A Panorama of Positivity. I: Dimension Free

Alexander Belton, Dominique Guillot, Apoorva Khare, and Mihai Putinar

Serguei Shimorin, in memoriam

Abstract This survey contains a selection of topics unified by the concept of positive semidefiniteness (of matrices or kernels), reflecting natural constraints imposed on discrete data (graphs or networks) or continuous objects (probability or mass distributions). We put emphasis on entrywise operations which preserve positivity, in a variety of guises. Techniques from harmonic analysis, function theory, operator theory, statistics, combinatorics, and group representations are invoked. Some partially forgotten classical roots in metric geometry and distance transforms are

D.G. is partially supported by a University of Delaware Research Foundation grant, by a Simons Foundation collaboration grant for mathematicians, and by a University of Delaware Research Foundation Strategic Initiative grant. A.K. is partially supported by Ramanujan Fellowship SB/S2/RJN-121/2017 and MATRICS grant MTR/2017/000295 from SERB (Govt. of India), by grant F.510/25/CAS-II/2018(SAP-I) from UGC (Govt. of India), and by a Young Investigator Award from the Infosys Foundation.

A. Belton
Department of Mathematics and Statistics, Lancaster University, Lancaster, UK
e-mail: a.belton@lancaster.ac.uk

D. Guillot
University of Delaware, Newark, DE, USA
e-mail: dguillot@udel.edu

A. Khare (✉)
Indian Institute of Science, Analysis and Probability Research Group, Bengaluru, India
e-mail: khare@iisc.ac.in

M. Putinar
University of California at Santa Barbara, Santa Barbara, CA, USA

Newcastle University, Newcastle upon Tyne, UK
e-mail: mputinar@math.ucsb.edu; mihai.putinar@ncl.ac.uk

A. Aleman et al. (eds.), *Analysis of Operators on Function Spaces*,
Trends in Mathematics, https://doi.org/10.1007/978-3-030-14640-5_5

presented with comments and full bibliographical references. Modern applications to high-dimensional covariance estimation and regularization are included.

Keywords Metric geometry · Positive semidefinite matrix · Toeplitz matrix · Hankel matrix · Positive definite function · Completely monotone functions · Absolutely monotonic functions · Entrywise calculus · Generalized Vandermonde matrix · Schur polynomials · Symmetric function identities · Totally positive matrices · Totally non-negative matrices · Totally positive completion problem · Sample covariance · Covariance estimation · Hard/soft thresholding · Sparsity pattern · Critical exponent of a graph · Chordal graph · Loewner monotonicity · Convexity · Super-additivity

2010 Mathematics Subject Classification 15-02, 26-02, 15B48, 51F99, 15B05, 05E05, 44A60, 15A24, 15A15, 15A45, 15A83, 47B35, 05C50, 30E05, 62J10

This is the first part of a two-part survey; we include on p. 165 the table of contents for the second part [10]. The survey in its unified form may be found online; see [9]. The abstract, keywords, MSC codes, and introduction are the same for both parts.

1 Introduction

Matrix positivity, or positive semidefiniteness, is one of the most wide-reaching concepts in mathematics, old and new. Positivity of a matrix is as natural as positivity of mass in statics or positivity of a probability distribution. It is a notion which has attracted the attention of many great minds. Yet, after at least two centuries of research, positive matrices still hide enigmas and raise challenges for the working mathematician.

The vitality of matrix positivity comes from its breadth, having many theoretical facets and also deep links to mathematical modelling. It is not our aim here to pay homage to matrix positivity in the large. Rather, the present survey, split for technical reasons into two parts, has a limited but carefully chosen scope.

Our panorama focuses on entrywise transforms of matrices which preserve their positive character. In itself, this is a rather bold departure from the dogma that canonical transformations of matrices are not those that operate entry by entry. Still, this apparently esoteric topic reveals a fascinating history, abundant characteristic phenomena, and numerous open problems. Each class of positive matrices or kernels (regarding the latter as continuous matrices) carries a specific toolbox of internal transforms. Positive Hankel forms or Toeplitz kernels, totally positive matrices, and group-invariant positive definite functions all possess specific *positivity preservers*. As we see below, these have been thoroughly studied for at least a century.

One conclusion of our survey is that the classification of positivity preservers is accessible in the dimension-free setting, that is, when the sizes of matrices are unconstrained. In stark contrast, precise descriptions of positivity preservers in fixed dimension are elusive, if not unattainable with the techniques of modern mathematics. Furthermore, the world of applications cares much more about matrices of fixed size than in the free case. The accessibility of the latter was by no means a sequence of isolated, simple observations. Rather, it grew organically out of distance geometry, and spread rapidly through harmonic analysis on groups, special functions, and probability theory. The more recent and highly challenging path through fixed dimensions requires novel methods of algebraic combinatorics and symmetric functions, group representations, and function theory.

As well as its beautiful theoretical aspects, our interest in these topics is also motivated by the statistics of big data. In this setting, functions are often applied entrywise to covariance matrices, in order to induce sparsity and improve the quality of statistical estimators (see [45, 46, 75]). Entrywise techniques have recently increased in popularity in this area, largely because of their low computational complexity, which makes them ideal to handle the ultra high-dimensional datasets arising in modern applications. In this context, the dimensions of the matrices are fixed, and correspond to the number of underlying random variables. Ensuring that positivity is preserved by these entrywise methods is critical, as covariance matrices must be positive semidefinite. Thus, there is a clear need to produce characterizations of entrywise preservers, so that these techniques are widely applicable and mathematically justified. We elaborate further on this in the second part of the survey [10].

We conclude by remarking that, while we have tried to be comprehensive in our coverage of the field of matrix positivity and the entrywise calculus, our panorama is far from being complete. We apologize for any omissions.

2 From Metric Geometry to Matrix Positivity

2.1 Distance Geometry

During the first decade of the twentieth century, the concept of a metric space emerged from the works of Fréchet and Hausdorff, each having different and well-anchored roots, in function spaces and in set theory and measure theory. We cannot think today of modern mathematics and physics without referring to metric spaces, which touch areas as diverse as economics, statistics, and computer science. Distance geometry is one of the early and ever-lasting by-products of metric-space theory. One of the key figures of the Vienna Circle, Karl Menger, started a systematic study in the 1920s of the geometric and topological features of spaces that are intrinsic solely to the distance they carry. Menger published his findings in a series of articles having the generic name “Untersuchungen über allgemeine Metrik,” the first one being [63]; see also his synthesis [64]. His work was very influential in the decades to come [16], and by a surprising and fortunate stroke not often encountered in mathematics, Menger's distance geometry has been resurrected in recent times by practitioners of convex optimization and network analysis [26, 60].

Let (X, ρ) be a metric space. One of the naive, yet unavoidable, questions arising from the very beginning concerns the nature of operations $\phi(\rho)$ which may be performed on the metric and which enhance various properties of the topological space X. We all know that $\rho/(\rho+1)$ and ρ^γ, if $\gamma \in (0, 1)$, also satisfy the axioms of a metric, with the former making it bounded. Less well known is an observation due to Blumenthal, that the new metric space (X, ρ^γ) has the four-point property if $\gamma \in (0, 1/2]$: every four-point subset of X can be embedded isometrically into Euclidean space [16, Section 49].

Metric spaces which can be embedded isometrically into Euclidean space, or into infinite-dimensional Hilbert space, are, of course, distinguished and desirable for many reasons. We owe to Menger a definitive characterization of this class of metric spaces. The core of Menger's theorem, stated in terms of certain matrices built from the distance function (known as Cayley–Menger matrices) was slightly reformulated by Fréchet and cast in the following simple form by Schoenberg.

Theorem 2.1 (Schoenberg [81]) *Let $d \geq 1$ be an integer and let (X, ρ) be a metric space. An $(n+1)$-tuple of points $x_0, x_1, \ldots, x_n$ in X can be isometrically embedded into Euclidean space $\mathbb{R}^d$, but not into $\mathbb{R}^{d-1}$, if and only if the matrix*

$$[\rho(x_0, x_j)^2 + \rho(x_0, x_k)^2 - \rho(x_j, x_k)^2]_{j,k=1}^n,$$

is positive semidefinite with rank equal to d.

Proof This is surprisingly simple. Necessity is immediate, since the Euclidean norm and scalar product in $\mathbb{R}^d$ give that

$$\begin{aligned}
&\rho(x_0, x_j)^2 + \rho(x_0, x_k)^2 - \rho(x_j, x_k)^2 \\
&= \|x_0 - x_j\|^2 + \|x_0 - x_k\|^2 - \|(x_0 - x_j) - (x_0 - x_k)\|^2 \\
&= 2\langle x_0 - x_j, x_0 - x_k\rangle,
\end{aligned}$$

and the latter are the entries of a positive semidefinite Gram matrix of rank less than or equal to d.

For the other implication, we consider first a full-rank $d \times d$ matrix associated with a $(d+1)$-tuple. The corresponding quadratic form

$$Q(\lambda) = \frac{1}{2} \sum_{j,k=1}^{d} (\rho(x_0, x_j)^2 + \rho(x_0, x_k)^2 - \rho(x_j, x_k)^2)\lambda_j \lambda_k$$

is positive definite. Hence there exists a linear change of variables

$$\lambda_k = \sum_{j=1}^{d} a_{jk}\mu_j \qquad (1 \leq j \leq d)$$

such that

$$Q(\lambda) = \mu_1^2 + \mu_2^2 + \cdots + \mu_d^2.$$

Interpreting $(\mu_1, \mu_2, \ldots, \mu_d)$ as coordinates in $\mathbb{R}^d$, the standard simplex with vertices

$$e_0 = (0, \ldots, 0), \quad e_1 = (1, 0, \ldots, 0), \quad \ldots, \quad e_d = (0, \ldots, 0, 1)$$

has the corresponding quadratic form (of distances) equal to $\mu_1^2 + \mu_2^2 + \cdots + \mu_d^2$. Now we perform the coordinate change $\mu_j \mapsto \lambda_j$. Specifically, set $P_0 = 0$ and let $P_j \in \mathbb{R}^d$ be the point with coordinates $\lambda_j = 1$ and $\lambda_k = 0$ if $k \neq j$. Then one identifies distances:

$$\begin{aligned} \|P_0 - P_j\| &= \rho(x_0, x_j) \qquad (0 \leq j \leq d) \\ \text{and} \quad \|P_j - P_k\| &= \rho(x_j, x_k) \qquad (1 \leq j, k \leq d). \end{aligned}$$

The remaining case with $n > d$ can be analyzed in a similar way, after taking an appropriate projection. □

In the conditions of the theorem, fixing a "frame" of d points and letting the $(d + 1)$-th point float, one obtains an embedding of the full metric space (X, ρ) into $\mathbb{R}^d$. This idea goes back to Menger, and it led, with Schoenberg's touch, to the following definitive statement. Here and below, all Hilbert spaces are assumed to be separable.

Corollary 2.2 (Schoenberg [81], following Menger) *A separable metric space (X, ρ) can be isometrically embedded into Hilbert space if and only if, for every $(n + 1)$-tuple of points $(x_0, x_1, \ldots, x_n)$ in X, where $n \geq 2$, the matrix*

$$[\rho(x_0, x_j)^2 + \rho(x_0, x_k)^2 - \rho(x_j, x_k)^2]_{j,k=1}^n$$

is positive semidefinite.

The notable aspect of the two previous results is the interplay between purely geometric concepts and matrix positivity. This will be a recurrent theme of our survey.

2.2 Spherical Distance Geometry

One can specialize the embedding question discussed in the previous section to submanifolds of Euclidean space. A natural choice is the sphere.

For two points x and y on the unit sphere $S^{d-1} \subset \mathbb{R}^d$, the rotationally invariant distance between them is

$$\rho(x, y) = \sphericalangle(x, y) = \arccos\langle x, y\rangle,$$

where the angle between the two vectors is measured on a great circle and is always less than or equal to π.

A straightforward application of the simple, but central, Theorem 2.1 yields the following result.

