. Below, after proving Theorem 23.5, we further strengthen it by proving a quantitative version – see Theorem 26.1 – which gives a sharp lower bound on $c_M$. For now, we list a special case of that result (without proof, as we show the more general Theorem 26.1 below):

Theorem 23.6. Theorem 23.5 holds even when the exponents $n_0, n_1, \ldots, n_{N-1}, M$ are real and lie in the set $\mathbb{Z}_{\geq 0} \cup [N - 2, \infty)$. 
With Theorem 23.6 in hand, it is possible to classify the sign patterns of a more general family of preservers, of the form \( f(x) = \sum_{j=0}^{\infty} c_{n_j} x^{n_j} \), where \( n_j \in \mathbb{Z}^{\geq 0} \cup \{N - 2, \infty\} \) are an arbitrary countable collection of pairwise distinct non-negative (real) exponents.

**Theorem 23.7** (Classification of sign patterns of power series preservers, fixed dimension). Let \( N \geq 2 \) and let \( n_0, n_1, \ldots \) be a sequence of pairwise distinct real numbers in \( \mathbb{Z}^{\geq 0} \cup \{N - 2, \infty\} \). For each \( j \geq 0 \), let \( \varepsilon_j \in \{-1, 0, 1\} \) be a sign such that whenever \( \varepsilon_{j_0} = -1 \), one has \( \varepsilon_j = +1 \) for at least \( N \) choices of \( j \) satisfying \( n_j < n_{j_0} \). Then for every \( \rho > 0 \), there exists a series with real exponents and real coefficients

\[
    f(x) = \sum_{j=0}^{\infty} c_{n_j} x^{n_j}
\]

which is convergent on \((0, \rho)\), which entrywise preserves positivity on \( \mathbb{P}_N((0, \rho)) \), and in which \( \text{sgn}(c_{n_j}) = \varepsilon_j \) for all \( j \geq 0 \).

That the sign patterns must satisfy the given hypotheses follows from Lemma 23.2. In particular, Theorem 23.7 shows that the Horn-type necessary condition in Lemma 23.2 remains the best possible in this generality as well.

**Remark 23.8.** A key difference between the classifications in Theorems 23.4 and 23.7 is that the latter is more flexible, since the sequence \( n_0, n_1, \ldots \) can now contain an infinite decreasing subsequence of exponents. This is more general than even Hahn or Puiseux series, not just power series. For instance, the sum may be over all rational powers in \( \mathbb{Z}^{\geq 0} \cup \{N - 2, \infty\} \).

**Proof of Theorem 23.7.** Given any set \( \{n_j : j \geq 0\} \) of (pairwise distinct) non-negative powers,

\[
    \sum_{j \geq 0} \frac{x^{n_j}}{j! |n_j|!} < \infty, \quad \forall x > 0.
\]  

(23.9)

Indeed, if we partition \( \mathbb{Z}^{\geq 0} \) into the disjoint union of \( J_k := \{ j \geq 0 : n_j \in (k-1, k]\}, k \geq 0 \), then using Tonelli’s theorem, we can estimate:

\[
    \sum_{j \geq 0} \frac{x^{n_j}}{j! |n_j|!} = \sum_{k \geq 0} \frac{1}{k!} \sum_{j \in J_k} \frac{x^{n_j}}{j!} \leq \sum_{k \geq 1} \frac{1}{k!} \sum_{j \in J_k} \frac{x^k + x^{k-1}}{j!} < e + e(x^e + x^{-1}e^x) < \infty.
\]

We now turn to the proof. Set \( J := \{ j : \varepsilon_j = -1\} \subset \mathbb{Z}^{\geq 0} \). For each \( l \in J \), by Lemma 23.2 there exist \( j_1(l), \ldots, j_N(l) \) such that \( \varepsilon_{j_k(l)} = 1 \) and \( n_{j_k(l)} < n_l \), for \( k = 1, \ldots, N \). Define

\[
    f_l(x) := \sum_{k=1}^{N} \frac{x^{n_{j_k(l)}}}{|n_{j_k(l)}|!} - \delta_l \frac{x^{n_l}}{|n_l|!},
\]

where \( \delta_l \in (0, 1) \) is chosen such that \( f_l[-] \) preserves positivity on \( \mathbb{P}_N((0, \rho)) \) by Theorem 23.6. Let \( J' \subset \mathbb{Z}^{\geq 0} \) consist of all \( j \geq 0 \) such that \( \varepsilon_j = +1 \) but \( j \neq j_l(l) \) for any \( l \in J, k \in [1, N] \). Finally, define

\[
    f(x) := \sum_{l \in J} f_l(x) \frac{x^l}{l!} + \sum_{j \in J'} \frac{x^{n_j}}{j! |n_j|!}, \quad x > 0.
\]

Repeating the calculation in (23.9), one can verify that \( f \) converges absolutely on \((0, \infty)\) and hence on \((0, \rho)\). By the above hypotheses and the Schur product theorem, it follows that \( f[-] \) preserves positivity on \( \mathbb{P}_N((0, \rho)) \).

Our goal in this section and the next is to prove the ‘qualitative’ Theorem 23.5 in the
previous section. Thus, we work with polynomials of the form
\[ f(x) = \sum_{j=0}^{N-1} c_{nj} x^{nj} + c_M x^M, \]
where \( N > 0 \) and \( 0 \leq n_0 < n_1 < \ldots < n_{N-1} < M \) are integers, and \( \rho, c_{n_0}, c_{n_1}, \ldots, c_{n_{N-1}} > 0 \)
are real.

24.1. Basic properties of Schur polynomials. We begin by defining the key tool required
in this section: Schur polynomials. We will then use these functions – via the Cauchy–Binet
identity – to understand when polynomials of the above form entrywise preserve positivity
on a ‘generic’ rank-one matrix in \( \mathbb{P}_N((0, \rho)) \).

Definition 24.1. Fix integers \( m, N > 0 \), and define \( n_{\text{min}} := (0, 1, \ldots, N-1) \). Now suppose
\( 0 \leq n'_0 \leq n'_1 \leq \cdots \leq n'_{N-1} \) are also integers.

1. A column-strict Young tableau, with shape \( n' := (n'_0, n'_1, \ldots, n'_{N-1}) \) and cell entries
   \( 1, 2, \ldots, m \), is a left-aligned two-dimensional rectangular array \( T \) of cells, with \( n'_0 \) cells
   in the bottom row, \( n'_1 \) cells in the second lowest row, and so on, such that:
   - Each cell in \( T \) has integer entry \( j \) with \( 1 \leq j \leq m \).
   - Entries weakly decrease in each row, from left to right.
   - Entries strictly decrease in each column, from top to bottom.

2. Given variables \( u_1, u_2, \ldots, u_m \) and a column-strict Young tableau \( T \) as above, define
   its weight to be
   \[ \text{wt}(T) := \prod_{j=1}^{m} u_j^{f_j}, \]
   where \( f_j \) equals the number of cells in \( T \) with entry \( j \).

3. Given an increasing sequence of integers \( 0 \leq n_0 < \cdots < n_{N-1} \), define the tuple
   \( n := (n_0, n_1, \ldots, n_{N-1}) \), and the corresponding Schur polynomial over \( u := (u_1, u_2, \ldots, u_m)^T \) to be
   \[ s_n(u) := \sum_T \text{wt}(T), \quad (24.2) \]
   where \( T \) runs over all column-strict Young tableaux of shape \( n' := n - n_{\text{min}} \) with
   cell entries \( 1, 2, \ldots, m \). (We will also abuse notation slightly and write \( s_n(u) = s_n(u_1, \ldots, u_m) \) on occasion.)

Example 24.3. Suppose \( N = m = 3 \), and \( n = (0, 2, 4) \). The column-strict Young Tableaux
with shape \( n - n_{\text{min}} = (0, 1, 2) \) and cell entries \( 1, 2, 3 \) are:

\[
\begin{array}{ccc}
3 & 3 & 3 \\
2 & 1 & 2 \\
\end{array} \quad \begin{array}{ccc}
3 & 3 & 3 \\
2 & 1 & 1 \\
\end{array} \quad \begin{array}{ccc}
3 & 2 & 2 \\
1 & 1 & 1 \\
\end{array} \quad \begin{array}{ccc}
3 & 1 & 1 \\
2 & 1 & 1 \\
\end{array} \quad \begin{array}{ccc}
2 & 2 & 1 \\
2 & 1 & 1 \\
\end{array}
\]

As a consequence,
\[
s_{(0,2,4)}(u_1, u_2, u_3) = u_3^2 u_2 + u_3^2 u_1 + u_3 u_2^2 + 2u_3 u_2 u_1 + u_3 u_1^2 + u_2^2 u_1 + u_2 u_1^2
= (u_1 + u_2)(u_2 + u_3)(u_3 + u_1).
\]
Remark 24.4. A visible notational distinction with the literature is that column-strict tableaux traditionally have entries that are \textit{increasing} down columns, and weakly increasing as one moves across rows. Since we only work with sets of tableaux through the sums of their weights occurring in Schur polynomials, this distinction is unimportant in these notes, for the following reason: define an involutive bijection \( \iota : j \mapsto m + 1 - j \), where \( \{1, \ldots, m\} \) is the alphabet of possible cell entries. Then the column-strict Young tableaux in our notation bijectively correspond under \( \iota \) – applied to each cell entry – to the ‘usual’ column-strict Young tableaux (in the literature); and as Schur polynomials are symmetric under permuting the variables by \( \iota \) (see Proposition 24.6), the sums of weights of the two sets of tableaux coincide.

Remark 24.5. Schur polynomials are fundamental objects in type \( A \) representation theory (of the general linear group, or the special linear Lie algebra), and are characters of irreducible finite-dimensional polynomial representations (over fields of characteristic zero). The above example 24.3 is a special case, corresponding to the adjoint representation for the Lie algebra of \( 3 \times 3 \) traceless matrices. This interpretation will not be used below.

Schur polynomials are always homogeneous – and also symmetric, because they can be written as a quotient of two generalized Vandermonde determinants. This is Cauchy’s definition; the definition using Young tableaux is by Littlewood. One can show that these two definitions are equivalent, among other basic properties:

\textbf{Proposition 24.6.} Fix integers \( m = N > 0 \) and \( 0 \leq n_0 < n_1 < \cdots < n_{N-1} \).

1. \textit{(Cauchy’s definition.)} If \( \mathbb{F} \) is a field and \( \mathbf{u} = (u_1, \ldots, u_N)^T \in \mathbb{F}^N \), then
\[
\det(u_1^{n_0} | u_2^{n_1} | \ldots | u_N^{n_{N-1}})_{N \times N} = V(\mathbf{u})s_n(\mathbf{u}),
\]
where \( V(\mathbf{u}) := \prod_{1 \leq j < k \leq N} (u_k - u_j) \) is the Vandermonde determinant as in (16.5). In particular, \( s_n(\mathbf{u}) \) is symmetric and homogeneous of degree \( \sum_{j=0}^{N-1} (n_j - j) \).