Theorem 2.3 (Schoenberg [81]) *Let (X, ρ) be a metric space and let $(x_1, \ldots, x_n)$ be an n-tuple of points in X. For any integer $d \geq 2$, there exists an isometric embedding of $(x_1, \ldots, x_n)$ into S^{d-1} endowed with the geodesic distance but not S^{d-2} if and only if*

$$\rho(x_j, x_k) \leq \pi \qquad (1 \leq j, k \leq n)$$

and the matrix $\left[\cos \rho(x_j, x_k)\right]_{j,k=1}^n$ is positive semidefinite of rank d.

Indeed, the necessity is assured by choosing x_0 to be the origin in $\mathbb{R}^d$. In this case,

$$\begin{aligned}\rho(x_0, x_j)^2 + \rho(x_0, x_k)^2 - \rho(x_j, x_k)^2 &= \|x_j\|^2 + \|x_k\|^2 - \|x_j - x_k\|^2 \\ &= 2\langle x_j, x_k\rangle \\ &= 2\cos\rho(x_j, x_k).\end{aligned}$$

The condition is also sufficient, by possibly adding an external point x_0 to the metric space, subject to the constraints that $\rho(x_0, x_j) = 1$ for all j. The details can be found in [81].[1]

2.3 Distance Transforms

A notable step forward in the study of the existence of isometric embeddings of a metric space into Euclidean or Hilbert space was made by Schoenberg. In a series of articles [82, 84, 85, 94], he changed the set-theoretic lens of Menger, by initiating a harmonic-analysis interpretation of this embedding problem. This was a major turning point, with long-lasting, unifying, and unexpected consequences.

[1]An alternate proof of sufficiency is to note that $A := [\cos\rho(x_j, x_k)]_{j,k=1}^n$ is a Gram matrix of rank r, hence equal to $B^T B$ for some $r \times n$ matrix B with unit columns. Denoting these columns by $\mathbf{b}_1, \ldots, \mathbf{b}_n \in S^{r-1}$, the map $x_j \mapsto b_j$ is an isometry since $\rho(x_j, x_k)$ and $\sphericalangle(y_j, y_k) \in [0, \pi]$. Moreover, since A has rank r, the $\mathbf{b}_j$ cannot all lie in a smaller-dimensional sphere.

We return to a separable metric space (X, ρ) and seek distance-function transforms $\rho \mapsto \phi(\rho)$ which enhance the geometry of X, to the extent that the new metric space $\big(X, \phi(\rho)\big)$ is isometrically equivalent to a subspace of Hilbert space. Schoenberg launched this whole new chapter from the observation that the Euclidean norm is such that the matrix

$$[\exp\big(-\|x_j - x_k\|^2\big)]_{j,k=1}^N$$

is positive semidefinite for any choice of points $x_1, \ldots, x_N$ in the ambient space. Once again, we see the presence of matrix positivity. While this claim may not be obvious at first sight, it is accessible once we recall a key property of Fourier transforms.

An even function $f : \mathbb{R}^d \to \mathbb{C}$ is said to be *positive definite* if the complex matrix $[f(x_j - x_k)]_{j,k=1}^N$ is positive semidefinite for any $N \geq 1$ and any choice of points $x_1, \ldots, x_N \in \mathbb{R}^d$. We will call $f(x - y)$ a *positive semidefinite kernel* on $\mathbb{R}^d \times \mathbb{R}^d$ in this case.

Bochner's theorem [18] characterizes positive definite functions on $\mathbb{R}^d$ as Fourier transforms of even positive measures of finite mass:

$$f(\xi) = \int e^{-\mathrm{i}x\cdot\xi}\, \mathrm{d}\mu(x).$$

Indeed,

$$f(\xi - \eta) = \int e^{-\mathrm{i}x\cdot\xi} e^{\mathrm{i}x\cdot\eta}\, \mathrm{d}\mu(x)$$

is a positive semidefinite kernel because it is the average over μ of the positive kernel $(\xi, \eta) \mapsto e^{-\mathrm{i}x\cdot\xi} e^{\mathrm{i}x\cdot\eta}$. Since the Gaussian e^{-x^2} is the Fourier transform of itself (modulo constants), it turns out that it is a positive definite function on $\mathbb{R}$, whence $\exp(-\|x\|^2)$ has the same property as a function on $\mathbb{R}^d$. Taking one step further, the function $x \mapsto \exp(-\|x\|^2)$ is positive definite on any Hilbert space.

With this preparation we are ready for a second characterization of metric subspaces of Hilbert space.

Theorem 2.4 (Schoenberg [84]) *A separable metric space (X, ρ) can be embedded isometrically into Hilbert space if and only if the kernel*

$$X \times X \to (0, \infty); \ (x, y) \mapsto \exp(-\lambda^2 \rho(x, y)^2)$$

is positive semidefinite for all $\lambda \in \mathbb{R}$.

Proof Necessity follows from the positive definiteness of the Gaussian discussed above. (We also provide an elementary proof below; see Lemma 4.7 and the subsequent discussion). To prove sufficiency, we recall the Menger–Schoenberg characterization of isometric subspaces of Hilbert space. We have to derive, from the positivity assumption, the positivity of the matrix

$$[\rho(x_0, x_j)^2 + \rho(x_0, x_k)^2 - \rho(x_j, x_k)^2]_{j,k=1}^n.$$

Elementary algebra transforms this constraint into the requirement that

$$\sum_{j,k=0}^n \rho(x_j, x_k)^2 c_j c_k \le 0 \qquad \text{whenever} \sum_{j=0}^n c_j = 0.$$

By expanding $\exp(-\lambda^2 \rho(x_j, x_k)^2)$ as a power series in λ^2, and invoking the positivity of the exponential kernel, we see that

$$0 \le -\lambda^2 \sum_{j,k=0}^n \rho(x_j, x_k)^2 c_j c_k + \frac{\lambda^4}{2} \sum_{j,k=0}^n \rho(x_j, x_k)^4 c_j c_k - \cdots$$

for all $\lambda > 0$. Hence the coefficient of $-\lambda^2$ is non-positive. □

The flexibility of the Fourier-transform approach is illustrated by the following application, also due to Schoenberg [84].

Corollary 2.5 *Let H be a Hilbert space with norm $\|\cdot\|$. For every $\delta \in (0, 1)$, the metric space $(H, \|\cdot\|^\delta)$ is isometric to a subspace of a Hilbert space.*

Proof Note first the identity

$$\xi^\alpha = c_\alpha \int_0^\infty (1 - e^{-s^2 \xi^2}) s^{-1-\alpha} \, \mathrm{d}s \qquad (\xi > 0, \ 0 < \alpha < 2),$$

where c_α is a normalization constant. Consequently,

$$\|x - y\|^\alpha = c_\alpha \int_0^\infty (1 - e^{-s^2 \|x-y\|^2}) s^{-1-\alpha} \, \mathrm{d}s.$$

Let $\delta = \alpha/2$. For points $x_0, x_1, \ldots, x_n$ in H and weights $c_0, c_1, \ldots, c_n$ satisfying

$$c_0 + c_1 + \cdots + c_n = 0,$$

it holds that

$$\sum_{j,k=0}^n \|x_j - x_k\|^{2\delta} c_j c_k = -c_\alpha \int_0^\infty \sum_{j,k=0}^n c_j c_k \, e^{-s^2 \|x_j - x_k\|^2} s^{-1-\alpha} \, \mathrm{d}s \le 0,$$

and the proof is complete. □

Several similar consequences of the Fourier-transform approach are within reach. For instance, Schoenberg observed in the same article that if the L^p norm is raised to the power γ, where $0 < \gamma \leq p/2$ and $1 \leq p \leq 2$, then $L^p(0, 1)$ is isometrically embeddable into Hilbert space.

2.4 Altering Euclidean Distance

By specializing the theme of the previous section to Euclidean space, Schoenberg and von Neumann discovered an arsenal of powerful tools from harmonic analysis that were able to settle the question of whether Euclidean space equipped with the altered distance $\phi\big(\|x - y\|\big)$ may be isometrically embedded into Hilbert space [83, 94]. The key ingredients are characterizations of Laplace and Fourier transforms of positive measures, that is, Bernstein's completely monotone functions [14] and Bochner's positive definite functions [18].

Here we present some highlights of the Schoenberg–von Neumann framework. First, we focus on an auxiliary class of distance transforms. A real continuous function ϕ is called *positive definite in Euclidean space* $\mathbb{R}^d$ if the kernel

$$(x, y) \mapsto \phi\big(\|x - y\|\big)$$

is positive semidefinite. Bochner's theorem and the rotation-invariance of this kernel prove that such a function ϕ is characterized by the representation

$$\phi(t) = \int_0^\infty \Omega_d(tu)\,\mathrm{d}\mu(u),$$

where μ is a positive measure and

$$\Omega_d\big(\|x\|\big) = \int_{\|\xi\|=1} e^{\mathrm{i}x\cdot\xi}\,\mathrm{d}\sigma(\xi),$$

with σ the normalized area measure on the unit sphere in $\mathbb{R}^d$; see [83, Theorem 1]. By letting d tend to infinity, one finds that positive definite functions on infinite-dimensional Hilbert space are precisely of the form

$$\phi(t) = \int_0^\infty e^{-t^2u^2}\,\mathrm{d}\mu(u),$$

with μ a positive measure on the semi-axis. Notice that positive definite functions in $\mathbb{R}^d$ are not necessarily differentiable more than $(d-1)/2$ times, while those which are positive definite in Hilbert space are smooth and even complex analytic in the sector $|\arg t| < \pi/4$.

The class of functions f which are continuous on $\mathbb{R}_+ := [0, \infty)$, smooth on the open semi-axis $(0, \infty)$, and such that

$$(-1)^n f^{(n)}(t) \geq 0 \qquad \text{for all } t > 0$$

was studied by Bernstein, who proved that they coincide with Laplace transforms of positive measures on $\mathbb{R}_+$:

$$f(t) = \int_0^\infty e^{-tu}\, \mathrm{d}\mu(u). \tag{2.1}$$

Such functions are called *completely monotonic* and have proved highly relevant for probability theory and approximation theory; see [14] for the foundational reference. Thus we have obtained a valuable equivalence.

Theorem 2.6 (Schoenberg) *A function f is completely monotone if and only if $t \mapsto f(t^2)$ is positive definite on Hilbert space.*

The direct consequences of this apparently innocent observation are quite deep. For example, the isometric-embedding question for altered Euclidean distances is completely answered via this route. The following results are from [83] and [94].

Theorem 2.7 (Schoenberg–von Neumann) *Let H be a separable Hilbert space with norm $\|\cdot\|$.*

(1) *For any integers $n \geq d > 1$, the metric space $(\mathbb{R}^d, \phi(\|\cdot\|))$ may be isometrically embedded into $(\mathbb{R}^n, \|\cdot\|)$ if and only if $\phi(t) = ct$ for some $c > 0$.*
(2) *The metric space $(\mathbb{R}^d, \phi(\|\cdot\|))$ may be isometrically embedded into H if and only if*

$$\phi(t)^2 = \int_0^\infty \frac{1 - \Omega_d(tu)}{u^2}\, \mathrm{d}\mu(u),$$

where μ is a positive measure on the semi-axis such that

$$\int_1^\infty \frac{1}{u^2}\, \mathrm{d}\mu(u) < \infty.$$

(3) *The metric space $(H, \phi(\|\cdot\|))$ may be isometrically embedded into H if and only if*

$$\phi(t)^2 = \int_0^\infty \frac{1 - e^{-t^2 u}}{u}\, \mathrm{d}\mu(u),$$

where μ is a positive measure on the semi-axis such that

$$\int_1^\infty \frac{1}{u}\, \mathrm{d}\mu(u) < \infty.$$

In von Neumann and Schoenberg's article [94], special attention is paid to the case of embedding a modified distance on the line into Hilbert space. This amounts to characterizing all *screw lines* in a Hilbert space H: the continuous functions

$$f : \mathbb{R} \to H;\ t \mapsto f_t$$

with the translation-invariance property

$$\| f_s - f_t \| = \| f_{s+r} - f_{t+r} \| \qquad \text{for all } s, r, t \in \mathbb{R}.$$

In this case, the gauge function ϕ is such that $\phi(t - s) = \| f_s - f_t \|$ and $t \mapsto f_t$ provides the isometric embedding of $(\mathbb{R}, \phi(| \cdot |))$ into H. Von Neumann seized the opportunity to use Stone's theorem on one-parameter unitary groups, together with the spectral decomposition of their unbounded self-adjoint generators, to produce a purely operator-theoretic proof of the following result.