2. \textit{(Principal Specialization Formula.)} For any \( q \in \mathbb{F} \) that is not a root of unity or else has order \( \geq N \), we have
\[
s_n(1, q, q^2, \ldots, q^{N-1}) = \prod_{0 \leq j < k \leq N-1} \frac{q^{n_k} - q^{n_j}}{q^k - q^j}.
\]

3. \textit{(Weyl Dimension Formula.)} Specialized to \( q = 1 \), we have:
\[
s_n(1, 1, \ldots, 1) = \frac{V(\mathbf{n})}{V(\mathbf{n}_{\text{min}})} \in \mathbb{N}.
\]
In particular, there are \( V(\mathbf{n})/V(\mathbf{n}_{\text{min}}) \) column-strict tableaux of shape \( \mathbf{n} - \mathbf{n}_{\text{min}} \) and cell entries \( 1, \ldots, N \). Here and below, we will mildly abuse notation and write \( V(\mathbf{n}) \) for a tuple/row vector \( \mathbf{n} \) to denote \( V(\mathbf{n}^T) \).

\textit{Proof.} The first part is proved in Theorem 27.1 below. Using this, we show the second part. Set \( \mathbf{u} := (1, q, q^2, \ldots, q^{N-1})^T \) with \( q \) as given. Then it is easy to verify that
\[
s_n(\mathbf{u}) = \frac{\det(u_1^{n_0} | u_2^{n_1} | \ldots | u_N^{n_{N-1}})}{V(\mathbf{u})} = \frac{V((q_0, \ldots, q_{N-1}))}{V((1, q^0, \ldots, q^{N-1}))} = \prod_{0 \leq j < k \leq N-1} \frac{q^{n_k} - q^{n_j}}{q^k - q^j},
\]
as desired.

Finally, to prove the Weyl Dimension Formula, notice that by the first part, the Schur polynomial has integer coefficients and hence makes sense over \( \mathbb{Z} \), and then specializes
to \( s_n(\mathbf{u}) \) over any ground field. Now work over the ground field \( \mathbb{Q} \), and let \( f_n(T) := s_n(1, T, \ldots, T^N) \in \mathbb{Z}[T] \) be the corresponding ‘principally specialized’ polynomial. Then,

\[
V((q^0, \ldots, q^{N-1})) f_n(q) = V((q^{n_0}, \ldots, q^{n_{N-1}})), \quad \forall q \in \mathbb{Q}.
\]

In particular, for every \( q \neq 1 \), dividing both sides by \((q - 1)_{1/2}^N\), we obtain:

\[
\prod_{0 \leq j < k \leq N-1} (q^{n_j} + q^{j+1} + \cdots + q^{nk-1}) - f_n(q) \prod_{0 \leq j < k \leq N-1} (q^j + q^{j+1} + \cdots + q^{k-1}) = 0,
\]

for all \( q \in \mathbb{Q} \setminus \{1\} \). This means that the left-hand side is (the specialization of) a polynomial with infinitely many roots, whence the polynomial vanishes identically on \( \mathbb{Q} \). Specializing this polynomial to \( q = 1 \) now yields the Weyl Dimension Formula:

\[
\frac{V(\mathbf{n})}{V(\mathbf{n}_{\text{min}})} = \prod_{0 \leq j < k \leq N-1} \frac{n_k-n_j}{k-j} = f_n(1) = s_n(1, \ldots, 1).
\]

The final assertion now follows from Littlewood’s definition \([24.2]\) of \( s_n(\mathbf{u}) \). \( \square \)

### 24.2. Polynomials preserving positivity on individual rank-one positive matrices.

We return to proving Theorem \([23.5]\) and hence Theorem \([23.4]\) on sign patterns. As we have shown, it suffices to prove the theorem for one higher degree (leading) term with a negative coefficient. Before proving the result in full, we tackle the following (simpler) versions. Thus, we are given a real polynomial as above:

\[
f(x) = \sum_{j=0}^{N-1} c_{n_j} x^{n_j} + c_M x^M, \quad \text{where } c_{n_j} > 0 \forall j.
\]

1. Does there exist \( c_M < 0 \) such that \( f[-] : \mathbb{P}_N((0, \rho)) \to \mathbb{P}_N? \)

   Here is a reformulation: dividing the expression for \( f(x) \) throughout by \( |c_M| = 1/t > 0 \), define

   \[
p_t(x) := t \sum_{j=0}^{N-1} c_{n_j} x^{n_j} - x^m, \quad \text{where } c_{n_j} > 0 \forall j. \quad (24.7)
   \]

   Then it is enough to ask for which \( t > 0 \) (if any) does \( p_t[-] : \mathbb{P}_N((0, \rho)) \to \mathbb{P}_N? \)

2. Here are two simplifications: Can we produce such a constant \( t > 0 \) for only the subset of rank-one matrices in \( \mathbb{P}_N((0, \rho))? \) How about for a single rank-one matrix \( \mathbf{u}\mathbf{u}^T? \)

3. An even more specific (and special) case: let \( \mathbf{u} \) be ‘generic’, in that \( \mathbf{u} \in (0, \infty)^N \) has distinct coordinates, and \( p_t \) is as above. Can one now compute all \( t > 0 \) such that

   \[
p_t[\mathbf{u}\mathbf{u}^T] \in \mathbb{P}_N? \quad \text{How about all } t > 0 \text{ such that } \det p_t[\mathbf{u}\mathbf{u}^T] \geq 0? \]

We begin by answering the last of these questions – the answer crucially uses Schur polynomials. The following result shows that in fact, \( \det p_t[\mathbf{u}\mathbf{u}^T] \geq 0 \) implies \( p_t[\mathbf{u}\mathbf{u}^T] \) is positive semidefinite!

**Proposition 24.8.** With notation as in \([24.7]\), define the vectors

\[
\mathbf{n} := (n_0, \ldots, n_{N-1}), \quad \mathbf{n}_j := (n_0, \ldots, n_{j-1}, \hat{n}_j, n_{j+1}, \ldots, n_{N-1}, M), \quad 0 \leq j < N, \quad (24.9)
\]

where \( 0 \leq n_0 < \cdots < n_{N-1} < M \). Now if the \( n_j \) and \( M \) are integers, and a vector \( \mathbf{u} \in (0, \infty)^N \) has pairwise distinct coordinates, then the following are equivalent.

1. \( p_t[\mathbf{u}\mathbf{u}^T] \) is positive semidefinite.
2. \( \det p_t[\mathbf{u}\mathbf{u}^T] \geq 0. \)
3. \( t \geq \sum_{j=0}^{N-1} c_{n_j} s_{n_j}(\mathbf{u})^2. \)

In particular, at least for ‘most’ rank-one matrices, it is possible to find polynomial pre-
servers of positivity (on that one matrix), with a negative coefficient.

The proof of Proposition 24.8 uses the following even more widely applicable equivalence
between the non-negativity of the determinant and of the entire spectrum for ‘special’ linear
pencils of matrices:

**Lemma 24.10.** Fix \( w \in \mathbb{R}^N \) and a positive definite matrix \( H \). Define \( P_t := tH - ww^T \), for \( t > 0 \). Then the following are equivalent:

1. \( P_t \) is positive semidefinite.
2. \( \det P_t \geq 0 \).
3. \( t \geq w^T H^{-1}w = 1 - \frac{\det(H - ww^T)}{\det H} \).

This lemma is naturally connected to the theory of \((generalized)\ Rayleigh\ quotients\), al-
though we do not pursue this further.

**Proof.** We show a cyclic chain of implications. That (1) \( \implies \) (2) is immediate.

(2) \( \implies \) (3) : Using the identity (2.32) from the section on Schur complements, we obtain
by taking determinants:

\[
det \begin{pmatrix} A & B \\ B' & D \end{pmatrix} = det D \cdot det(A - BD^{-1}B')
\]

whenever \( A, D \) are square matrices, with \( D \) invertible. Using this, we compute:

\[
0 \leq \det(tH - ww^T) = det \begin{pmatrix} tH & w \\ w^T & 1 \end{pmatrix} = det \begin{pmatrix} 1 & w^T \\ tH & 1 \end{pmatrix} = \det(tH) \det(1 - w^T(tH)^{-1}w).
\]

Since the last quantity is a scalar, and \( \det(tH) > 0 \) by assumption, it follows that

\[1 \geq t^{-1}(w^T H^{-1}w) \implies t \geq w^T H^{-1}w.\]

Now substitute \( t = 1 \) in the above computation, to obtain:

\[
\frac{\det(H - ww^T)}{\det H} = \det(H) \det(1 - w^T H^{-1}w)
\]

\[
\implies \frac{\det(H - ww^T)}{\det H} = 1 - w^T H^{-1}w \geq 1 - t,
\]

which implies (3).

(3) \( \implies \) (1) : It suffices to show that \( x^T P_t x \geq 0 \) for all nonzero vectors \( x \in \mathbb{R}^N \). Using a change of variables \( y = \sqrt{H} x \neq 0 \), we compute:

\[
x^T P_t x = ty^T y - (y^T \sqrt{H}^{-1}w)^2 \\
= \|y\|^2(t - ((y')^T \sqrt{H}^{-1}w)^2), \quad \text{where} \quad y' := \frac{y}{\|y\|}
\]

\[
\geq \|y\|^2(t - \|y'\|^2 \|\sqrt{H}^{-1}w\|^2) \quad \text{ (using Cauchy–Schwarz)}
\]

\[
\geq \|y\|^2(w^T H^{-1}w - w^T H^{-1}w) = 0 \quad \text{ (by assumption).} \]

We can now answer the last of the above questions on positivity preservers, for generic
rank-one matrices.
Proof of Proposition 24.8. We are interested in the following matrix and its determinant:

\[
p_t[uu^T] = t \sum_{j=0}^{N-1} c_{nj} (u^{on_j})(u^{on_j})^T - (u^{oM})(u^{oM})^T.
\]

We first work more generally: over any field \( F \), and with matrices \( uv^T \), where \( u, v \in F^N \). Thus, we study

\[
p_t[uv^T] = t \sum_{j=0}^{N-1} c_{nj} u^{on_j} (v^{on_j})^T - u^{oM} (v^{oM})^T,
\]

where \( t, c_{nj} \in F \), and \( c_{nj} \neq 0 \ \forall j \). Setting \( D = \text{diag}(tc_{n_0}, \ldots, tc_{n_{N-1}}, -1) \), we have the decomposition

\[
p_t[uv^T] = U(u) DU(v)^T, \quad \text{where } U(u)_{N \times (N+1)} := (u^{on_0} | \ldots | u^{on_{N-1}} | u^{oM}).
\]