Corollary 2.8 *The metric space* $(\mathbb{R}, \phi(| \cdot |))$ *isometrically embeds into Hilbert space if and only if*

$$\phi(t)^2 = \int_0^\infty \frac{\sin^2(tu)}{u^2}\, \mathrm{d}\mu(u) \qquad (t \in \mathbb{R}),$$

where μ *is a positive measure on* $\mathbb{R}_+$ *satisfying*

$$\int_1^\infty \frac{1}{u^2}\, \mathrm{d}\mu(u) < \infty.$$

Moreover, in the conditions of the corollary, the space $(\mathbb{R}, \phi(| \cdot |))$ embeds isometrically into $\mathbb{R}^d$ if and only if the measure μ consists of finitely many point masses, whose number is roughly $d/2$; see [94, Theorem 2] for the precise statement. To give a simple example, consider the function

$$\phi : \mathbb{R} \to \mathbb{R}_+;\ t \mapsto \sqrt{t^2 + \sin^2 t}.$$

This is indeed a screw function, because

$$\begin{aligned}\phi(t-s)^2 &= (t-s)^2 + \sin^2(t-s) \\ &= (t-s)^2 + \frac{1}{4}\big(\cos(2t) - \cos(2s)\big)^2 + \frac{1}{4}\big(\sin(2t) - \cos(2s)\big)^2.\end{aligned}$$

Note that a screw line is periodic if and only if it is not injective. Furthermore, one may identify screw lines with period $\tau > 0$ by the geometry of the support of the representing measure: this support must be contained in the lattice $(\pi/\tau)\mathbb{Z}_+$, where $\mathbb{Z}_+ := \mathbb{Z} \cap \mathbb{R}_+ = \{0, 1, 2, \ldots\}$. Consequently, all periodic screw lines in Hilbert space have a gauge function ϕ such that

$$\phi(t)^2 = \sum_{k=1}^{\infty} c_k \sin^2(k\pi t/\tau), \tag{2.2}$$

where $c_k \geq 0$ and $\sum_{k=1}^{\infty} c_k < \infty$; see [94, Theorem 5].

2.5 *Positive Definite Functions on Homogeneous Spaces*

Having resolved the question of isometrically embedding Euclidean space into Hilbert space, a natural desire was to extend the analysis to other special manifolds with symmetry. This was done almost simultaneously by Schoenberg on spheres [86] and by Bochner on compact homogeneous spaces [19].

Let X be a compact space endowed with a transitive action of a group G and an invariant measure. We seek G-invariant distance functions, and particularly those which identify X with a subspace of a Hilbert space. To simplify terminology, we call the latter *Hilbert distances*.

The first observation of Bochner is that a G-invariant symmetric kernel $f : X \times X \to \mathbb{R}$ satisfies the Hilbert-space embeddability condition,

$$\sum_{k=0}^{n} c_k = 0 \quad \Longrightarrow \quad \sum_{j,k=0}^{n} f(x_j, x_k) c_j c_k \geq 0,$$

for all choices of weights c_j and points $x_j \in X$, if and only if f is of the form

$$f(x, y) = h(x, y) - h(x_0, x_0) \qquad (x, y \in X),$$

where h is a G-invariant positive definite kernel and x_0 is a point of X. One implication is clear. For the other, we start with a G-invariant function f subject to the above constraint and prove, using G-invariance and integration over X, the existence of a constant c such that $h(x, y) = f(x, y) + c$ is a positive semidefinite kernel. This gives the following result.

Theorem 2.9 (Bochner [19]) *Let X be a compact homogeneous space. A continuous invariant function ρ on $X \times X$ is a Hilbert distance if and only if there exists a continuous, real-valued, invariant, positive definite kernel h on X and a point $x_0 \in X$, such that*

$$\rho(x, y) = \sqrt{h(x_0, x_0) - h(x, y)} \qquad (x, y \in X).$$

Privileged orthonormal bases of G-invariant functions, in the L^2 space associated with the invariant measure, provide a canonical decompositions of positive definite kernels. These generalized spherical harmonics were already studied by Cartan, Weyl, and von Neumann; see, for instance [96]. We elaborate on two important particular cases.

Let $X = \mathbb{T} = \{e^{i\theta} : \theta \in \mathbb{R}\}$ be the unit torus, endowed with the invariant arc-length measure. A continuous positive definite function $h : \mathbb{T} \times \mathbb{T} \to \mathbb{R}$ admits a Fourier decomposition

$$h(e^{ix}, e^{iy}) = \sum_{j,k \in \mathbb{Z}} a_{jk} e^{ijx} e^{-iky}.$$

If h is further required to be rotation invariant, we find that

$$h(e^{ix}, e^{iy}) = \sum_{k \in \mathbb{Z}} a_k e^{ik(x-y)},$$

where $a_k \geq 0$ for all $k \in \mathbb{Z}$ and $a_k = a_{-k}$ because h takes real values. Moreover, the series is Abel summable: $\sum_{k=0}^{\infty} a_k = h(1,1) < \infty$. Therefore, a rotation-invariant Hilbert distance ρ on the torus has the expression (after taking its square):

$$\begin{aligned}
\rho(e^{ix}, e^{iy})^2 = h(1,1) - h(e^{ix}, e^{iy}) &= \sum_{k=1}^{\infty} a_k (2 - e^{ik(x-y)} - e^{-ik(x-y)}) \\
&= 2 \sum_{k=1}^{\infty} a_k (1 - \cos k(x-y)) \\
&= 4 \sum_{k=1}^{\infty} a_k \sin^2(k(x-y)/2).
\end{aligned}$$

These are the periodic screw lines (2.2) already investigated by von Neumann and Schoenberg.

As a second example, we follow Bochner in examining a separable, compact group G. A real-valued, continuous, positive definite, and G-invariant kernel h admits the decomposition

$$h(x, y) = \sum_{k \in \mathbb{Z}} c_k \chi_k(yx^{-1}),$$

where $c_k \geq 0$ for all $k \in \mathbb{Z}$, $\sum_{k \in Z} c_k < \infty$ and χ_k denote the characters of irreducible representations of G. In conclusion, an invariant Hilbert distance ρ on G

is characterized by the formula

$$\rho(x, y)^2 = \sum_{k \in \mathbb{Z}} c_k \big(1 - \frac{\chi_k(yx^{-1}) + \chi_k(xy^{-1})}{2\chi_k(1)}\big),$$

where $c_k \geq 0$ and $\sum_{k \in \mathbb{Z}} c_k < \infty$.

For details and an analysis of similar decompositions on more general homogeneous spaces, we refer the reader to [19].

The above analysis of positive definite functions on homogeneous spaces was carried out separately by Schoenberg in [86]. First, he remarks that a continuous, real-valued, rotationally invariant, and positive definite kernel f on the sphere S^{d-1} has a distinguished Fourier-series decomposition with non-negative coefficients. Specifically,

$$f(\cos\theta) = \sum_{k=0}^{\infty} c_k P_k^{(\lambda)}(\cos\theta) \tag{2.3}$$

where $\lambda = (d-2)/2$, $P_k^{(\lambda)}$ are the ultraspherical orthogonal polynomials, $c_k \geq 0$ for all $k \geq 0$ and $\sum_{k=0}^{\infty} c_k < \infty$. This decomposition is in accord with Bochner's general framework, with the difference lying in Schoenberg's elementary proof, based on induction on dimension. As with all our formulas concerning the sphere, θ represents the geodesic distance (arc length along a great circle) between two points.

To convince the reader that expressions in the cosine of the geodesic distance are positive definite, let us consider points $x_1, \ldots, x_n \in S^{d-1}$. The Gram matrix with entries

$$\langle x_j, x_k \rangle = \cos\theta(x_j, x_k)$$

is obviously positive semidefinite, with constant diagonal elements equal to 1. According to the Schur product theorem [90], all functions of the form $\cos^k\theta$, where k is a non-negative integer, are therefore positive definite on the sphere.

At this stage, Schoenberg makes a leap forward and studies invariant positive definite kernels on S^∞, that is, functions $f(\cos\theta)$ which admit representations as above for all $d \geq 2$. His conclusion is remarkable in its simplicity.

Theorem 2.10 (Schoenberg [86]) *A real-valued function $f(\cos\theta)$ is positive definite on all spheres, independent of their dimension, if and only if*

$$f(\cos\theta) = \sum_{k=0}^{\infty} c_k \cos^k\theta, \tag{2.4}$$

where $c_k \geq 0$ for all $k \geq 0$ and $\sum_{k=0}^{\infty} c_k < \infty$.

This provides a return to the dominant theme, of isometric embedding into Hilbert space.

Corollary 2.11 *The function $\rho(\theta)$ is a Hilbert distance on S^∞ if and only if*

$$\rho(\theta)^2 = \sum_{k=0}^{\infty} c_k(1 - \cos^k \theta),$$

where $c_k \geq 0$ for all $k \geq 0$ and $\sum_{k=0}^{\infty} c_k < \infty$.

However, there is much more to derive from Schoenberg's theorem, once it is freed from the spherical context.

Theorem 2.12 (Schoenberg [86]) *Let $f : [-1, 1] \to \mathbb{R}$ be a continuous function. If the matrix $[f(a_{jk})]_{j,k=1}^n$ is positive semidefinite for all $n \geq 1$ and all positive semidefinite matrices $[a_{jk}]_{j,k=1}^n$ with entries in $[-1, 1]$, then, and only then,*

$$f(x) = \sum_{k=0}^{\infty} c_k x^k \qquad (x \in [-1, 1]),$$

where $c_k \geq 0$ for all $k \geq 0$ and $\sum_{k=0}^{\infty} c_k < \infty$.

Proof One implication follows from the Schur product theorem [90], which says that if the $n \times n$ matrices A and B are positive semidefinite, then so is their entrywise product $A \circ B := [a_{jk}b_{jk}]_{j,k=1}^n$. Indeed, inductively setting $B = A^{\circ k} = A \circ \cdots \circ A$, the k-fold entrywise power shows that every monomial x^k preserves positivity when applied entrywise. That the same property holds for functions $f(x) = \sum_{k \geq 0} c_k x^k$, with all $c_k \geq 0$, now follows from the fact that the set of positive semidefinite $n \times n$ matrices forms a closed convex cone, for all $n \geq 1$.

For the non-trivial, reverse implication we restrict the test matrices to those with leading diagonal terms all equal to 1. By interpreting such a matrix A as a Gram matrix, we identify n points on the sphere $x_1, \ldots, x_n \in S^{n-1}$ satisfying

$$a_{jk} = \langle x_j, x_k \rangle = \cos \theta(x_j, x_k) \quad (1 \leq j, k \leq n).$$

Then we infer from Schoenberg's theorem that f admits a uniformly convergent Taylor series with non-negative coefficients. □

We conclude this section by mentioning some recent avenues of research that start from Bochner's theorem (and its generalization in 1940, by Weil, Povzner, and Raikov, to all locally compact abelian groups) and Schoenberg's classification of positive definite functions on spheres. On the theoretical side, there has been a profusion of recent mathematical activity on classifying positive definite functions (and strictly positive definite functions) in numerous settings, mostly related to spheres [4, 5, 23, 98–100], two-point homogeneous spaces[2] [2, 3, 21], locally compact abelian groups and homogeneous spaces [28, 41], and products of these [11, 13, 40, 42–44].

Moreover, this line of work directly impacts applied fields. For instance, in climate science and geospatial statistics, one uses positive definite kernels and Schoenberg's results (and their sequels) to study trends in climate behavior on the Earth, since it can be modelled by a sphere, and positive definite functions on $S^2 \times \mathbb{R}$ characterize space-time covariance functions on it. See [39, 65, 71] for more details on these applications. Other applied fields include genomics and finance, through high-dimensional covariance estimation. We elaborate on this in the second part of the survey: see [10] or the full version [9, Chapter 7].

There are several other applications of Schoenberg's work on positive definite functions on spheres (his paper [86] has more than 160 citations) and we mention here just a few of them. Schoenberg's results were used by Musin [66] to compute the kissing number in four dimensions, by an extension of Delsarte's linear-programming method. Moreover, the results also apply to obtain new bounds on spherical codes [67], with further applications to sphere packing [25]. There are also applications to approximating functions and interpolating data on spheres, pseudodifferential equations with radial basis functions, and Gaussian random fields.

Remark 2.13 Another modern-day use of Schoenberg's results in [86] is in Machine Learning; see [91, 92], for example. Given a real inner-product space H and a function $f : \mathbb{R} \to \mathbb{R}$, an alternative notion of f being *positive definite* is as follows: for any finite set of vectors $x_1, \ldots, x_n \in H$, the matrix

$$[f(\langle x_j, x_k \rangle)]_{j,k=1}^n$$

is positive semidefinite. This is in contrast to the notion promoted by Bochner, Weil, Schoenberg, Pólya, and others, which concerns positivity of the matrix with entries $f(\langle x_j - x_k, x_j - x_k \rangle^{1/2})$. It turns out that every positive definite kernel on H, given by

$$(x, y) \mapsto f(\langle x, y \rangle)$$

for a function f which is positive definite in this alternate sense, gives rise to a reproducing-kernel Hilbert space, which is a central concept in Machine Learning. We restrict ourselves here to mentioning that, in this setting, it is desirable for the kernel to be strictly positive definite; see [68] for further clarification and theoretical results along these lines.

[2]Recall [95] that a metric space (X, ρ) is *n-point homogeneous* if, given finite sets $X_1, X_2 \subset X$ of equal size no more than n, every isometry from X_1 to X_2 extends to a self-isometry of X. This property was first considered by Birkhoff [15], and of course differs from the more common usage of the terminology of a homogeneous space G/H, whose study by Bochner was mentioned above.

2.6 Connections to Harmonic Analysis

Positivity and sharp continuity bounds for linear transformations between specific normed function spaces go hand in hand, especially when focusing on the kernels of integral transforms. The end of 1950s marked a fortunate condensation of observations, leading to a quasi-complete classification of preservers of positive or bounded convolution transforms acting on spaces of functions on locally compact abelian groups. In particular, these results can be interpreted as Schoenberg-type theorems for Toeplitz matrices or Toeplitz kernels. We briefly recount the main developments.

A groundbreaking theorem of the 1930s attributed to Wiener and Levy asserts that the pointwise inverse of a non-vanishing Fourier series with coefficients in L^1 exhibits the same summability behavior of the coefficient sequence. To be more precise, if ϕ is never zero and has the representation

$$\phi(\theta) = \sum_{n=-\infty}^{\infty} c_n e^{\mathrm{i}n\theta}, \quad \text{where} \sum_{n=-\infty}^{\infty} |c_n| < \infty,$$

then its reciprocal has a representation of the same form:

$$(1/\phi)(\theta) = \sum_{n=-\infty}^{\infty} d_n e^{\mathrm{i}n\theta}, \quad \text{where} \sum_{n=-\infty}^{\infty} |d_n| < \infty.$$

It was Gelfand [38] who in 1941 cast this permanence phenomenon in the general framework of commutative Banach algebras. Gelfand's theory applied to the Wiener algebra $W := \widehat{L^1(\mathbb{Z})}$ of Fourier transforms of L^1 functions on the dual of the unit torus proves the following theorem.

Theorem 2.14 (Gelfand [38]) *Let $\phi \in W$ and let $f(z)$ be an analytic function defined in a neighborhood of $\phi(\mathbb{T})$. Then $f(\phi) \in W$.*

The natural inverse question of deriving smoothness properties of inner transformations of Lebesgue spaces of Fourier transforms was tackled almost simultaneously by several analysts. For example, Rudin proved in 1956 [76] that a coefficient-wise transformation $c_n \mapsto f(c_n)$ mapping the space $\widehat{L^1(\mathbb{T})}$ into itself implies the analyticity of f in a neighborhood of zero. In a similar vein, Rudin and Kahane proved in 1958 [53] that a coefficient-wise transformation $c_n \mapsto f(c_n)$ which preserves the space of Fourier transforms $\widehat{M(\mathbb{T})}$ of finite measures on the torus implies that f is an entire function. In the same year, Kahane [52] showed that no quasi-analytic function (in the sense of Denjoy–Carleman) preserves the space $\widehat{L^1(\mathbb{Z})}$ and Katznelson [56] refined an inverse to Gelfand's theorem above, by showing the semi-local analyticity of transformers of elements of $\widehat{L^1(\mathbb{Z})}$ subject to some support conditions.

Soon after, the complete picture emerged in full clarity. It was unveiled by Helson, Kahane, Katznelson, and Rudin in an Acta Mathematica article [48]. Given a function f defined on a subset E of the complex plane, we say that f operates on the function algebra A, if $f(\phi) \in A$ for every $\phi \in A$ with range contained in E. The following metatheorem is proved in the cited article.

Theorem 2.15 (Helson–Kahane–Katznelson–Rudin [48]) *Let G be a locally compact abelian group and let Γ denote its dual, and suppose both are endowed with their respective Haar measures. Let $f : [-1, 1] \to \mathbb{C}$ be a function satisfying $f(0) = 0$.*

(1) *If Γ is discrete and f operates on $\widehat{L^1(G)}$, then f is analytic in some neighborhood of the origin.*
(2) *If Γ is not discrete and f operates on $\widehat{L^1(G)}$, then f is analytic in $[-1, 1]$.*
(3) *If Γ is not compact and f operates on $\widehat{M(G)}$, then f can be extended to an entire function.*

Rudin refined the above results to apply in the case of various L^p norms [78, 79], by stressing the lack of continuity assumption for the transformer f in all results (similar in nature to the statements in the above theorem). From Rudin's work we extract a highly relevant observation, *à la* Schoenberg's theorem, aligned to the spirit of the present survey.

Theorem 2.16 (Rudin [77]) *Suppose $f : (-1, 1) \to \mathbb{R}$ maps every positive semidefinite Toeplitz kernel with elements in $(-1, 1)$ into a positive semidefinite kernel:*

$$[a_{j-k}]_{j,k=-\infty}^{\infty} \geq 0 \qquad \Longrightarrow \qquad [f(a_{j-k})]_{j,k=-\infty}^{\infty} \geq 0.$$

Then f is absolutely monotonic, that is analytic on $(-1, 1)$ with a Taylor series having non-negative coefficients:

$$f(x) = \sum_{n=0}^{\infty} c_k x^k, \qquad \text{where } c_k \geq 0 \text{ for all } k \geq 0.$$

The converse is obviously true by the Schur product theorem. The elementary proof, quite independent of the derivation of the metatheorem stated above, is contained in [77]. Notice again the lack of a continuity assumption in the hypotheses.

In fact, Rudin proves more, by restricting the test domain of positive semidefinite Toeplitz kernels to the two-parameter family

$$a_n = \alpha + \beta \cos(n\theta) \qquad (n \in \mathbb{Z}) \tag{2.5}$$

with θ fixed so that θ/π is irrational and $\alpha, \beta \geq 0$ such that $\alpha + \beta < 1$. Rudin's proof commences with a mollifier argument to deduce the continuity of the transformer, then uses a development in spherical harmonics very similar to the original argument of Schoenberg. We will resume this topic in Sect. 3.3, setting it in a wider context.

With the advances in abstract duality theory for locally convex spaces, it is not surprising that proofs of Schoenberg-type theorems should be accessible with the aid of such versatile tools. We will confine ourselves here to mentioning one pertinent convexity-theoretic proof of Schoenberg's theorem, due to Christensen and Ressel [24].

Skipping freely over the details, the main observation of these two authors is that the multiplicatively closed convex cone of positivity preservers of positive semidefinite matrices of any size, with entries in $[-1, 1]$, is closed in the product topology of $\mathbb{R}^{[-1,1]}$, with a compact base K defined by the normalization $f(1) = 1$. The set of extreme points of K is readily seen to be closed, and an elementary argument identifies it as the set of all monomials x^n, where $n \geq 0$, plus the characteristic functions $\chi_1 \pm \chi_{-1}$. An application of Choquet's representation theorem now provides a proof of a generalization of Schoenberg's theorem, by removing the continuity assumption in the statement.

3 Entrywise Functions Preserving Positivity in All Dimensions

3.1 History

With the above history to place the present survey in context, we move to its dominant theme: entrywise positivity preservers. In analysis and in applications in the broader mathematical sciences, one is familiar with applying functions to the spectrum of diagonalizable matrices: $A = UDU^*$ then $f(A) = Uf(D)U^*$. More formally, one uses the Riesz–Dunford holomorphic functional calculus to define $f(A)$ for classes of matrices A and functions f.

Our focus in this survey will be on the parallel philosophy of *entrywise calculus*. To differentiate this from the functional calculus, we use the notation $f[A]$.

Definition 3.1 Fix a domain $I \subset \mathbb{C}$ and integers $m, n \geq 1$. Let $\mathcal{P}_n(I)$ denote the set of $n \times n$ Hermitian positive semidefinite matrices with all entries in I.

A function $f : I \to \mathbb{C}$ acts *entrywise* on a matrix

$$A = [a_{jk}]_{1 \leq j \leq m,\ 1 \leq k \leq n} \in I^{m \times n}$$

by setting

$$f[A] := [f(a_{jk})]_{1 \leq j \leq m,\ 1 \leq k \leq n} \in \mathbb{C}^{m \times n}.$$

Below, we allow the dimensions m and n to vary, while keeping the uniform notation $f[-]$.

We also let $\mathbf{1}_{m\times n}$ denote the $m \times n$ matrix with each entry equal to one. Note that $\mathbf{1}_{n\times n} \in \mathcal{P}_n(\mathbb{R})$.

In this survey, we explore the following overarching question in several different settings.

Which functions preserve positive semidefiniteness when applied entrywise to a class of positive matrices?

This question was first asked by Pólya and Szegö in their well-known book [70]. The authors observed that Schur's product theorem, together with the fact that the positive matrices form a closed convex cone, has the following consequence: if $f(x)$ is any power series with non-negative Maclaurin coefficients that converges on a domain $I \subset \mathbb{R}$, then f preserves positivity (that is, preserves positive semidefiniteness) when applied entrywise to positive semidefinite matrices with entries in I. Pólya and Szegö then asked if there are any other functions that possess this property. As discussed above, Schoenberg's theorem 2.12 provides a definitive answer to their question (together with the improvements by Rudin or Christensen–Ressel to remove the continuity hypothesis). Thanks to Pólya and Szegö's observation, Schoenberg's result may be considered as a rather challenging converse to the Schur product theorem.

In a similar vein, Rudin [77] observed that if one moves to the complex setting, then the conjugation map also preserves positivity when applied entrywise to positive semidefinite complex matrices. Therefore the maps

$$z \mapsto z^j \overline{z}^k \qquad (j, k \geq 0)$$

preserve positivity when applied entrywise to complex matrices of all dimensions, again by the Schur product theorem. The same property is now satisfied by non-negative linear combinations of these functions. In [77], Rudin made this observation and conjectured, *à la* Pólya–Szegö, that these are all of the preservers. This was proved by Herz in 1963.

Theorem 3.2 (Herz [49]) *Let $D(0,1)$ denote the open unit disc in $\mathbb{C}$, and suppose $f : D(0,1) \to \mathbb{C}$. The entrywise map $f[-]$ preserves positivity on $\mathcal{P}_n\big(D(0,1)\big)$ for all $n \geq 1$, if and only if*

$$f(z) = \sum_{j,k\geq 0} c_{jk} z^j \overline{z}^k \qquad \text{for all } z \in D(0,1),$$

where $c_{jk} \geq 0$ for all $j, k \geq 0$.

Akin to the above results by Schoenberg, Rudin, Christensen and Ressel, and Herz, we mention one more Schoenberg-type theorem, for matrices with positive entries. The following result again demonstrates the rigid principle that analyticity and absolute monotonicity follow from the preservation of positivity in all dimensions.