Applying the Cauchy–Binet identity to \( A = U(u), B = DU(v)^T \), as well as Cauchy’s definition in Proposition 24.6(1), we obtain the following general determinantal identity, valid over any field:

\[
\det p_t[uv^T] = V(u)V(v)t^{N-1} \prod_{j=0}^{N-1} c_{nj} \cdot \left( s_n(u)s_n(v)t - \sum_{j=0}^{N-1} \frac{s_{nj}(u)s_{nj}(v)}{c_{nj}} \right), \tag{24.11}
\]

Now specialize this identity to \( F = \mathbb{R} \), with \( t, c_{nj} > 0 \) and \( u = v \in \mathbb{R}^N \) having distinct coordinates. From this we deduce the following consequences. First, set

\[
H := \sum_{j=0}^{N-1} c_{nj} (uu^T)^{on_j} = U'(u) D'U'(u)^T, \quad \mathbf{w} := u^{oM},
\]

where \( U'(u) := (u^{on_0} | \ldots | u^{on_{N-1}}) \) is a non-singular generalized Vandermonde matrix and \( D' := \text{diag}(c_{n_0}, \ldots, c_{n_{N-1}}) \) is positive definite. From this it follows that \( H \) is positive definite, whence Lemma 24.10 applies. Moreover, \( H - \mathbf{w} \mathbf{w}^T = p_t[uu^T] \), so using the above calculation (24.11) and the Cauchy–Binet identity respectively, we have:

\[
\det(H - \mathbf{w} \mathbf{w}^T) = V(u)^2 \prod_{j=0}^{N-1} c_{nj} \cdot s_n(u)^2 \left( 1 - \sum_{j=0}^{N-1} \frac{s_{nj}(u)^2}{c_{nj}s_n(u)^2} \right),
\]

\[
\det H = V(u)^2 \prod_{j=0}^{N-1} c_{nj} \cdot s_n(u)^2.
\]

In particular, from Lemma 24.10(3) we obtain:

\[
\mathbf{w}^T H^{-1} \mathbf{w} = \sum_{j=0}^{N-1} \frac{s_{nj}(u)^2}{c_{nj}s_n(u)^2}.
\]

Now the proposition follows directly from Lemma 24.10 since \( P_t = p_t[uu^T] \) for all \( t > 0 \). □
25. First-order approximation / leading term of Schur polynomials.
From rank-one matrices to all matrices.

In the previous section, we computed the exact threshold for the leading term of a polynomial
\[ p_t(x) := t \sum_{j=0}^{N-1} c_{n_j} x^{n_j} - x^M, \] where \( c_{n_j} > 0 \) \( \forall j \)
(and where \( 0 \leq n_0 < \cdots < n_{N-1} < M \) are integers), such that \( p_t(\mathbf{uu}^T) \in \mathbb{P}_N \) for a vector \( \mathbf{u} \in (0, \infty)^N \) with pairwise distinct coordinates. Recall that our (partial) goal is to find a threshold that works for all rank-one matrices \( \mathbf{uu}^T \in \mathbb{P}_N((0, \rho)) \) \( \rho \) \( \in (0, \sqrt{\rho})^N \).

Thus, we need to show that the supremum of the threshold over all such \( \mathbf{u} \) is bounded:
\[ \sup_{\mathbf{u} \in (0, \sqrt{\rho})^N} \sum_{j=0}^{N-1} \frac{s_{n_j}(\mathbf{u})^2}{c_{n_j}} < \infty. \]

Since we only consider vectors \( \mathbf{u} \) with positive coordinates, it suffices to bound \( s_{n_j}(\mathbf{u})/s_n(\mathbf{u}) \) from above, for each \( j \). In turn, for this it suffices to find lower and upper bounds for every Schur polynomial evaluated at \( \mathbf{u} \in (0, \infty)^N \). This is achieved by the following result:

**Theorem 25.1** (First-order approximation / Leading term of Schur polynomials). Say \( N > 0 \) and \( 0 \leq n_0 < \cdots < n_{N-1} \) are integers. Then for all real numbers \( 0 < u_1 \leq u_2 \leq \cdots \leq u_N \), we have the bounds
\[ 1 \times \mathbf{u}^{n-n_{\min}} \leq s_n(\mathbf{u}) \leq \frac{V(n)}{V(n_{\min})} \times \mathbf{u}^{n-n_{\min}}, \]
where \( \mathbf{u}^{n-n_{\min}} = u_1^{n_0} u_2^{n_1-1} \cdots u_N^{n_{N-1}-(N-1)} \), and \( V(n) \) is as in (16.5). Moreover the constants 1 and \( \frac{V(n)}{V(n_{\min})} \) cannot be improved.

**Proof.** Recall that \( s_n(\mathbf{u}) \) is obtained by summing the weights of all column-strict Young tableaux of shape \( n - n_{\min} \) with cell entries 1, \( \ldots, N \). Moreover, by the Weyl Dimension Formula in Proposition 24.6(3), there are precisely \( V(n)/V(n_{\min}) \) such tableaux. Now each such tableau can have weight at most \( \mathbf{u}^{n-n_{\min}} \), as follows: the cells in the top row each have entries at most \( N \); the cells in the next row at most \( N-1 \), and so on. The tableau \( T_{\max} \) obtained in this fashion has weight precisely \( \mathbf{u}^{n-n_{\min}} \). Hence by definition, we have:
\[ \mathbf{u}^{n-n_{\min}} = \text{wt}(T_{\max}) \leq \sum_T \text{wt}(T) = s_n(\mathbf{u}) \leq \sum_T \text{wt}(T_{\max}) = \frac{V(n)}{V(n_{\min})} \mathbf{u}^{n-n_{\min}}. \]

This proves the bounds; we claim that both bounds are sharp. If \( n = n_{\min} \) then all terms in the claimed inequalities are 1, and we are done. Thus, assume henceforth that \( n \neq n_{\min} \).

Let \( A > 1 \) and define \( \mathbf{u}(A) := (A, A^2, \ldots, A^N) \). Then \( \text{wt}(T_{\max}) = A^M \) for some \( M > 0 \). Hence for every column-strict tableau \( T \neq T_{\max} \) as above, \( \text{wt}(T) \) is at most \( \text{wt}(T_{\max})/A \) and at least 1 = \( \text{wt}(T_{\max})/A^M \). Now summing over all such tableaux \( T \) yields:
\[ s_n(\mathbf{u}(A)) \leq \mathbf{u}(A)^{n-n_{\min}} \left( 1 + \left( \frac{V(n)}{V(n_{\min})} - 1 \right) \frac{1}{A} \right), \]
\[ s_n(\mathbf{u}(A)) \geq \mathbf{u}(A)^{n-n_{\min}} \left( 1 + \left( \frac{V(n)}{V(n_{\min})} - 1 \right) \frac{1}{A^M} \right). \]

Divide throughout by \( \mathbf{u}(A)^{n-n_{\min}} \); now taking the limit as \( A \to \infty \) yields the sharp lower bound 1, while taking the limit as \( A \to 1^+ \) yields the sharp upper bound \( V(n)/V(n_{\min}) \). \( \square \)
We now use Theorem 25.1 and Proposition 24.8 in the previous section, to find a threshold for \( t > 0 \) beyond which \( p_t(\cdot) \) preserves positivity on all rank-one matrices in \( \mathbb{P}_N((0, \rho)) \) — and in fact, on all matrices in \( \mathbb{P}_N((0, \rho)) \).

**Theorem 25.2.** Fix integers \( N > 0 \) and \( 0 \leq n_0 < n_1 < \cdots < n_{N-1} < M \), and scalars \( \rho, t, c_{n_0}, \ldots, c_{n_{N-1}} > 0 \). The polynomial \( p_t(x) := t \sum_{j=0}^{N-1} c_{n_j} x^{n_j} - x^M \) entrywise preserves positivity on \( \mathbb{P}_N((0, \rho)) \), if \( t \geq t_0 := \sum_{j=0}^{N-1} \frac{V(n_j)^2}{c_{n_j} V(n_{\min})^2} \rho^{M-n_j} \).

The following notation is useful in the sequel, including in the proof of Theorem 25.2.

**Definition 25.3.** Given a subset \( S \subset \mathbb{R} \) and an integer \( N > 0 \), define the sets

\[
S^N_{\neq} := \{(s_1, \ldots, s_N) \in S^N : s_j \text{ are pairwise distinct}\},
\]

\[
S^N_{<} := \{(s_1, \ldots, s_N) \in S^N : s_1 < \cdots < s_N\}.
\]

**Proof of Theorem 25.2**. Given \( u \in (0, \sqrt{\rho})^N_{\neq} \), from Proposition 24.8 it follows that \( p_t(uu^T) \in \mathbb{P}_N \) if and only if \( t \geq \sum_{j=0}^{N-1} s_{n_j}(u)^2 \). Now suppose \( u \in (0, \sqrt{\rho})^N_{<} \). Then by Theorem 25.1,

\[
\sum_{j=0}^{N-1} \frac{s_{n_j}(u)^2}{c_{n_j}s_n(u)^2} \leq \sum_{j=0}^{N-1} \frac{u^{2(n_j-n_{\min})}/V(n_j)^2/V(n_{\min})^2}{c_{n_j}2^{2(n-n_{\min})}} = \sum_{j=0}^{N-1} \frac{V(n_j)^2}{c_{n_j}V(n_{\min})^2} \rho^{M-n_j}
\]

and this is bounded above by \( t_0 \), since if \( v := \sqrt{\rho}(1, \ldots, 1)^T \) then \( u^{2(n_j-n)} \leq v^{2(n_j-n)} = \rho^{M-n_j} \) for all \( j \). Thus, we conclude that

\[
t \geq t_0 \implies p_t(uu^T) \in \mathbb{P}_N \forall u \in (0, \sqrt{\rho})^N_{<} \implies p_t(uu^T) \in \mathbb{P}_N \forall u \in (0, \sqrt{\rho})^N_{\neq} \implies p_t(uu^T) \in \mathbb{P}_N \forall u \in (0, \sqrt{\rho})^N_{\neq},
\]

where the first implication was proved above, the second follows by (the symmetric nature of Schur polynomials and by) relabelling the rows and columns of \( uu^T \) to rearrange the entries of \( u \) in increasing order, and the third implication follows from the continuity of \( p_t \) and the density of \( (0, \sqrt{\rho})^N_{\neq} \) in \( (0, \sqrt{\rho})^N \).

This validates our claimed threshold \( t_0 \) for all rank-one matrices. To prove the result on all of \( \mathbb{P}_N((0, \rho)) \), we use induction on \( N \geq 1 \), with the base case of \( N = 1 \) already done since 1 \times 1 matrices have rank one.

For the induction step, recall the Extension Theorem 9.10 which said that: "Suppose \( I = (0, \rho) \) or (\(-\rho, \rho\)) or its closure, for some \( 0 < \rho \leq \infty \). If \( h \in C^1(I) \) is such that \( h(\cdot) \) preserves positivity on rank-one matrices in \( \mathbb{P}_N(I) \) and \( h'(\cdot) : \mathbb{P}_N(I) \to \mathbb{P}_{N-1} \), then \( h(\cdot) : \mathbb{P}_N(I) \to \mathbb{P}_{N} \).

We will apply this result to \( h(x) = p_{t_0}(x) \), with \( t_0 \) as above. By the extension theorem, we need to show that \( h'(\cdot) : \mathbb{P}_{N-1}((0, \rho)) \to \mathbb{P}_{N-1} \). Note that

\[
h'(x) = t_0 \sum_{j=0}^{N-1} n_j c_{n_j} x^{n_j-1} - M x^{M-1} = Mg(x) + t_0 n_0 c_{n_0} x^{n_0-1}.
\]

where we define

\[
g(x) := \frac{t_0}{M} \sum_{j=1}^{N-1} n_j c_{n_j} x^{n_j-1} - x^{M-1}.
\]
25. First-order approximation / leading term of Schur polynomials.

From rank-one matrices to all matrices. 137

We claim that the entrywise polynomial map \( g[-] : \mathbb{P}_{N-1}((0, \rho)) \to \mathbb{P}_{N-1} \). If this holds, then by the Schur product theorem, the same property is satisfied by \( M g(x) + t_0 n_0 c_{n_0} x^{n_0-1} \) (regardless of whether \( n_0 = 0 \) or \( n_0 > 0 \)). But this function is precisely \( h' \), and the theorem would follow.

It thus remains to prove the claim, and we do so via a series of reductions and simplifications – i.e., “working backward”. By the induction hypothesis, Theorem 25.2 holds in dimension \( N - 1 \geq 1 \), for the polynomials

\[
q_t(x) := t \sum_{j=1}^{N-1} n_j c_{n_j} x^{n_j-1} - x^{M-1}.
\]

For this family, the threshold is now given by

\[
\sum_{j=1}^{N-1} \frac{V(n'_j)^2}{n_j c_{n_j} V(n'_{\min})^2} \rho^{M-1-(n_j-1)},
\]

where

\[
n'_{\min} := (0, 1, \ldots, N-2), \quad n'_j := (n_1, \ldots, n_{j-1}, \widehat{n_j}, n_{j+1}, \ldots, n_{N-1}, M) \quad \forall j > 0.
\]

Thus, the proof is complete if we show that

\[
\sum_{j=1}^{N-1} \frac{V(n'_j)^2}{n_j c_{n_j} V(n'_{\min})^2} \rho^{M-1-(n_j-1)} \leq \frac{t_0}{M} = \sum_{j=0}^{N-1} \frac{V(n_j)^2}{M c_{n_j} V(n_{\min})^2} \rho^{M-n_j}.
\]

In turn, comparing just the \( j \)th summand for each \( j > 0 \), it suffices to show that

\[
\frac{V(n'_j)}{\sqrt{n'_j} V(n'_{\min})} \leq \frac{V(n_j)}{\sqrt{M V(n_{\min})}}, \quad \forall j > 0.
\]

Dividing the right-hand side by the left-hand side, and cancelling common factors, we obtain the expression

\[
\prod_{k=1}^{N-1} \frac{n_k - n_0}{k} \cdot \frac{\sqrt{n_j}}{M \sqrt{n_j}} \cdot \frac{M - n_0}{n_j - n_0}.
\]

Since every factor in the product term is at least 1, it remains to show that

\[
\frac{M - n_0}{n_j - n_0} \geq \frac{\sqrt{M}}{\sqrt{n_j}}, \quad \forall j > 0.
\]

But this follows from a straightforward calculation:

\[
(M - n_0)^2 n_j - (n_j - n_0)^2 M = (M - n_j)(M n_j - n_0^2) > 0,
\]

and the proof is complete. \( \square \)

Finally, we recall our original goal of classifying the sign patterns of positivity preservers in fixed dimension – see Theorem 23.4. We showed this result holds if one can prove its special case, Theorem 23.5. Now this latter result follows from Theorem 25.2, by setting

\[
c_M := -t_0^{-1}, \quad \text{where } t_0 = \sum_{j=0}^{N-1} \frac{V(n_j)^2}{c_{n_j} V(n_{\min})^2} \rho^{M-n_j} \text{ as in Theorem 25.2}.
\]

(\( \square \))

In the last few sections, we proved the existence of a negative threshold $c_M$ for polynomials of the form

$$f(x) = \sum_{j=0}^{N-1} c_{n_j} x^{n_j} + c_M x^M$$

to entrywise preserve positivity on $\mathbb{P}_N((0, \rho))$. (Here, $N > 0$ and $0 \leq n_0 < \cdots < n_{N-1} < M$ are integers.) We now compute the exact value of this threshold, more generally for real powers; this has multiple consequences which are described below. Thus, the goal in this section is to prove the following quantitative result, for real powers – including negative powers:

**Theorem 26.1.** Fix an integer $N > 0$ and real powers $n_0 < \cdots < n_{N-1} < M$. Also fix real scalars $\rho > 0$ and $c_{n_0}, \ldots, c_{n_{N-1}}, c_M$, and define

$$f(x) := \sum_{j=0}^{N-1} c_{n_j} x^{n_j} + c_M x^M. \quad (26.2)$$

Then the following are equivalent:

1. The entrywise map $f[-]$ preserves positivity on all rank-one matrices in $\mathbb{P}_N((0, \rho))$.
2. The entrywise map $f[-]$ preserves positivity on rank-one Hankel TN matrices in $\mathbb{P}_N((0, \rho))$.
3. Either all $c_{n_j}, c_M \geq 0$; or $c_{n_j} > 0 \forall j$ and $c_M \geq -C^{-1}$, where

$$C = \sum_{j=0}^{N-1} \frac{V(n_j)^2}{c_{n_j} V(n)^2} \rho^{M-n_j}. \quad (26.3)$$

Here $V(u), n, n_j$ are defined as in $(16.5)$ and $(24.9)$. If moreover we assume that $n_j \in \mathbb{Z}^2 \cup [N - 2, \infty)$ for all $j$, then the above conditions are further equivalent to the ‘full-rank’ version:

4. The entrywise map $f[-]$ preserves positivity on $\mathbb{P}_N([0, \rho])$.

Theorem 26.1 is a powerful result. It has multiple applications, some of which are now listed.

1. Suppose $M = N$ and $n_j = j$ for $0 \leq j \leq N - 1$. Then the result provides a complete characterization of which polynomials of degree $\leq N$ entrywise preserve positivity on $\mathbb{P}_N((0, \rho))$ – or more generally, on any intermediate set between $\mathbb{P}_N((0, \rho))$ and the rank-one Hankel TN matrices inside it.
2. In fact a similar result to the previous characterization is implied, whenever one considers linear combinations of at most $N + 1$ monomial powers.
3. The result provides information on positivity preservers beyond polynomials, since the powers $n_j, M$ are now allowed to be real, even negative if one works with rank-one matrices.
4. In particular, the result implies Theorem 23.6 and hence Theorem 23.7 (see its proof). This latter theorem provides a full classification of the sign patterns of possible “countable sums of real powers” which entrywise preserve positivity on $\mathbb{P}_N((0, \rho))$.
5. The result also provides information on preservers of total non-negativity on Hankel matrices in fixed dimension; see Corollary 26.11 below.
26. Exact quantitative bound: monotonicity of Schur ratios.

Real powers and power series.

(6) There are further applications, two of which are (i) to the matrix cube problem and to sharp linear matrix inequalities/spectrahedra involving Hadamard powers; and (ii) to computing the simultaneous kernels of Hadamard powers and a related “Schubert cell-type” stratification of the cone \( \mathbb{P}_N(\mathbb{C}) \). These are explained in the 2016 paper of Belton, Guillot, Khare, and Putinar in Adv. in Math.; see also the 2017 preprint by Khare and Tao (mentioned a few lines above (23.1)).

(7) Theorem 26.1 is proved using a monotonicity phenomenon for ratios of Schur polynomials; see Theorem 26.6 below. This latter result is also useful in extending a 2011 conjecture by Cuttler–Greene–Skandera (and its proof). In fact, this line of attack ends up characterizing weak majorization using Schur polynomials. See the 2017 preprint by Khare and Tao for more details.

(8) One further application is Theorem 26.14 which finds a threshold for bounding by \( \sum_{j=0}^{N-1} c_{n_j} A^{on_j} \), any power series – and more general ‘Laplace transforms’ – applied entrywise to a positive matrix \( A \). This extends Theorem 26.1 where the ‘power series’ is simply \( x^M \), because Theorem 26.1 says in particular that \( (x^M)[A] = A^{0M} \) is dominated by a multiple of \( \sum_{j=0}^{N-1} c_{n_j} A^{on_j} \).

(9) As mentioned in the remarks prior to Theorem 23.4, Theorem 26.1 also provides examples of ‘power series’ preservers on \( \mathbb{P}_N((0,\rho)) \) with negative coefficients; and of such functions which preserve positivity on \( \mathbb{P}_N((0,\rho)) \) but not on \( \mathbb{P}_{N+1}((0,\rho)) \).

26.1. Monotonicity of ratios of Schur polynomials. The proof of Theorem 26.1 uses the same ingredients as developed in previous sections. A summary of what follows is now provided. In the rank-one case, we use a variant of Proposition 24.8 for an individual matrix; the result does not apply as is, since the powers may now be real. Next, in order to find the sharp threshold for all rank-one matrices, even for real powers we crucially appeal to the result does not apply as is, since the powers may now be real. Next, in order to find the sharp threshold for all rank-one matrices, even for real powers we crucially appeal to the extension principle in Theorem 9.10, and hence to work for all matrices; the further hope would be that this threshold bound is tight enough to behave well with respect to the extension principle in Theorem 9.10 and hence to work for all matrices in \( \mathbb{P}_N((0,\rho)) \). Remarkably, these two hopes are indeed justified, proving the theorem.

We begin with the key result required to be able to take suprema over ratios of Schur polynomials \( s_m(u)/s_n(u) \). To motivate the result, here is a special case.