Theorem 3.3 (Vasudeva [93]) *Let $f : (0, \infty) \to \mathbb{R}$. Then $f[-]$ preserves positivity on $\mathcal{P}_n\big((0, \infty)\big)$ for all $n \geq 1$, if and only if $f(x) = \sum_{k=0}^{\infty} c_k x^k$ on $(0, \infty)$, where $c_k \geq 0$ for all $k \geq 0$.*

3.2 The Horn–Loewner Necessary Condition in Fixed Dimension

The previous section contains several variants of a "dimension-free" result: namely, the classification of entrywise maps that preserve positivity on test sets of matrices of all sizes. In the next section, we discuss a dimension-free result that parallels Rudin's work in [77], by approaching the problem via preservers of moment sequences for positive measures on the real line. In other words, we will work with Hankel instead of Toeplitz matrices.

In the later part of this survey, we focus on entrywise functions that preserve positivity when the test set consists of matrices of a fixed size. For both of these settings, the starting point is an important result found in the PhD thesis of Roger Horn, which he attributes to his advisor, Charles Loewner.

Theorem 3.4 ([50]) *Let $f : (0, \infty) \to \mathbb{R}$ be continuous. Fix a positive integer n and suppose $f[-]$ preserves positivity on $\mathcal{P}_n\big((0, \infty)\big)$. Then $f \in C^{n-3}(I)$,*

$$f^{(k)}(x) \geq 0 \qquad \textit{whenever } x \in (0, \infty) \textit{ and } 0 \leq k \leq n-3,$$

and $f^{(n-3)}$ is a convex non-decreasing function on $(0, \infty)$. Furthermore, if $f \in C^{n-1}\big((0, \infty)\big)$, then $f^{(k)}(x) \geq 0$ whenever $x \in (0, \infty)$ and $0 \leq k \leq n-1$.

This result and its variations are the focus of the present section.

Theorem 3.4 is remarkable for several reasons.

(1) Modulo variations, it remains to this day the only known criterion for a general entrywise function to preserve positivity in a fixed dimension. Later on, we will see more precise conclusions drawn when f is a polynomial or a power function, but for a general function there are essentially no other known results.
(2) While Theorem 3.4 is a fixed-dimension result, it can be used to prove some of the aforementioned dimension-free characterizations. For instance, if $f[-]$ preserves positivity on $\mathcal{P}_n\big((0, \infty)\big)$ for all $n \geq 1$, then, by Theorem 3.4, the function f is absolutely monotonic on $(0, \infty)$. A classical result of Bernstein on absolutely monotonic functions now implies that f is necessarily given by a power series with non-negative coefficients, which is precisely Vasudeva's Theorem 3.3.

    In the next section, we will outline an approach to prove a stronger version of Schoenberg's Theorem 2.12 (in the spirit of Theorem 2.16 by Rudin), starting from Theorem 3.3.

(3) Theorem 3.4 is also significant because there is a sense in which it is sharp. We elaborate on this when studying polynomial and power-function preservers; this is discussed in the second part of the survey: see [10] or [9, Chapters 4 and 6].

Remark 3.5 There are other, rather unexpected consequences of Theorem 3.4 as well. It was recently shown that the key determinant computation underlying Theorem 3.4 can be generalized to yield a new class of symmetric function identities for any formal power series. The only such identities previously known were for the case $f(x) = \frac{1-cx}{1-x}$. This is discussed in the second part of this survey [10] and the full version [9, Section 4.6].

We next explain the steps behind the proof of the Horn–Loewner Theorem 3.4. These also help in proving certain strengthenings of Theorem 3.4, which are mentioned below. In turn, these strengthenings additionally serve to clarify the nature of the Horn–Loewner necessary condition.

Proof of Theorem 3.4 The proof by Loewner is in two steps. First he assumes f to be smooth and shows the result by induction on n. The base case of $n = 1$ is immediate, and for the induction step one proceeds as follows. Fix $a > 0$, choose any vector $\mathbf{u} = (u_1, \ldots, u_n)^T \in \mathbb{R}^n$ with distinct coordinates, and define

$$\Delta(t) := \det[f(a + tu_j u_k)]_{j,k=1}^n = \det f[a\mathbf{1}_{n\times n} + \mathbf{u}\mathbf{u}^T] \qquad (0 < t \ll 1).$$

Then Loewner shows that

$$\begin{aligned} \Delta(0) = \Delta'(0) = \cdots = \Delta^{\binom{n}{2}-1}(0) = 0, \\ \Delta^{\binom{n}{2}}(0) = cf(a)f'(a)\cdots f^{(n-1)}(a) \qquad \text{for some } c > 0. \end{aligned} \tag{3.1}$$

(See Remark 3.5 above.)

Returning to the proof of Theorem 3.4 for smooth functions: apply the above treatment not to f but to $g_\tau(x) := f(x) + \tau x^n$, where $\tau > 0$. By the Schur product theorem, g_τ satisfies the hypotheses, whence $\Delta(t)/t^{\binom{n}{2}} \geq 0$ for $t > 0$. Taking $t \to 0^+$, by L'Hôpital's rule we obtain

$$g_\tau(a)g_\tau'(a)\cdots g_\tau^{(n-1)}(a) \geq 0, \qquad \text{for all } \tau > 0.$$

Finally, the induction hypothesis implies that $f, f', \ldots, f^{(n-2)}$ are non-negative at a, whence $g_\tau(a), \ldots, g_\tau^{(n-2)}(a) > 0$. It follows that $g_\tau^{(n-1)}(a) \geq 0$ for all $\tau > 0$, and hence, $f^{(n-1)}(a) \geq 0$, as desired.

Remark 3.6 The above argument is amenable to proving more refined results. For example, it can be used to prove the positivity of the first n non-zero derivatives of a smooth preserver f; see Theorem 3.10.

The second step of Loewner's proof begins by using mollifiers. Suppose f is continuous; approximate it by a mollified family $f_\delta \to f$ as $\delta \to 0^+$. Thus f_δ is smooth and its first n derivatives are non-negative on $(0, \infty)$. By the mean-value theorem for divided differences, this implies that the divided differences of each f_δ, of orders up to $n-1$ are non-negative. Since f is continuous, the same holds for f.

Now one invokes a rather remarkable result by Boas and Widder [17], which can be viewed as a converse to the mean-value theorem for divided differences. It asserts that given an integer $k \geq 2$ and an open interval $I \subset \mathbb{R}$, if all kth order "equi-spaced" forward differences (whence divided differences) of a continuous function $f : I \to \mathbb{R}$ are non-negative on I, then f is $k-2$ times differentiable on I; moreover, $f^{(k-2)}$ is continuous and convex on I, with non-decreasing left- and right-hand derivatives. Applying this result for each $2 \leq k \leq n-1$ concludes the proof of Theorem 3.4. $\square$

Note that this proof only uses matrices of the form $a\mathbf{1}_{n\times n} + t\mathbf{u}\mathbf{u}^T$, and the arguments are all local. Thus it is unsurprising that strengthened versions of the Horn–Loewner theorem can be found in the literature; see [7, 47], for example. We present here the stronger of these variants.

Theorem 3.7 (See [7, Section 3]) *Suppose* $0 < \rho \leq \infty$, $I = (0, \rho)$, *and* $f : I \to \mathbb{R}$. *Fix* $u_0 \in (0, 1)$ *and an integer* $n \geq 1$, *and define* $\mathbf{u} := (1, u_0, \ldots, u_0^{n-1})^T$. *Suppose* $f[A] \in \mathcal{P}_2(\mathbb{R})$ *for all* $A \in \mathcal{P}_2(I)$, *and also that* $f[A] \in \mathcal{P}_n(\mathbb{R})$ *for all Hankel matrices* $A = a\mathbf{1}_{n\times n} + t\mathbf{u}\mathbf{u}^T$, *with* $a, t \geq 0$ *such that* $a + t \in I$. *Then the conclusions of Theorem 3.4 hold.*

Beyond the above strengthenings, the notable feature here is that the continuity hypothesis has been removed, akin to the Rudin and Christensen–Ressel results. We reproduce here an elegant argument to show continuity; this can be found in Vasudeva's paper [93], and uses only the test set $\mathcal{P}_2(I)$. By considering $f[A]$ for $A = \begin{bmatrix} a & b \\ b & a \end{bmatrix}$ with $0 < b < a < \rho$, it follows that f is non-negative and non-decreasing on I. One also shows that f is either identically zero or never zero on I. In the latter case, considering $f[A]$ for $A = \begin{bmatrix} a & \sqrt{ab} \\ \sqrt{ab} & b \end{bmatrix} \in \mathcal{P}_2(I)$ shows that f is *multiplicatively mid-convex*: the function

$$g(y) := \log f(e^y) \qquad (y < \log \rho)$$

is midpoint convex and locally bounded on the interval $\log I$. Now the following classical result [74, Theorem 71.C] shows that g is continuous on $\log I$, so f is continuous on I.

Proposition 3.8 *Let* U *be a convex open set in a real normed linear space. If* $g : U \to \mathbb{R}$ *is midpoint convex on* U *and bounded above in an open neighborhood of a single point in* U, *then* g *is continuous, so convex, on* U.

We now move to variants of the Horn–Loewner result. Notice that Theorems 3.4 and 3.7 are results for arbitrary positivity preservers $f(x)$. When more is known about f, such as smoothness or even real analyticity, stronger conclusions can be drawn from smaller test sets of matrices. A recent variant is the following lemma, shown by evaluating $f[-]$ at matrices $(tu_j u_k)_{j,k=1}^n$ and using the invertibility of "generic" generalized Vandermonde matrices.

Lemma 3.9 (Belton–Guillot–Khare–Putinar [6] and Khare–Tao [58]) *Let* $n \geq 1$ *and* $0 < \rho \leq \infty$. *Suppose* $f(x) = \sum_{k \geq 0} c_k x^k$ *is a convergent power series on* $I = [0, \rho)$ *that is positivity preserving entrywise on rank-one matrices in* $\mathcal{P}_n(I)$. *Further assume that* $c_{m'} < 0$ *for some* m'.

(1) *If* $\rho < \infty$, *then we have* $c_m > 0$ *for at least* n *values of* $m < m'$. *(In particular, the first* n *non-zero Maclaurin coefficients of* f, *if they exist, must be positive.)*
(2) *If instead* $\rho = \infty$, *then we have* $c_m > 0$ *for at least* n *values of* $m < m'$ *and at least* n *values of* $m > m'$. *(In particular, if* f *is a polynomial, then the first* n *non-zero coefficients and the last* n *non-zero coefficients of* f, *if they exist, are all positive.)*

Notice that this lemma (a) talks about the derivatives of f at 0 and not in $(0, \rho)$; and moreover, (b) considers not the first few derivatives, but the first few *non-zero* derivatives. Thus, it is morally different from the preceding two theorems, and one naturally seeks a common unification of these three results. This was recently achieved.

Theorem 3.10 (Khare [57]) *Let* $0 \leq a < \infty$, $\epsilon \in (0, \infty)$, $I = [a, a + \epsilon)$, *and let* $f : I \to \mathbb{R}$ *be smooth. Fix integers* $n \geq 1$ *and* $0 \leq p \leq q \leq n$, *with* $p = 0$ *if* $a = 0$, *and such that* $f(x)$ *has* $q - p$ *non-zero derivatives at* $x = a$ *of order at least* p. *Now let*

$$m_0 := 0, \quad \dots \quad m_{p-1} := p - 1;$$

suppose further that

$$p \leq m_p < m_{p+1} < \cdots < m_{q-1}$$

are the lowest orders (above p*) of the first* $q - p$ *non-zero derivatives of* $f(x)$ *at* $x = a$.

Also fix distinct scalars $u_1, \dots, u_n \in (0, 1)$, *and let* $\mathbf{u} := (u_1, \dots, u_n)^T$. *If* $f[a\mathbf{1}_{n \times n} + t\mathbf{u}\mathbf{u}^T] \in \mathcal{P}_n(\mathbb{R})$ *for all* $t \in [0, \epsilon)$, *then the derivative* $f^{(k)}(a)$ *is non-negative whenever* $0 \leq k \leq m_{q-1}$.

Notice that varying p allows one to control the number of initial derivatives versus the number of subsequent non-zero derivatives of smallest order. In particular, if $p = q = n$, then the result implies the "stronger" Horn–Loewner Theorem 3.7 (and so Theorem 3.4) pointwise at every $a > 0$. At the other extreme is the special case of $p = 0$ (at any $a \geq 0$), which strengthens the conclusions of Theorems 3.4 and 3.7 for smooth functions.