**Example 26.4.** Suppose \( N = 3, n = (0, 2, 3), m = (0, 2, 4) \). As above, we have \( u = (u_1, u_2, u_3)^T \) and \( \mathbf{n}_{\min} = (0, 1, 2) \). Now let \( f(u) := \frac{s_m(u)}{s_n(u)} : (0, \infty)^N \to (0, \infty) \). This is a rational function, whose numerator sums weights over tableaux of shape \((2, 1)\), and hence by Example 24.3 above, equals \((u_1 + u_2)(u_2 + u_3)(u_3 + u_1)\). The denominator sums weights over tableaux of shape \((1, 1)\); there are only three such:

\[
\begin{bmatrix}
3 \\
2 \\
3 \\
1 \\
2 \\
1
\end{bmatrix}
\]

and hence,

\[
f(u) := \frac{(u_1 + u_2)(u_2 + u_3)(u_3 + u_1)}{u_1 u_2 + u_2 u_3 + u_3 u_1}, \quad u_1, u_2, u_3 > 0.
\]

Notice that the numerator and denominator are both Schur polynomials, hence positive combinations of monomials (this is called *monomial positivity*). In particular, they are both non-decreasing in each coordinate. One can verify that their ratio \( f(u) \) is not a polynomial;
moreover, it is not \textit{a priori} clear if $f(u)$ shares the same coordinatewise monotonicity property. However, we claim that this does hold, i.e., $f(u)$ is non-decreasing in each coordinate on $u \in (0, \infty)^N$.

To see why: by symmetry, it suffices to show that $f$ is non-decreasing in $u_3$. Using the quotient rule of differentiation, we claim that the expression

$$s_n(u) \partial_{u_3} s_m(u) - s_m(u) \partial_{u_3} s_n(u)$$

is non-negative on $(0, \infty)^3$. Indeed, computing this expression yields:

$$(u_1 + u_2)(u_1 u_3 + 2u_1 u_2 + u_2 u_3)u_3,$$

and this is clearly non-negative as desired. More strongly, the expression turns out to be monomial-positive, which implies non-negativity.

Here is the punchline: an even stronger phenomenon holds. Namely, when we write the expression (26.5) in the form $\sum_{j \geq 0} p_j(u_1, u_2)u_3^j$, each polynomial $p_j$ is \textit{Schur-positive}! This means that it is a non-negative integer-linear combination of Schur polynomials:

- $p_0(u_1, u_2) = 0$,
- $p_1(u_1, u_2) = 2u_1 u_2^2 + 2u_1^2 u_2 = 2 \begin{array}{c} 2 \\ 1 \end{array} + 2 \begin{array}{c} 2 \\ 1 \end{array} = 2s_{(1,3)}(u_1, u_2),$
- $p_2(u_1, u_2) = (u_1 + u_2)^2 = 2 \begin{array}{c} 2 \\ 1 \end{array} + 2 \begin{array}{c} 2 \\ 1 \end{array} + 1 \begin{array}{c} 1 \\ 1 \end{array} + 1 \begin{array}{c} 1 \\ 1 \end{array} + 2 \begin{array}{c} 2 \\ 1 \end{array} = s_{(0,3)}(u_1, u_2) + s_{(1,2)}(u_1, u_2),$

where the mild abuse of notation is clear in its meaning. This yields the sought-for non-negativity, as each $s_n(u)$ is monomial-positive by definition.

The remarkable fact is that the phenomena described in the above example also occur for \textit{every} pair of Schur polynomials $s_m(u), s_n(u)$ for which $m \geq n$ coordinatewise:

\textbf{Theorem 26.6} (Monotonicity of Schur polynomial ratios). Suppose $0 \leq n_0 < \cdots < n_{N-1}$ and $0 \leq m_0 < \cdots < m_{N-1}$ are integers satisfying: $n_j \leq m_j \forall j$. Then the function

$$f : (0, \infty)^N \to \mathbb{R}, \quad f(u) := \frac{s_m(u)}{s_n(u)}$$

is non-decreasing in each coordinate.

More strongly, viewing the expression

$$s_n(u) \partial_{u_N} s_m(u) - s_m(u) \partial_{u_N} s_n(u)$$

as a polynomial in $u_N$, the coefficient of each monomial $u_N^j$ is a Schur-positive polynomial in $(u_1, u_2, \ldots, u_{N-1})^T$.

Theorem 26.6 is an application of a deep result in representation theory/symmetric function theory, by Lam, Postnikov, and Pylyavskyy (Amer. J. Math., 2007). The proof of this latter result is beyond the scope of these notes, and hence is not pursued further; but its usage means that in the spirit of the previous sections, the proof of Theorem 26.6 once again combines analysis with symmetric function theory.

To proceed further, we introduce the following notation:

\textbf{Definition 26.7}. Given a vector $u = (u_1, \ldots, u_m)^T \in (0, \infty)^m$ and a real tuple $n = (n_0, \ldots, n_{N-1})$ for integers $m, N > 0$, define

$$u^{\circ n} := (u^{\circ n_0} | \ldots | u^{\circ n_{N-1}})_{m \times N} = (u_j^{n_{j-1}})_{j=1, k=1}^{m, N}.$$
We now extend Theorem 26.6 to arbitrary real powers (instead of non-negative integer powers). As one can no longer use Schur polynomials, the next result uses generalized Vandermonde determinants instead:

**Proposition 26.8.** Fix real tuples \( \mathbf{n} = (n_0 < n_1 < \cdots < n_{N-1}) \) and \( \mathbf{m} = (m_0 < m_1 < \cdots < m_{N-1}) \) with \( n_j \leq m_j \forall j \). Then the function

\[
f_{\frac{\chi}{\chi}}(\mathbf{u}) := \frac{\det(\mathbf{u^m})}{\det(\mathbf{u^n})}
\]

is non-decreasing in each coordinate on \((0, \infty)^N_{\neq}\). (See Definition 25.3.)

**Proof.** For a fixed \( t \in \mathbb{R} \), if for each \( j \) we multiply the \( j \)th row of the matrix \( \mathbf{u^m} \) by \( u_j^t \), we obtain a matrix \( \mathbf{u^{m_t}} \) where \( m_j^t = m_j + t \forall j \). In particular, if we start with real powers \( n_j, m_j \), then multiplying the numerator and denominator of \( f_{\frac{\chi}{\chi}} \) by \( (u_1 \cdots u_N)^{-m_0} \) reduces the situation to working with the non-negative real tuples \( \mathbf{n}^t := (n_j - n_0)_{j=0}^{N-1} \) and \( \mathbf{m}^t := (m_j - n_0)_{j=0}^{N-1} \).

Thus, we suppose henceforth that \( n_j, m_j \geq 0 \forall j \).

First suppose \( n_j, m_j \) are all integers. Then the result is an immediate reformulation of the first part of Theorem 26.6 via Proposition 24.6(1).

Next suppose \( n_j, m_j \) are rational. Choose a (large) integer \( L > 0 \) such that \( L n_j, L m_j \in \mathbb{Z} \forall j \), and define \( y_j := u_j^{1/L} \). By the previous case (i.e., Theorem 26.6), the function

\[
f(y) := \frac{\det(\mathbf{y^{Lm}})}{\det(\mathbf{y^{Ln}})} = \frac{\det(\mathbf{u^m})}{\det(\mathbf{u^n})}, \quad y := (y_1, \ldots, y_N)^T \in (0, \infty)^N_{\neq}
\]

is coordinatewise non-decreasing on \((0, \infty)^N_{\neq}\) in the \( y_j \), whence on \((0, \infty)^N_{\neq}\) in the \( u_j \), as desired.

Finally, in the general case, given non-negative real powers \( n_j, m_j \) satisfying the hypotheses, choose sequences \( 0 \leq n_{0,k} < n_{1,k} < \cdots < n_{N-1,k} \) and \( 0 \leq m_{0,k} < m_{1,k} < \cdots < m_{N-1,k} \) for each \( k = 1, 2, \ldots, \), which further satisfy:

1. \( n_{j,k}, m_{j,k} \) are rational for \( 0 \leq j \leq N - 1, \ k \geq 1 \);
2. \( n_{j,k} \leq m_{j,k} \forall j, k \);
3. \( n_{j,k} \to n_j \) and \( m_{j,k} \to m_j \) as \( k \to \infty \), for each \( j = 0, 1, \ldots, N - 1 \).

By the rational case above, for each \( k \geq 1 \) the function

\[
f_k(\mathbf{u}) := \frac{\det(\mathbf{u^{m_k}})}{\det(\mathbf{u^{n_k}})}
\]

is coordinatewise non-decreasing, where \( \mathbf{m}_k := (m_{0,k}, \ldots, m_{N-1,k}) \) and similarly for \( \mathbf{n}_k \). But then their limit \( \lim_{k \to \infty} f_k(\mathbf{u}) = f_{\frac{\chi}{\chi}}(\mathbf{u}) \) is also coordinatewise non-decreasing, as claimed. \( \square \)

### 26.2. Proof of the quantitative bound.

Using Proposition 26.8, we can now prove the main result in this section.

**Proof of Theorem 26.1.** We first work only with rank-one matrices. Clearly (1) \( \rightarrow \) (2), and we show that (2) \( \Rightarrow \) (3) \( \Rightarrow \) (1).

If all coefficients \( c_{n_j}, c_M \geq 0 \) then \( f[-] \) preserves positivity on rank-one matrices. Otherwise, by the Horn-type necessary conditions in Lemma 23.2 (now for real powers, possibly negative!), it follows that \( c_{n_0}, \ldots, c_{n_{N-1}} > 0 > c_M \). In this case, the discussion that opens
Section 24.2 allows us to reformulate the problem using

\[ p_t(x) := t \sum_{j=0}^{N-1} c_{nj} x^{n_j} - x^M, \quad t > 0, \]

and the goal is to find a sharp positive lower bound for \( t \), above which \( p_t[-] \) preserves positivity on rank-one Hankel TN matrices \( uu^T \in \mathbb{F}_N((0, \rho)) \).

But now one can play the same game as in Section 24.2. In other words, Lemma 24.10 shows that the ‘real powers analogue’ of Proposition 24.8 holds: \( p_t[uu^T] \geq 0 \) if and only if

\[ t \geq \sum_{j=0}^{N-1} \frac{\det(u^{cn_j})^2}{c_{nj} \det(u^{cn_j})^2}, \]

for all generic rank-one matrices \( uu^T \), with \( u \in (0, \sqrt{\rho})^N \). By the same reasoning as in the proof of Theorem 25.2 (see the previous section), \( p_t[-] \) preserves positivity on a given test set of rank-one matrices \( \{uu^T : u \in S \subset (0, \sqrt{\rho})^N\} \), if and only if (by density and continuity,) \( t \) exceeds the following supremum:

\[ t \geq \sup_{S \cap (0, \sqrt{\rho})^N} \sum_{j=0}^{N-1} \frac{\det(u^{cn_j})^2}{c_{nj} \det(u^{cn_j})^2}. \tag{26.9} \]

This is, of course, subject to \( S \cap (0, \sqrt{\rho})^N \) being dense in the set \( S \), which is indeed the case if \( S \) equals the set of rank-one Hankel TN matrices as in assertion (2).

Thus, to prove (2) \( \implies \) (3) \( \implies \) (1) in the theorem, it suffices to prove: (i) the supremum (26.9) is bounded above by the value \( \sum_{j=0}^{N-1} \frac{V(n_j)^2}{c_{nj} V(n_j)^2} \rho^{M-n_j} \); and (ii) this value is attained on (a countable set of) rank-one Hankel TN matrices, whence it equals the supremum.

We now prove both of these assertions. By Proposition 26.3, each ratio \( \det(u^{cn_j})/\det(u^{cn_n}) \) is coordinatewise non-decreasing, hence its supremum on \( (0, \sqrt{\rho})^N \) is bounded above by (and in fact equals) its limit as \( u \to \sqrt{\rho}(1^-, \ldots, 1^-) \). Thus, one may compute this limit inside \((0, \sqrt{\rho})^N\) along any sequence of vectors, and we will work with the rank-one Hankel TN family \( u(\epsilon)u(\epsilon)^T \), where

\[ u(\epsilon) := \sqrt{\rho}(\epsilon, \epsilon^2, \ldots, \epsilon^N)^T, \quad \epsilon \in (0, 1). \]

More precisely, we work with a countable sequence of \( \epsilon \to 0^+ \) in place of Lemma 23.2 above; and another countable sequence of \( \epsilon \to 1^- \) in what follows. First observe:

**Lemma 26.10** (Principal Specialization Formula for real powers). Suppose \( q > 0 \) and \( n_0 < n_1 < \cdots < n_{N-1} \) are real. If \( u := (1, q, \ldots, q^{N-1})^T \), then

\[ \det(u^{cn}) = \prod_{0 \leq j < k \leq N-1} (q^{n_k} - q^{n_j}) = V(q^{cn}). \]

The proof is exactly the same as of Proposition 24.6(2), since the transpose of \( u^{cn} \) is a usual Vandermonde matrix.

We can now complete the proof of Theorem 26.1. The above lemma immediately implies that

\[ \frac{\det(u(\epsilon)^{cn_j})}{\det(u(\epsilon)^{cn_n})} = \sqrt{\rho}^{M-n_j} \frac{V(e^{cn_j})}{V(e^{cn_n})}, \quad \forall 0 \leq j \leq N - 1. \]
Dividing the numerator and denominator by \((1 - \epsilon)^{\binom{N}{2}}\) and taking the limit as \(\epsilon \to 1^-\) using L'Hôpital's rule, we obtain the expression

\[
\sqrt{\rho^{M-n_j}} \frac{V(n_j)}{V(n)}.
\]

Since all of these suprema/limits occur as \(\epsilon \to 1^-\), we finally have:

\[
\sup_{(0, \sqrt{\rho})^N} \sum_{j=0}^{N-1} \frac{\det(u^{cn_j})^2}{c_{n_j} \det(u^{cn_j})^2} = \lim_{\epsilon \to 1^-} \sum_{j=0}^{N-1} \frac{\det(u(\epsilon)^{cn_j})^2}{c_{n_j} \det(u(\epsilon)^{cn_j})^2} = \sum_{j=0}^{N-1} \frac{V(n_j)^2 \rho^{M-n_j}}{V(n)^2 c_{n_j}}.
\]

This proves the equivalence of assertions (1)–(3) in the theorem, for rank-one matrices.

Finally, suppose all \(n_j \in \mathbb{Z}^\geq 0 \cup [N-2, \infty)\). In this case \((4) \implies (1)\) is immediate. Conversely, given that (1) holds, we prove (4) using once again the integration trick of FitzGerald–Horn, as isolated in Theorem 9.10. The proof and calculation are very similar to that of Theorem 25.2 above, and are left to the interested reader as an exercise.

26.3. Applications to Hankel TN preservers in fixed dimension and to power series preservers. We conclude by discussing some applications of Theorem 26.1. First, the result implies in particular that \(A^{cn_j}\) is bounded above by a multiple of \(\sum_{j=0}^{N-1} c_{n_j} A^{cn_j}\).

In particular, the proof of Theorem 23.7 above goes through; thus, we have classified the sign patterns of all entrywise ‘power series’ preserving positivity on \(\mathbb{F}_N((0, \rho))\).

Second, the equivalent conditions in Theorem 26.1 classifying the (entrywise) polynomial positivity preservers on \(\mathbb{F}_N((0, \rho))\) – or on rank-one matrices – also end up classifying the polynomial preservers of total non-negativity on the corresponding test sets:

**Corollary 26.11.** With notation as in Theorem 26.1, if we restrict to all real powers and only the rank-one matrices, then the assertions (1)–(3) in Theorem 26.1 are further equivalent to:

\[
(1') f[-] \text{ preserves total non-negativity on all rank-one matrices in } \mathbb{HTN}_{N} \text{ with entries in } (0, \rho).
\]

If moreover the powers \(n_j\) lie in \(\mathbb{Z}^\geq 0 \cup [N-2, \infty)\), then these conditions are further equivalent to:

\[
(4') f[-] \text{ preserves total non-negativity on all matrices in } \mathbb{HTN}_{N} \text{ with entries in } [0, \rho].
\]

Recall here that by Definition 12.10 \(\mathbb{HTN}_N\) denotes the set of \(N \times N\) Hankel totally nonnegative matrices.

**Proof.** Clearly, (1') implies assertion (2) in Theorem 26.1. Conversely, we claim that assertion (1) in Theorem 26.1 implies (1') via Theorem 4.1. Indeed, if \(A \in \mathbb{HTN}_N\) has rank one and entries in \((0, \rho)\), then \(f[A] \in \mathbb{P}_N\) by Theorem 26.1(1). Similarly, \(A^{(1)} \oplus (0)_{1 \times 1} \in \mathbb{P}_N((0, \rho))\) and has rank one, whence \(f[A^{(1)}]\) is also positive semidefinite, and hence Theorem 4.1 applies as desired.

The same proof works to show the equivalence between (4') and Theorem 26.1(1). □

Our third and final application is to bounding \(g[A]\), where \(g(x)\) is a power series – or more generally, a sum of real powers – by a threshold times \(\sum_{j=0}^{N-1} c_{n_j} A^{cn_j}\). This extends Theorem 26.1 in which \(g(x) = x^M\). The idea is that if we fix \(0 \leq n_0 < \cdots < n_{N-1}\) and \(c_{n_j}\)
for \( j = 0, \ldots, N - 1 \), then
\[
A^\circ M \leq t_M \sum_{j=0}^{N-1} c_{n_j}A^{\circ n_j}, \quad \text{where } t_M := \sum_{j=0}^{N-1} \frac{V(n_j)^2}{c_{n_j}V(n)}\rho^{M-n_j}, \tag{26.12}
\]
and this linear matrix inequality holds for all \( A \in \mathbb{P}_N((0, \rho)) \) – possibly of rank one if the \( n_j \) are allowed to be arbitrary non-negative real numbers, else of all ranks if all \( n_j \in \mathbb{Z}^\geq 0 \cup [N - 2, \infty) \). Here the \( \leq \) stands for the positive semidefinite ordering, or Loewner ordering – see Definition 14.7, for instance. Moreover, the constant \( t_M \) depends on \( M \) through \( n_j \) and \( \rho^{M-n_j} \).

If now we consider a power series \( g(x) := \sum_{M \geq n_{N-1}+1} c_M x^M \), then by adding several linear matrix inequalities of the form \( (26.12) \), it follows that
\[
g[A] \leq t_g \sum_{j=0}^{N-1} c_{n_j}A^{\circ n_j}, \quad \text{where } t_g := \sum_{M \geq n_{N-1}+1} \max(c_M, 0)t_M,
\]
and this is a valid linear matrix inequality as long as the sum \( t_g \) is convergent. Thus, we now explore when this sum converges.

Even more generally: notice that a power series is the sum/integral of the power function, over a measure on the powers which is supported on the integers. Thus, given any real measure \( \mu \) supported in \([n_{N-1} + \varepsilon, \infty) \), one can consider its corresponding ‘Laplace transform’
\[
g_\mu(x) := \int_{n_{N-1}+\varepsilon}^\infty x^t \, d\mu(t). \tag{26.13}
\]
Our final application of Theorem 26.1 explores in this generality, when a threshold exists to bound \( g_\mu[A] \) by a sum of \( N \) lower powers.

**Theorem 26.14.** Fix \( N \geq 2 \) and real exponents \( 0 \leq n_0 < \cdots < n_{N-1} \) in the set \( \mathbb{Z}^\geq 0 \cup [N - 2, \infty) \). Also fix scalars \( \rho, c_{n_j} > 0 \) for all \( j \).

Now suppose \( \varepsilon, \varepsilon' > 0 \) and \( \mu \) is a real measure supported on \([n_{N-1} + \varepsilon, \infty) \) such that \( g_\mu(x) \) – defined as in \( (26.13) \) – is absolutely convergent at \( \rho(1 + \varepsilon') \). Then there exists a finite constant \( t_\mu \in (0, \infty) \), such that the map
\[
t_\mu \sum_{j=0}^{N-1} c_{n_j}x^{n_j} - g_\mu(x)
\]
entrywise preserves positivity on \( \mathbb{P}_N((0, \rho)) \). Equivalently, \( g_\mu[A] \leq t_\mu \sum_{j=0}^{N-1} c_{n_j}A^{\circ n_j} \), for all \( A \in \mathbb{P}_N((0, \rho)) \).

**Proof.** If \( \mu = \mu_+ - \mu_- \) denotes the decomposition of \( \mu \) into its positive and negative parts, then notice (e.g. by the FitzGerald–Horn Theorem 9.2) that
\[
\int_{\mathbb{R}} A^\circ M \, d\mu_-(M) \in \mathbb{P}_N, \quad \forall A \in \mathbb{P}_N((0, \rho)).
\]
Hence, it suffices to show that
\[
t_\mu := \int_{n_{N-1}+\varepsilon}^{\infty} t_M \, d\mu_+(M) = \int_{n_{N-1}+\varepsilon}^{\infty} \sum_{j=0}^{N-1} \frac{V(n_j)^2}{c_{n_j}V(n)}\rho^{M-n_j} \, d\mu_+(M) < \infty, \tag{26.15}
\]
since this would imply:
\[ t\mu \sum_{j=0}^{N-1} c_{nj} A^{\circ n_j} - g_\mu[A] = \int_{n_{N-1}+\varepsilon}^\infty \left( t_M \sum_{j=0}^{N-1} c_{nj} A^{\circ n_j} - A^{\circ M} \right) d\mu_+(M) + \int_{n_{N-1}+\varepsilon}^\infty A^{\circ M} d\mu_-(M), \]
and both integrands and integrals are positive semidefinite.