Corollary 3.11 *Suppose* a, ϵ, I, f, n, *and* $\mathbf{u}$ *are as in Theorem 3.10. If* $f[a\mathbf{1}_{n \times n} + t\mathbf{u}\mathbf{u}^T] \in \mathcal{P}_n(\mathbb{R})$ *for all* $t \in [0, \epsilon)$, *then the first* n *non-zero derivatives of* $f(x)$ *at* $x = a$ *are positive.*

Remark 3.12 Theorem 3.10 further clarifies the nature of the Horn–Loewner result and its proof. The reduction from arbitrary functions, to continuous functions, to smooth functions, requires an open domain $(0, \rho)$, in order to use mollifiers, for example. However, the result for smooth functions actually holds pointwise, as shown by Theorem 3.10.

The proof of Theorem 3.10 combines novel arguments together with the previously mentioned techniques of Loewner. The refinement of the determinant computations (3.1) is of particular note; see the second part of this survey ([10] or [9, Section 4.6]).

3.3 Schoenberg Redux: Moment Sequences and Hankel Matrices

In this section, we outline another approach to proving Schoenberg's Theorem 2.12, which yields a stronger version parallel to the strengthening by Rudin of Theorem 2.16. The present section reveals connections between positivity preservers, totally non-negative Hankel matrices, moment sequences of positive measures on the real line, and also a connection to semi-algebraic geometry.

We begin with Rudin's Theorem 2.16 and the family (2.5). Notice that the positive definite sequences in (2.5) give rise to the Toeplitz matrices $A(n, \alpha, \beta, \theta)$ with (j, k) entry equal to $\alpha + \beta \cos\big((j - k)\theta\big)$. From the elementary identity

$$\cos(p - q) = \cos p \cos q + \sin p \sin q \qquad (p, q \in \mathbb{R}),$$

it follows that these Toeplitz matrices have rank at most three:

$$A(n, \alpha, \beta, \theta) = \alpha \mathbf{1}_{n\times n} + \beta \mathbf{u}\mathbf{u}^T + \beta \mathbf{v}\mathbf{v}^T, \tag{3.2}$$

where

$$\mathbf{u} := \big(\cos\theta, \cos(2\theta), \ldots, \cos(n\theta)\big)^T \text{ and } \mathbf{v} := \big(\sin\theta, \sin(2\theta), \ldots, \sin(n\theta)\big)^T.$$

In particular, Rudin's work (see Theorem 2.16 and the subsequent discussion) implies the following result.

Proposition 3.13 *Let $\theta \in \mathbb{R}$ such that θ/π is irrational. An entrywise map $f : \mathbb{R} \to \mathbb{R}$ preserves positivity on the set of Toeplitz matrices*

$$\{A(n, \alpha, \beta, \theta) : n \geq 1,\ \alpha, \beta > 0\}$$

if and only if $f(x) = \sum_{k=0}^\infty c_k x^k$ is a convergent power series on $\mathbb{R}$, with $c_k \geq 0$ for all $k \geq 0$.

Thus, one can significantly reduce the set of test matrices.

Proof Given $0 < \rho < \infty$, let the restriction $f_\rho := f|_{(-\rho,\rho)}$. Observe from the discussion following Theorem 2.16 that Rudin's work explicitly shows the result for f_1, whence for any f_ρ by a change of variables. Thus,

$$f_\rho(x) = \sum_{k=0}^{\infty} c_{k,\rho} x^k, \qquad c_{k,\rho} \geq 0 \text{ for all } k \geq 0 \text{ and } \rho > 0.$$

Given $0 < \rho < \rho' < \infty$, it follows by the identity theorem that $c_{k,\rho} = c_{k,\rho'}$ for all k. Hence $f(x) = \sum_{k\geq 0} c_{k,1} x^k$ (which was Rudin's $f_1(x)$), now on all of $\mathbb{R}$. $\square$

In a parallel vein to Rudin's results and Proposition 3.13, the following strengthening of Schoenberg's result can be shown, using a different (and perhaps more elementary) approach than those of Schoenberg and Rudin.

Theorem 3.14 (Belton–Guillot–Khare–Putinar [7]) *Suppose* $0 < \rho \leq \infty$ *and* $I = (-\rho, \rho)$. *Then the following are equivalent for a function* $f : I \to \mathbb{R}$.

(1) *The entrywise map* $f[-]$ *preserves positivity on* $\mathcal{P}_n(I)$, *for all* $n \geq 1$.
(2) *The entrywise map* $f[-]$ *preserves positivity on the Hankel matrices in* $\mathcal{P}_n(I)$ *of rank at most* 3, *for all* $n \geq 1$.
(3) *The function* f *is real analytic on* I *and absolutely monotonic on* $(0, \rho)$. *In other words,* $f(x) = \sum_{k\geq 0} c_k x^k$ *on* I, *with* $c_k \geq 0\ \forall k$.

Remark 3.15 Recall the alternate notion of positive definite functions discussed in Remark 2.13. In [68] and related works, Pinkus and other authors study this alternate notion of positive definite functions on H. Notice that such matrices form precisely the set of positive semidefinite symmetric matrices of rank at most $\dim H$. In particular, Theorem 3.14 and the far earlier 1959 paper [77] of Rudin both provide a characterization of these functions, on every Hilbert space of dimension 3 or more.

Parallel to the discussions of the proofs of Schoenberg's and Rudin's results (see the previous chapter), we now explain how to prove Theorem 3.14. Clearly, (3) $\implies$ (1) $\implies$ (2) in the theorem. We first outline how to weaken the condition (2) even further and still imply (3). The key idea is to consider *moment sequences* of certain non-negative measures on the real line. This parallels Rudin's considerations of Fourier–Stieltjes coefficients of non-negative measures on the circle.

Definition 3.16 A measure μ with support in $\mathbb{R}$ is said to be *admissible* if $\mu \geq 0$ on $\mathbb{R}$, and all moments of μ exist and are finite:

$$s_k(\mu) := \int_{\mathbb{R}} x^k \, \mathrm{d}\mu(x) < \infty \qquad (k \geq 0).$$

The sequence $\mathbf{s}(\mu) := \big(s_k(\mu)\big)_{k=0}^{\infty}$ is termed the *moment sequence* of μ. Corresponding to μ and this moment sequence is the *moment matrix* of μ:

$$H_\mu := \begin{bmatrix} s_0(\mu) & s_1(\mu) & s_2(\mu) & \cdots \\ s_1(\mu) & s_2(\mu) & s_3(\mu) & \cdots \\ s_2(\mu) & s_3(\mu) & s_4(\mu) & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix};$$

note that $H_\mu = [s_{i+j}(\mu)]_{i,j\geq 0}$ is a semi-infinite Hankel matrix. Finally, a function $f : \mathbb{R} \to \mathbb{R}$ acts entrywise on moment sequences, to yield real sequences:

$$f[\mathbf{s}(\mu)] := \big(f\big(s_0(\mu)\big), \ldots, f\big(s_k(\mu)\big), \ldots\big).$$

We are interested in understanding which entrywise functions preserve the space of moment sequences of admissible measures. The connection to positive semidefinite matrices is made through Hamburger's theorem, which says that a real sequence $(s_0, s_1, \ldots)$ is the moment sequence of an admissible measure on $\mathbb{R}$ if and only if every (finite) principal minor of the moment matrix H_μ is positive semidefinite. For simplicity, this last will be reformulated below to saying that H_μ is positive semidefinite.

The weakening of Theorem 3.14(2) is now explained: it suffices to consider the reduced test set of those Hankel matrices, which arise as the moment matrices of admissible measures supported at three points. Henceforth, let δ_x denote the Dirac probability measure supported at $x \in \mathbb{R}$. It is not hard to verify that the m-point measure $\mu = \sum_{j=1}^m c_j \delta_{x_j}$ has Hankel matrix H_μ with rank no more than m:

$$\begin{aligned} & s_k(\mu) = \sum_{j=1}^{m} c_j x_j^k \qquad (k \geq 0) \\ \implies \quad & H_\mu = \sum_{j=1}^{m} c_j \mathbf{u}_j \mathbf{u}_j^T, \text{ where } \mathbf{u}_j := (1, x_j, x_j^2, \ldots)^T. \end{aligned} \tag{3.3}$$

Thus, a further strengthening of Schoenberg's result is as follows.

Theorem 3.17 (Belton–Guillot–Khare–Putinar [7]) *In the setting of Theorem 3.14, the three assertions contained therein are also equivalent to*

(4) *For each measure*

$$\mu = a\delta_1 + b\delta_{u_0} + c\delta_{-1}, \quad \textit{with } u_0 \in (0,1),\ a, b, c \geq 0,\ a+b+c \in (0,\rho), \tag{3.4}$$

there exists an admissible measure σ_μ on $\mathbb{R}$ such that $f\big(s_k(\mu)\big) = s_k(\sigma_\mu)$ for all $k \geq 0$.

In fact, we will see in Sect. 3.4 below that this assertion (4) can be simplified to just assert that $f[H_\mu]$ is positive semidefinite, and so completely avoid the use of Hamburger's theorem.

We now discuss the proof of these results, working with $\rho = \infty$ for ease of exposition. The first observation is that the strengthening of the Horn–Loewner Theorem 3.7, together with the use of Bernstein's theorem (see remark (2) following Theorem 3.4), implies the following "stronger" form of Vasudeva's Theorem 3.3:

Theorem 3.18 (See [7]) *Suppose $I = (0, \infty)$ and $f : I \to \mathbb{R}$. Also fix $u_0 \in (0, 1)$. The following are equivalent:*

(1) *The entrywise map $f[-]$ preserves positivity on $\mathcal{P}_n(I)$ for all $n \geq 1$.*
(2) *The entrywise map $f[-]$ preserves positivity on all moment matrices H_μ for $\mu = a\delta_1 + b\delta_{u_0},\ a, b > 0$.*
(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 $k \geq 0$.*

Notice that the test matrices in assertion (2) are all Hankel, and of rank at most two. This severely weakens Vasudeva's original hypotheses.

Now suppose the assertion in Theorem 3.17(4) holds. By the preceding result, $f(x)$ is given on $(0, \infty)$ by an absolutely monotonic function $\sum_{k \geq 0} c_k x^k$. The next step is to show that f is continuous. For this, we will crucially use the following "integration trick." Suppose for each admissible measure μ as in (3.4), there is a non-negative measure σ_μ supported on $[-1, 1]$ such that $f\big(s_k(\mu)\big) = s_k(\sigma_\mu)$ for all $k \geq 0$. (Note here that it is not immediate that the support is contained in $[-1, 1]$.)

Now let $p(t) = \sum_{k \geq 0} b_k t^k$ be a polynomial that takes non-negative values on $[-1, 1]$. Then,

$$0 \leq \int_{-1}^{1} p(t)\,\mathrm{d}\sigma_\mu(t) = \sum_{k=0}^{\infty} \int_{-1}^{1} b_k t^k\,\mathrm{d}\sigma_\mu(t) = \sum_{k=0}^{\infty} b_k s_k(\sigma_\mu) = \sum_{k=0}^{\infty} b_k f\big(s_k(\mu)\big). \tag{3.5}$$

Remark 3.19 For example, suppose $p(t) = 1 - t^d$ for some $d \geq 1$. If $\mu = a\delta_1 + b\delta_{u_0} + c\delta_{-1}$, where $u_0 \in (0, 1)$ and $a, b, c > 0$, then the inequality (3.5) gives that

$$0 \leq f\big(s_0(\mu)\big) - f\big(s_d(\mu)\big) = f(a + b + c) - f(a + b u_0^d + c(-1)^d).$$

It is not clear *a priori* how to deduce this inequality using the fact that $f[-]$ preserves matrix positivity and the Hankel moment matrix of μ. The explanation, which we provide in Sect. 3.4 below, connects moment problems, matrix positivity, and real algebraic geometry.

We now outline how (3.5) can be used to prove the continuity of f. First note that $|s_k(\mu)| \leq s_0(\mu)$ for μ as above and all $k \geq 0$. This fact and the easy observation that f is bounded on compact subsets of $\mathbb{R}$ together imply that all moments of σ_μ are uniformly bounded. From this we deduce that σ_μ is necessarily supported on $[-1, 1]$.

The inequality (3.5) now gives the left-continuity of f at $-\beta$, for every $\beta \geq 0$. Fix $u_0 \in (0, 1)$, and let

$$\mu_b := (\beta + bu_0)\delta_{-1} + b\delta_{u_0} \qquad (b > 0).$$

Applying (3.5) to the polynomials $p_{\pm,1}(t) := (1 \pm t)(1 - t^2)$, we deduce that

$$f\big(\beta + b(1 + u_0)\big) - f\big(\beta + b(u_0 + u_0^2)\big) \geq |f(-\beta) - f\big(-\beta - bu_0(1 - u_0^2)\big)|.$$

Letting $b \to 0^+$, the left continuity of f at $-\beta$ follows. Similarly, to show that f is right continuous at $-\beta$, we apply the integral trick to $p_{\pm,1}(t)$ and to $\mu_b' := (\beta + bu_0^3)\delta_{-1} + b\delta_{u_0}$ instead of μ_b.

Having shown continuity, to prove the stronger Schoenberg theorem, we next assume that f is smooth on $\mathbb{R}$. For all $a \in \mathbb{R}$, define the function

$$H_a : \mathbb{R} \to \mathbb{R}; \; x \mapsto f(a + e^x).$$

The function H_a satisfies the estimates

$$|H_a^{(n)}(x)| \leq H_{|a|}^{(n)}(x) \qquad (a, x \in \mathbb{R}, \; n \in \mathbb{Z}_+). \tag{3.6}$$

This is shown by another use of the integration trick (3.5), this time for the polynomials $p_{\pm,n}(t) := (1 \pm t)(1 - t^2)^n$ for all $n \geq 0$. In turn, the estimates (3.6) lead to showing that H_a is real analytic on $\mathbb{R}$, for all $a \in \mathbb{R}$. Now composing H_{-a} for $a > |x|$ with the function $L_a(y) := \log(a + y)$ shows that $f(x)$ is real analytic on $\mathbb{R}$ and agrees with $\sum_{k \geq 0} a_k x^k$ on $(0, \infty)$. This concludes the proof for smooth functions.

Finally, to pass from smooth functions to continuous functions, we again use a mollified family $f_\delta \to f$ as $\delta \to 0^+$. Each f_δ is the restriction of an entire function, say $\widetilde{f_\delta}$, and the family $\{\widetilde{f_{1/n}} : n \geq 1\}$ forms a normal family on each open disc $D(0, r)$. It follows from results by Montel and Morera that $\widetilde{f_{1/n}}(z)$ converges uniformly to a function g_r on each closed disc $\overline{D(0, r)}$, and g_r is analytic. Since g_r restricts to f on $(-r, r)$, it follows that f is necessarily also real analytic on $\mathbb{R}$, and we are done.

3.4 The Integration Trick and Positivity Certificates

Observe that the inequality (3.5) can be written more generally as follows.

Given a polynomial $p(t) = \sum_{k \geq 0} b_k t^k$ *which takes non-negative values on* $[-1, 1]$, *as well as a positive semidefinite Hankel matrix* $H = (s_{i+j})_{i,j \geq 0}$, *we have that*

$$\sum_{k \geq 0} b_k s_k \geq 0. \tag{3.7}$$

As shown in (3.5), this assertion is clear via an application of Hamburger's theorem. We now demonstrate how the assertion can instead be derived from first principles, with interesting connections to positivity certificates.

First note that the inequality (3.7) holds if $p(t)$ is the square of a polynomial. For instance, if $p(t) = (1-3t)^2 = 1 - 6t + 9t^2$ on $[-1, 1]$, then

$$s_0 - 6s_1 + 9s_2 = (e_0 - 3e_1)^T H(e_0 - 3e_1), \tag{3.8}$$

where $e_0 = (1, 0, 0, \dots)$ and $e_1 = (0, 1, 0, 0, \dots)$. The non-negativity of (3.8) now follows immediately from the positivity of the matrix H. The same reasoning applies if $p(t)$ is a sum of squares of polynomials, or even the limit of a sequence of sums of squares. Thus, one approach to showing the inequality (3.7) for an arbitrary polynomial $p(t)$ which is non-negative on $[-1, 1]$ is to seek a *limiting sum-of-squares representation*, which is also known as a *positivity certificate*, for p.

If a d-variate real polynomial is a sum of squares of real polynomials, then it is clearly non-negative on $\mathbb{R}^d$, but the converse is not true for $d > 1$.[3] Even when $d = 1$, while a sum-of-squares representation is an equivalent characterization for one-variable polynomials that are non-negative on $\mathbb{R}$, here we are working on the compact semi-algebraic set $[-1, 1]$. We now give three proofs of the existence of such a positivity certificate in the setting used above.

Proof 1. A result of Berg, Christensen, and Ressel (see the end of [12]) shows more generally that, for every dimension $d \geq 1$, any non-negative polynomial on $[-1, 1]^d$ has a limiting sum-of-squares representation. □

Proof 2. The only polynomials used in proving the stronger form of Schoenberg's theorem, Theorems 3.14 and 3.17, appear following (3.6):

$$p_{\pm,n}(t) := (1 \pm t)(1 - t^2)^n \qquad (n \geq 0).$$

Each of these polynomials is composed of factors of the form $p_{\pm,0}(t) = 1 \pm t$, so it suffices to produce a limiting sum-of-squares representation for these two polynomials on $[-1, 1]$. Note that

$$\begin{aligned}
\frac{1}{2}(1 \pm t)^2 &= \frac{1}{2} \pm t + \frac{t^2}{2}, \\
\frac{1}{4}(1 - t^2)^2 &= \frac{1}{4} - \frac{t^2}{2} + \frac{t^4}{4}, \\
\frac{1}{8}(1 - t^4)^2 &= \frac{1}{8} - \frac{t^4}{4} + \frac{t^8}{8},
\end{aligned}$$

[3] This is connected to semi-algebraic geometry and to Hilbert's seventeenth problem: recall the famous result of Motzkin that there are non-negative polynomials on $\mathbb{R}^d$ that are not sums of squares, such as $x^4y^2 + x^2y^4 - 3x^2y^2 + 1$. Such phenomena have been studied in several settings, including polytopes (by Farkas, Handelman, and Pólya) and more general semi-algebraic sets (by Putinar, Schmüdgen, Stengel, Vasilescu, and others).

and so on. Adding the first n equations shows that $(1 \pm t) + 2^{-n}(t^{2^n} - 1)$ is a sum-of-squares polynomial for all n. Taking $n \to \infty$ finishes the proof. $\square$

Proof 3. In fact, for any $d \geq 1$ and any compact set $K \subset \mathbb{R}^d$, if f is a non-negative continuous function on K, then f has a positivity certificate. The Stone–Weierstrass theorem gives a sequence of polynomials which converges to $\sqrt{f}$, and the squares of these polynomials then provide the desired limiting representation for f. This is a simpler proof than Proof 1 from [12], but the convergence here is uniform, whereas the convergence in [12] is stronger. $\square$

Remark 3.20 In (3.5), we used $H = H_{\sigma_\mu}$, which was positive semidefinite by assumption. The previous discussion shows that Theorem 3.17(4) can be further weakened, by requiring only that $f[H_\mu]$ is positive semidefinite, as opposed to being equal to H_σ for some admissible measure σ. Hence we do not require Hamburger's theorem in order to prove the strengthening of Schoenberg's theorem that uses the test set of low-rank Hankel matrices.

3.5 Variants of Moment-Sequence Transforms

We now present a trio of results on functions which preserve moment sequences.

For $K \subset \mathbb{R}$, let $\mathcal{M}(K)$ denote the set of moment sequences corresponding to admissible measures with support in K. We say that F maps $\mathcal{M}(K)$ into $\mathcal{M}(L)$, where $K, L \subset \mathbb{R}$, if for every admissible measure μ with support in K there exists an admissible measure σ with support in L such that

$$F(s_k(\mu)) = s_k(\sigma) \quad \text{for all } k \in \mathbb{Z}_+,$$

where $s_k(\mu)$ is the kth-power moment of μ, as in Definition 3.16.

Theorem 3.21 *A function $F : \mathbb{R} \to \mathbb{R}$ maps $\mathcal{M}([-1, 1])$ into itself if and only if F is the restriction to $\mathbb{R}$ of an absolutely monotonic entire function.*

Theorem 3.22 *A function $F : \mathbb{R}_+ \to \mathbb{R}$ maps $\mathcal{M}([0, 1])$ into itself if and only if F is absolutely monotonic on $(0, \infty)$ and $0 \leq F(0) \leq \lim_{\epsilon \to 0^+} F(\epsilon)$.*

Theorem 3.23 *A function $F : \mathbb{R} \to \mathbb{R}$ maps $\mathcal{M}([-1, 0])$ into $\mathcal{M}((-\infty, 0])$ if and only if there exists an absolutely monotonic entire function $\widetilde{F} : \mathbb{C} \to \mathbb{C}$ such that*

$$F(x) = \begin{cases} \widetilde{F}(x) & \text{if } x \in (0, \infty), \\ 0 & \text{if } x = 0, \\ -\widetilde{F}(-x) & \text{if } x \in (-\infty, 0). \end{cases}$$

It is striking to observe the possibility of a discontinuity at the origin which may occur in the latter two of these three theorems.

We will content ourselves here with sketching the proof of the second result. For the others, see [7], noting that the first of the results follows from Theorems 3.14 and 3.17 for $\rho = \infty$.

Proof of Theorem 3.22 Note that the moment matrix corresponding to an element of $\mathcal{M}([0,1])$ has a zero entry if and only if $\mu = a\delta_0$ for some $a \geq 0$. This and the Schur product theorem give one implication.

For the converse, suppose F preserves $\mathcal{M}([0,1])$. Fix finitely many scalars c_j, $t_j > 0$ and an integer $n \geq 0$, and set

$$p(t) = (1-t)^n \quad \text{and} \quad \mu = \sum_j e^{-t_j\alpha} c_j \delta_{e^{-t_j h}}, \tag{3.9}$$

where $\alpha > 0$ and $h > 0$. If $g(x) := \sum_j c_j e^{-t_j x}$ then the integration trick (3.5), but working on $[0,1]$, shows that the forward finite differences of $F \circ g$ alternate in sign:

$$\sum_{k=0}^{n} (-1)^k \binom{n}{k} F\Big(\sum_j c_j e^{-t_j(\alpha+kh)}\Big) \geq 0,$$

so $(-1)^n \Delta_h^n (F \circ g)(\alpha) \geq 0$. As this holds for all $\alpha, h > 0$ and all $n \geq 0$, it follows that $F \circ g : (0,\infty) \to (0,\infty)$ is completely monotonic. The weak density of measures of the form μ, together with Bernstein's Theorem (2.1), gives that $F \circ g$ is completely monotonic on $(0,\infty)$ for every completely monotonic function $g : (0,\infty) \to (0,\infty)$. Finally, a theorem of Lorch and Newman [61, Theorem 5] now gives that $F : (0,\infty) \to (0,\infty)$ is absolutely monotonic. □

3.6 Multivariable Positivity Preservers and Moment Families

We now turn to the multivariable case, and begin with two results of FitzGerald, Micchelli, and Pinkus [33]. We first introduce some notation and a piece of terminology.