In turn, isolating the terms in (26.15) that depend on $M$, it suffices to show for each $j$ that
\[ \int_{n_{N-1}+\varepsilon}^\infty \prod_{k=0, k\neq j}^{N-1} (M - n_k)^2 \rho^M \, d\mu_+(M) < \infty. \]

By linearity, it suffices to examine the finiteness of the integrals
\[ \int_{n_{N-1}+\varepsilon}^\infty M^k \rho^M \, d\mu_+(M), \quad k \geq 0. \]

But by assumption, $\int_{n_{N-1}+\varepsilon}^\infty \rho^M (1 + \varepsilon')^M \, d\mu_+(M)$ is finite; and moreover, for any fixed $k \geq 0$ there is a threshold $M_k$ beyond which $(1 + \varepsilon')^M \geq M^k$. (Indeed, this happens when $\frac{\log M}{M} \leq \frac{\log(1 + \varepsilon')}{k}$. Therefore,
\[ \int_{n_{N-1}+\varepsilon'}^\infty M^k \rho^M \, d\mu_+(M) \leq \int_{n_{N-1}+\varepsilon'}^{M_k} M^k \rho^M \, d\mu_+(M) + \int_{M_k}^\infty \rho^M (1 + \varepsilon')^M \, d\mu_+(M) < \infty, \]
which concludes the proof. \qed
27. Appendix B: Cauchy’s and Littlewood’s definitions of Schur polynomials.
The (dual) Jacobi–Trudi identity and Lindström–Gessel–Viennot bijections. 147

27. Appendix B: Cauchy’s and Littlewood’s definitions of Schur polynomials.

For completeness, in this section we show the equivalence of four definitions of Schur polynomials, two of which are named identities. To proceed, first recall two other families of symmetric polynomials: the elementary symmetric polynomials are simply
\[ e_1(u_1, u_2, \ldots) := u_1 + u_2 + \cdots, \quad e_2(u_1, u_2, \ldots) := u_1u_2 + u_1u_3 + u_2u_3 + \cdots, \]
and in general,
\[ e_k(u_1, u_2, \ldots) := \sum_{1 \leq j_1 < j_2 < \cdots} u_{j_1}u_{j_2}\cdots \]
These symmetric functions crucially feature while decomposing polynomials into linear factors.

We also recall the complete homogeneous symmetric polynomials
\[ h_k(u_1, u_2, \ldots) := \sum_{1 \leq j_1 \leq j_2 \leq \cdots} u_{j_1}u_{j_2}\cdots \]
By convention, we set \( e_0 = h_0 = 1 \), and \( e_k = h_k = 0 \) for \( k < 0 \). Now we have:

**Theorem 27.1.** Fix an integer \( N \geq 1 \) and any unital commutative ground ring. Given a partition of \( N \) i.e., an \( N \)-tuple of non-increasing non-negative integers \( \lambda = (\lambda_1 \geq \cdots \geq \lambda_N) \) with \( \sum_j \lambda_j = N \), the following four definitions give the same expression \( s_{\lambda+\delta}(u_1, u_2, \ldots, u_N) \).

1. (Littlewood’s definition.) The sum of weights over all column-strict Young tableaux of shape \( \lambda \) with cell entries \( u_1, \ldots, u_N \).
2. (Cauchy’s definition, aka the type ~A Weyl Character Formula.) The ratio of the (generalized) Vandermonde determinants \( a_{\lambda+\delta}/a_\delta \), where \( \delta := (N-1, N-2, \ldots, 0) \), and \( a_\lambda := \det(u_{\lambda_j+N-k}^{j}) \).
3. (The Jacobi–Trudi identity.) The determinant \( \det(h_{\lambda_j-j+k}^{N})_{j,k=1}^{N} \).
4. (The dual Jacobi–Trudi, or von Nägelsbach–Kostka identity.) The determinant \( \det(e_{\lambda_j-j+k}^{N}) \), where \( \lambda' \) is the dual partition, meaning \( \lambda'_k := \#\{j : \lambda_j \geq k\} \).

From this result, we deduce the equivalence of these definitions of the Schur polynomial for fewer numbers of variables \( u_1, \ldots, u_n \), where \( n \leq N \).

**Corollary 27.2.** Suppose \( 1 \leq r < N \) and \( \lambda_{r+1} = \cdots = \lambda_N = 0 \). Then the four definitions in Theorem 27.1 agree for the smaller set of variables \( u_1, \ldots, u_r \).

**Proof.** Using fewer numbers of variables in the definitions (3), (4) amounts to specializing the remaining variables \( u_{r+1}, \ldots, u_N \) to zero. The same holds for definition (1) since weights involving the ‘extra’ variables \( u_{r+1}, \ldots, u_N \) now get set to zero. It follows that definitions (1), (3), and (4) agree for fewer numbers of variables.

We will show that Cauchy’s definition (2) in Theorem 27.1 has the same property. In this case the definitions are different: Given \( u_1, \ldots, u_r \) for \( 1 \leq r \leq N \), the corresponding ratio of alternating polynomials would only involve \( \lambda_1 \geq \cdots \geq \lambda_r \), and would equal \( \det(u_{\lambda_j+N-r}^{j})_{j,k=1}^{r} / \det(u_{\lambda_j-r}^{j})_{j,k=1}^{r} \). We now claim that this equals the ratio in (2), by downward induction on \( r \leq N \). Note that it suffices to show the claim for \( r = N-1 \). But here, if we set \( u_N := 0 \) then both generalized Vandermonde matrices have last column \((0, \ldots, 0, 1)^T\). In particular, we may expand along their last columns. Now cancelling the common factors
of $u_1 \cdots u_{N-1}$ from each of the previous columns reduces to the case of $r = N - 1$, and the proof is completed by similarly continuing inductively. □

The remainder of this section is devoted to proving Theorem [27.1] We will show that (4) $\iff$ (1) $\iff$ (3) $\iff$ (2), and over the ground ring $\mathbb{Z}$, which then carries over to arbitrary ground rings. To do so, we use an idea due to Karlin–Macgregor (1959), Lindström (1973), and Gessel–Viennot (1985), which interprets determinants in terms of tuples of weighted lattice paths. The approach below is due to Bressoud–Wei (1993).

**Proposition 27.3.** The definitions (1) and (3) are equivalent.

*Proof.* The proof is divided into steps, for ease of exposition.

**Step 1:** In this step we define the formalism of lattice paths and their weights. Define points in the plane

$$P_k := (N - k + 1, N), \quad Q_k := (N - k + 1 + \lambda_k, 1), \quad k = 1, 2, \ldots, N,$$

and consider (ordered) $N$-tuples $p$ of (directed) lattice paths satisfying the following properties:

1. The $k$th path starts at some $P_j$ and ends at $Q_k$, for each $k$.
2. No two paths start at the same point $P_j$.
3. From $P_j$, and at each point $(a, b)$, a path can go either East or South. Weight each East step at height $(a, b)$ by $u_{N+1-b}$.

Notice that one can assign a unique permutation $\sigma = \sigma_p \in S_N$ to each tuple of paths $p$, so that paths go from $P_{\sigma(k)}$ to $Q_k$ for each $k$.

We now assign a weight to each tuple $p$, defined to be $(-1)^{\sigma_p}$ times the product of the weights at all East steps in $p$. For instance, if $\lambda = (3, 1, 1, 0, 0)$ partitions $N = 5$, then here is a typical tuple of paths:

- For $k = 4, 5$, $P_k, Q_k$ are each connected by vertical straight lines (i.e., four South steps each).
- $P_2$ and $Q_3$ are connected by a vertical straight line (i.e., four South steps).
- The steps from $P_3$ to $Q_2$ are $SESESS$.
- The steps from $P_1$ to $Q_1$ are $SEESSES$.

This tuple $p$ corresponds to the permutation $\sigma_p = (13245)$, and has weight $-u_2^3u_3u_4$.

**Step 2:** Our goal is to examine the generating function of the tuples, i.e., $\sum_\mathcal{P} \operatorname{wt}(p)$. Note that given $\sigma$, among all tuples $p$ with $\sigma_p = \sigma$, the $k$th path contributes a monomial of total degree $\lambda_k - k + \sigma(k)$, which can be any monomial in $u_1, \ldots, u_N$ of this total degree. It follows that the generating function equals

$$\sum_\mathcal{P} \operatorname{wt}(p) = \sum_{\sigma \in S_N} (-1)^{\sigma} \prod_{k=1}^N h_{\lambda_k - k + \sigma(k)} = \det(h_{\lambda_k - k + j})_{j,k=1}^N.$$

**Step 3:** We next rewrite the above generating function to obtain $\sum_T \operatorname{wt}(T)$ (the sum of weights over all column-strict Young tableaux of shape $\lambda$ with cell entries $u_1, \ldots, u_N$), which is precisely $s_{\lambda+\delta}(u_1, \ldots, u_N)$ by definition. To do so, we will pair off the tuples $p$ of intersecting paths into pairs, whose weights cancel one another.

Suppose $p$ consists of intersecting paths. Define the final intersection point of $p$ to be the lattice point with maximum $x$-coordinate where at least two paths intersect, and if there are more than one such points, then the one with minimal $y$ coordinate. We claim that exactly
two paths in $p$ intersect at this point. Indeed, if three paths intersect at any point, then all of them have to go either East or South at the next step. By the pigeonhole principle, there are at least two paths that proceed in the same direction. It follows that a point common to three paths in $p$ cannot be the final intersection point, as desired.

Define the tail of $p$ to be the two paths to the East and South of the final intersection point in $p$. Given an intersecting tuple of paths $p$, there exists a unique other tuple $p'$ with the same final intersection point between the same two paths, but with the tails swapped. It is easy to see that the paths $p, p'$ satisfy have opposite signs (for their permutations $\sigma_p, \sigma'_p$), but the same monomials in their weights. Therefore $\text{wt}(p) = -\text{wt}(p')$, and the intersecting paths pair off as desired.

**Step 4:** From Step 3, the generating function $\sum_p \text{wt}(p)$ equals the sum over only tuples of non-intersecting paths. Each of these tuples necessarily has $\sigma_p = \text{id}$, so all signs are positive. In such a tuple, the monomial weight for the $k$th path naturally corresponds to a weakly increasing sequence of $\lambda_k$ integers in $[1, N]$. That the paths do not intersect corresponds to the entries in the $k$th sequence being strictly smaller than the corresponding entries in the $(k+1)$st sequence. This yields a natural weight-preserving bijection from the tuples of non-intersecting paths to the ‘column-strict’ Young tableaux of shape $\lambda$ with cell entries $1, \ldots, N$. (Notice that these tableaux are in direct bijection to the column-strict Young tableaux studied earlier in this chapter, by switching the cell entries $j \leftrightarrow N+1-j$.) This concludes the proof. 

**Proposition 27.4.** The definitions (1) and (4) are equivalent.

**Proof.** The proof is a variant of that of Proposition 27.3. Now we consider all tuples of paths such that the $k$th path goes from $P_{\sigma(k)}$, to the point

$$Q'_k := (N - k + 1 + \lambda'_k, 1),$$

and moreover, each of these paths has at most one East step at each fixed height – i.e., no two East steps are consecutive.

Once again, in summing to obtain the generating function, given a permutation $\sigma = \sigma_p$, the $k$th path in $p$ contributes a monomial of total degree $\lambda'_k - k + \sigma(k)$, but now runs over all monomials with individual variables of degree at most $1$ – i.e., all monomials in $e_{\lambda'_k - k + \sigma(k)}$. It follows that

$$\sum_p \text{wt}(p) = \sum_{\sigma \in S_N} (-1)^\sigma \prod_{k=1}^N e_{\lambda'_k - k + \sigma(k)} = \text{det}(e_{\lambda'_k - k + j})_{j,k=1}^N.$$

On the other side, we once again pair off tuples – this time, leaving the ones that do not overlap. In other words, paths in tuples may intersect at a point, but do not share an East/South line segment. As in the previous proof, notice that at most two paths can overlap at a given point. Now given a tuple containing two overlapping paths, define the final overlap point similarly as in Proposition 27.3. Then for every tuple of paths $p$ that overlaps, there exists a unique other tuple $p'$ with the same final overlap point between the same two paths, but with the (new version of) tails swapped. It is easy to see that $p, p'$ have the same monomials as weights, but with opposite signs, so they pair off and cancel weights.

This leaves us with tuples of non-overlapping paths, all of which again corresponding to $\sigma_p = \text{id}$. In such a tuple, from the $k$th path we obtain a strictly increasing sequence of $\lambda'_k$ integers in $[1, N]$. That the paths do not overlap corresponds to the entries in the $k$th sequence being at most as large as the corresponding entries in the $(k+1)$st sequence. This
gives a bijection to the conjugates of column-strict Young tableaux of shape $\lambda$, and hence we once again have $\sum_p \text{wt}(p) = \sum_T \text{wt}(T)$ in this setting. \hfill \Box

**Corollary 27.5.** Schur polynomials are symmetric and homogeneous.

*Proof.* This follows because Definition (4) is symmetric and homogeneous in the variables $u_j$. \hfill \Box

Finally, we show:

**Proposition 27.6.** The definitions (2) and (3) are equivalent.

*Proof.* Once again, this proof is split into steps, for ease of exposition. In the proof below, we use the above results and assume that the definitions (1), (3), and (4) are all equivalent. Thus, our goal is to show that

$$\det(u_j^{\lambda_k-k})_{j,k=1}^{N} \cdot \det(h_{\lambda_j-j+k})_{j,k=1}^{N} = \det(u_j^{\lambda_k-N-k})_{j,k=1}^{N}.$$  

**Step 1:** We explain the formalism, which is a refinement of the one in the proof of Proposition 27.3. Thus, we return to the setting of paths between $P_k = (N - k + 1, N)$ and $Q_k = (N - k + 1 + \lambda_k, 1)$ for $k = 1, \ldots, N$, but now equipped also with a permutation $\tau \in S_N$. The weight of an East step now depends on its height: at height $N + 1 - b$, an East step has weight $u_{\tau(b)}$ instead of $u_b$. We will consider tuples of paths over all $\tau$; let us write their weights as $\text{wt}_\tau(p)$ for notational clarity. In what follows, we also use $p$ or $(p, \tau)$ depending on the need to specify and work with $\tau \in S_N$.

For each fixed $\tau \in S_N$, notice first that the generating function $\sum_p \text{wt}_\tau(p)$ of the $\tau$-permuted paths is independent of $\tau$, by Corollary 27.5.

Now we define a new weight for these $\tau$-permuted paths $p$. Namely, given $p = (p, \tau)$, recall there exists a unique permutation $\sigma_p \in S_N$; now define

$$\text{wt}_\tau'(p) := (-1)^\tau \mu(\tau) \cdot \text{wt}_\tau(p),$$

where $\mu(\tau) := u_{\tau(1)}^{N-1} u_{\tau(2)}^{N-2} \cdots u_{\tau(N-1)}$.

The new generating function is

$$\sum_{\tau \in S_N} \sum_p \text{wt}_\tau'(p) = \sum_{\tau \in S_N} (-1)^\tau \mu(\tau) \sum_p \text{wt}_\tau(p) = \det(h_{\lambda_k-k+j})_{j,k=1}^{N} \cdot \det(u_j^{\lambda_k-N-k})_{j,k=1}^{N},$$

where the final equality follows from the above propositions, given that the inner sum is independent of $\tau$ from above.

**Step 2:** Say that a tuple $p = (P_{\sigma_p(k)} \rightarrow Q_k)_k$ is high enough if for every $1 \leq k \leq N$, the $k$th path has no East steps below height $N + 1 - k$. We **claim** that (summing over all $\tau \in S_N$) the $\tau$-tuples that are not high enough once again pair up, with cancelling weights.

Modulo the claim, we prove the theorem. The first reduction is that for a fixed $\tau$, we may further restrict to the $\tau$-tuples that are high enough and are non-intersecting (as in the proof of Proposition 27.3). Indeed, defining the final intersection point and the tail of $p$ as in that proof, it follows that switching tails in tuples $p$ of intersecting paths changes neither the monomial part of the weight, nor the high-enough property; and it induces the opposite sign to that of $p$.

Thus, the generating function of all $\tau$-tuples (over all $\tau$) equals that of all non-intersecting, high-enough $\tau$-tuples (also summed over all $\tau \in S_N$). But each such tuple corresponds to $\sigma_p = \text{id}$, and in it, all East steps in the first path must occur in the topmost row/height/Y-coordinate of $N$. Hence all East steps in the second path must occur in the next highest row, and so on. It follows that the non-intersecting, high-enough $\tau$-tuples $p = (p, \tau)$ are in
27. Appendix B: Cauchy’s and Littlewood’s definitions of Schur polynomials.

The (dual) Jacobi–Trudi identity and Lindström–Gessel–Viennot bijections. 151

bijection with \( \tau \in S_N \); moreover, each such tuple has weight \((-1)^{\tau} \mu(\tau) u^{\lambda_1}_{\tau(1)} u^{\lambda_2}_{\tau(2)} \cdots u^{\lambda_N}_{\tau(N)}\). Thus, the above generating function is shown to equal

\[
\det(u^{\lambda_k+N-k}_{\tau,j})_{j,k=1}^{N},
\]

and the proof is complete.

**Step 3:** It thus remains to show the claim in Step 2 above. Given parameters

\[ \sigma \in S_N, \quad k \in [1, N], \quad j \in [1, N-k], \]

let \( NH_{\sigma,k,j} \) denote the \( \tau \)-tuples of paths \( p = (p, \tau) \) (with \( \tau \) running over \( S_N \)), which satisfy the following properties:

1. \( p \) is not high (\( NH \)) enough.
2. In \( p \), the \( k \)th path has an East step at most by height \( N-k \), but the paths labelled \( 1, \ldots, k-1 \) are all high enough.
3. Moreover, \( j \) is the height of the lowest East step in the \( k \)th path; thus \( j \in [1, N-k] \).
4. The permutation associated to the start and end points of the paths in the tuple is \( \sigma_p = \sigma \in S_N \).

Note that the set \( NH \) of tuples of paths that are not high enough can be partitioned as:

\[ NH = \bigcup_{\sigma \in S_N, \quad k \in [1, N], \quad j \in [1, N-k]} NH_{\sigma,k,j}. \]

We now construct an involution of sets \( \iota : NH \to NH \) which permutes each subset \( NH_{\sigma,k,j} \), and such that \( p \) and \( \iota(p) \) have the same monomial attached to them but different \( \tau, \tau' \), leading to cancelling signs \((-1)^{\tau} \neq (-1)^{\tau'}\).

Thus, suppose \( p \) is a \( \tau \)-tuple in \( NH_{\sigma,k,j} \). Now define \( \tau' := \tau \circ (N-j, N+1-j) \); in other words,

\[
\tau'(i) := \begin{cases} 
\tau(i+1), & \text{if } i = N-j; \\
\tau(i-1), & \text{if } i = N-j+1; \\
\tau(i), & \text{otherwise}.
\end{cases}
\]

In particular,

\[ (-1)^{\tau'} = -(-1)^{\tau} \quad \text{and} \quad \mu(\tau') = \mu(\tau) u_{\tau(N+1-j)} u_{\tau(N-j)}^{-1}. \]

With \( \tau' \) in hand, we can define the tuple \( \iota(p) = (\iota(p), \tau') \in NH_{\sigma,k,j} \). First, change the weight of each East step at height \( N+1-b \), from \( u_{\tau(b)} \) to \( u_{\tau'(b)} \). Next, we keep unchanged the paths labelled \( 1, \ldots, k-1 \), and in the remaining paths we do not change the source and target nodes either (since \( \sigma \) is fixed). Notice that weights change at only two heights \( j, j+1 \); hence the first \( k-1 \) paths do not see any weights change.

The changes in the (other) paths are now described. In the \( k \)th path, change only the numbers \( n_l \) of East steps at height \( l = j, j+1 \), via: \((n_j, n_{j+1}) \mapsto (n_{j+1}+1, n_j-1)\). Note, the product of weights of all East steps in this path changes by a multiplicative factor of \( u_{\tau(N+1-j)}^{-1} u_{\tau(N-j)} \) – which cancels the above change from \( \mu(\tau) \) to \( \mu(\tau') \). Finally, in the \( m \)th path for each \( m > k \), if \( n_l \) again denotes the number of East steps at height \( l \), then we swap \( n_j \leftrightarrow n_{j+1} \) steps in the \( m \)th path. This leaves unchanged the weight of those paths, and hence of the tuple \( p \) overall.

It is now straightforward to verify that the map \( \iota \) is an involution that preserves each of the sets \( NH_{\sigma,k,j} \). Since \( \text{wt}(\iota(p)) = -\text{wt}(p) \) for all \( p \in NH \), the claim in Step 2 is true, and the proof of the theorem is complete. \( \square \)