Fix $I \subset \mathbb{C}$ and an integer $m \geq 1$, and let

$$A^k = (a_{ij}^k)_{i,j=1}^N \in I^{N \times N} \quad \text{for } k = 1, \ldots, m.$$

For any function $f : I^m \to \mathbb{C}$, we have the $N \times N$ matrix

$$f(A^1, \ldots, A^m) := \big(f(a_{ij}^1, \ldots, a_{ij}^m)\big)_{i,j=1}^N \in \mathbb{C}^{N \times N}.$$

We say that $f : \mathbb{R}^m \to \mathbb{R}$ is *real positivity preserving* if

$$f(A^1, \ldots, A^m) \in \mathcal{P}_N(\mathbb{R}) \text{ for all } A^1, \ldots, A^m \in \mathcal{P}_N(\mathbb{R}) \text{ and all } N \geq 1,$$

where, as above $\mathcal{P}_N(\mathbb{R})$ is the collection of $N \times N$ positive semidefinite matrices with real entries. Similarly, we say that $f : \mathbb{C}^m \to \mathbb{C}$ is positivity preserving if

$$f(A^1, \ldots, A^m) \in \mathcal{P}_N \qquad \text{for all } A^1, \ldots, A^m \in \mathcal{P}_N \text{ and all } N \geq 1,$$

where $\mathcal{P}_N$ is the collection of $N \times N$ positive semidefinite matrices with complex entries. Finally, recall that a function $f : \mathbb{R}^m \to \mathbb{R}$ is said to be *real entire* if there exists an entire function $F : \mathbb{C}^m \to \mathbb{C}$ such that $F|_{\mathbb{R}^m} = f$. We will also use the multi-index notation

$$\mathbf{x}^\alpha := x_1^{\alpha_1} \cdots x_m^{\alpha_m} \quad \text{if } \mathbf{x} = (x_1, \ldots, x_m) \text{ and } \alpha = (\alpha_1, \ldots, \alpha_m).$$

The following theorems are natural extensions of Schoenberg's theorem and Herz's theorem, respectively.

Theorem 3.24 ([33, Theorem 2.1]) *Let $f : \mathbb{R}^m \to \mathbb{R}$, where $m \geq 1$. Then f is real positivity preserving if and only f is real entire of the form*

$$f(\mathbf{x}) = \sum_{\alpha \in \mathbb{Z}_+^m} c_\alpha \mathbf{x}^\alpha \qquad (\mathbf{x} \in \mathbb{R}^m),$$

where $c_\alpha \geq 0$ for all $\alpha \in \mathbb{Z}_+^m$.

Theorem 3.25 ([33, Theorem 3.1]) *Let $f : \mathbb{C}^m \to \mathbb{C}$, where $m \geq 1$. Then f is positivity preserving if and only f is of the form*

$$f(\mathbf{z}) = \sum_{\alpha, \beta \in \mathbb{Z}_+^m} c_{\alpha\beta} \mathbf{z}^\alpha \overline{\mathbf{z}}^\beta \qquad (\mathbf{z} \in \mathbb{C}^m),$$

where $c_{\alpha\beta} \geq 0$ for all $\alpha, \beta \in \mathbb{Z}_+^m$ and the power series converges absolutely for all $\mathbf{z} \in \mathbb{C}$.

We now consider the notion of moment family for measures on $\mathbb{R}^d$. As above, a measure on $\mathbb{R}^d$ is said to be *admissible* if it is non-negative and has moments of all orders. Given such a measure μ, we define the *moment family*

$$s_\alpha(\mu) := \int \mathbf{x}^\alpha \, \mathrm{d}\mu(\mathbf{x}) \qquad \text{for all } \alpha \in \mathbb{Z}_+^m.$$

In line with the above, we let $\mathcal{M}(K)$ denote the set of all moment families of admissible measures supported on $K \subset \mathbb{R}^d$.

Note that a measure μ is supported in $[-1, 1]^d$ if and only if its moment family is uniformly bounded:

$$\sup\{|s_\alpha(\mu)| : \alpha \in \mathbb{Z}_+^m\} < \infty.$$

Theorem 3.26 ([7, Theorem 8.1]) *A function $F : \mathbb{R} \to \mathbb{R}$ maps $\mathcal{M}\big([-1,1]^d\big)$ to itself if and only if F is absolutely monotonic and entire.*

Proof Since $[-1,1]$ can be identified with $[-1,1] \times \{0\}^{d-1} \subset [-1,1]^d$, the forward implication follows from the one-dimensional result, Theorem 3.21.

For the converse, we use the fact [73] that a collection of real numbers $(s_\alpha)_{\alpha \in \mathbb{Z}_+^d}$ is an element of $\mathcal{M}\big([-1,1]^d\big)$ if and only if the weighted Hankel-type kernels on $\mathbb{Z}_+^d \times \mathbb{Z}_+^d$

$$(\alpha, \beta) \mapsto s_{\alpha+\beta} \quad \text{and} \quad (\alpha, \beta) \mapsto s_{\alpha+\beta} - s_{\alpha+\beta+2\mathbf{1}_j} \quad (1 \leq j \leq d)$$

are positive semidefinite, where

$$\mathbf{1}_j := (0, \ldots, 0, 1, 0, \ldots, 0) \in \mathbb{Z}_+^d$$

with 1 in the jth position. Now suppose F is absolutely monotonic and entire; given a family $(s_\alpha)_{\alpha \in \mathbb{Z}_+^d}$ subject to these positivity constraints, we have to verify that the family $(F(s_\alpha))_{\alpha \in \mathbb{Z}_+^d}$ satisfies them as well.

Theorem 3.14 gives that $(\alpha, \beta) \mapsto F(s_{\alpha+\beta})$ and $(\alpha, \beta) \mapsto F(s_{\alpha+\beta+2\mathbf{1}_j})$ are positive semidefinite, so we must show that

$$(\alpha, \beta) \mapsto F(s_{\alpha+\beta}) - F(s_{\alpha+\beta+2\mathbf{1}_j})$$

is positive semidefinite for $j = 1, \ldots, d$. As F is absolutely monotonic and entire, it suffices to show that

$$(\alpha, \beta) \mapsto (s_{\alpha+\beta})^{\circ n} - (s_{\alpha+\beta+2\mathbf{1}_j})^{\circ n}$$

is positive semidefinite for any $n \geq 0$, but this follows from the Schur product theorem: if $A \geq B \geq 0$, then

$$A^{\circ n} \geq A^{\circ(n-1)} \circ B \geq A^{\circ(n-2)} \circ B^{\circ 2} \geq \cdots \geq B^{\circ n}.$$

□

We next consider characterizations of real-valued multivariable functions which map tuples of moment sequences to moment sequences.

Let $K_1, \ldots, K_m \subset \mathbb{R}$. A function $F : \mathbb{R}^m \to \mathbb{R}$ acts on tuples of moment sequences of (admissible) measures $\mathcal{M}(K_1) \times \cdots \times \mathcal{M}(K_m)$ as follows:

$$F[\mathbf{s}(\mu_1), \ldots, \mathbf{s}(\mu_m)]_k := F\big(s_k(\mu_1), \ldots, s_k(\mu_m)\big) \quad \text{for all } k \geq 0. \tag{3.10}$$

Given $I \subset \mathbb{R}^m$, a function $F : I \to \mathbb{R}$ is *absolutely monotonic* if F is continuous on I, and for all interior points $\mathbf{x} \in I$ and $\alpha \in \mathbb{Z}_+^m$, the mixed partial derivative

$D^\alpha F(\mathbf{x})$ exists and is non-negative, where

$$D^\alpha F(\mathbf{x}) := \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \cdots \partial x_m^{\alpha_m}} F(x_1, \ldots, x_m) \quad \text{and } |\alpha| := \alpha_1 + \cdots + \alpha_m.$$

With this definition, the multivariable analogue of Bernstein's theorem is as one would expect; see [20, Theorem 4.2.2].

To proceed further, it is necessary to introduce the notion of a *facewise absolutely monotonic function* on $\mathbb{R}_+^m$. Observe that the orthant $\mathbb{R}_+^m$ is a convex polyhedron, and is therefore the disjoint union of the relative interiors of its faces. These faces are in one-to-one correspondence with subsets of $[m] := \{1, \ldots, m\}$:

$$J \mapsto \mathbb{R}_+^J := \{(x_1, \ldots, x_m) \in \mathbb{R}_+^m : x_i = 0 \text{ if } i \notin J\}; \tag{3.11}$$

note that this face has relative interior $\mathbb{R}_{>0}^J := (0, \infty)^J \times \{0\}^{[m] \setminus J}$.

Definition 3.27 A function $F : \mathbb{R}_+^m \to \mathbb{R}$ is *facewise absolutely monotonic* if, for every $J \subset [m]$, there exists an absolutely monotonic function g_J on $\mathbb{R}_+^J$ which agrees with F on $\mathbb{R}_{>0}^J$.

Thus a facewise absolutely monotonic function is piecewise absolutely monotonic, with the pieces being the relative interiors of the faces of the orthant $\mathbb{R}_+^m$. See [7, Example 8.4] for further discussion. In the special case $m = 1$, this broader class of functions (than absolutely monotonic functions on $\mathbb{R}_+$) coincides precisely with the maps which are absolutely monotonic on $(0, \infty)$ and have a possible discontinuity at the origin, as in Theorem 3.22 above.

This definition allows us to characterize the preservers of m-tuples of elements of $\mathcal{M}\big([0, 1]\big)$; the preceding observation shows that Theorem 3.22 is precisely the $m = 1$ case.

Theorem 3.28 ([7, Theorem 8.5]) *Let $F : \mathbb{R}_+^m \to \mathbb{R}$, where the integer $m \geq 1$. The following are equivalent.*

(1) *F maps $\mathcal{M}([0, 1])^m$ into $\mathcal{M}([0, 1])$.*
(2) *F is facewise absolutely monotonic, and the functions $\{g_J : J \subset [m]\}$ are such that $0 \leq g_J \leq g_K$ on $\mathbb{R}_+^J$ whenever $J \subset K \subset [m]$.*
(3) *F is such that*

$$F\big(\sqrt{x_1 y_1}, \ldots, \sqrt{x_m y_m}\big)^2 \leq F(x_1, \ldots, x_m) F(y_1, \ldots, y_m)$$

for all $\mathbf{x}$, $\mathbf{y} \in \mathbb{R}_+^m$ and there exists some $\mathbf{z} \in (0, 1)^m$ such that the products $\mathbf{z}^\alpha := z_1^{\alpha_1} \cdots z_m^{\alpha_m}$ are distinct for all $\alpha \in \mathbb{Z}_+^m$ and F maps $\mathcal{M}\big(\{1, z_1\}\big) \times \cdots \times \mathcal{M}\big(\{1, z_m\}\big) \cup \mathcal{M}(\{0, 1\})^m$ to $\mathcal{M}(\mathbb{R})$.

The heart of Theorem 3.28 can be deduced from the following result on positivity preservation on tuples of low-rank Hankel matrices. In a sense, it is the multi-dimensional generalization of the "stronger Vasudeva Theorem" 3.18.

Fix $\rho \in (0, \infty]$, an integer $m \geq 1$ and a point $\mathbf{z} \in (0, 1)^m$ with distinct products, as in Theorem 3.28(3). For all $N \geq 1$, let

$$\mathcal{H}_N := \{a\mathbf{1}_{N\times N} + b\mathbf{u}_{l,N}\mathbf{u}_{l,N}^T : a \in (0, \rho),\ b \in [0, \rho - a),\ 1 \leq l \leq m\},$$

where $\mathbf{u}_{l,N} := (1, z_l, \dots, z_l^{N-1})^T$.

Theorem 3.29 ([7, Theorem 8.6]) *If $F : (0, \rho)^m \to \mathbb{R}$ preserves positivity on $\mathcal{P}_2\big((0, \rho)\big)^m$ and $\mathcal{H}_N^m$ for all $N \geq 1$, then F is absolutely monotonic and is the restriction of an analytic function on the polydisc $D(0, \rho)^m$.*

The notion of facewise absolute monotonicity emerges from the study of positivity preservers of tuples of moment sequences. If one focuses instead on maps preserving positivity of tuples of all positive semidefinite matrices, or even all Hankel matrices, then this richer class of maps does not appear.

Proposition 3.30 *Suppose $\rho \in (0, \infty]$ and $F : [0, \rho)^m \to \mathbb{R}$. The following are equivalent.*

(1) *$F[-]$ preserves positivity on the space of m-tuples of Hankel matrices with entries in $[0, \rho)$.*
(2) *F is absolutely monotonic on $[0, \rho)^m$.*
(3) *$F[-]$ preserves positivity on the space of m-tuples of all matrices with entries in $[0, \rho)$.*

Proof Clearly (2) $\implies$ (3) $\implies$ (1), so suppose (1) holds. It follows from Theorem 3.29 that F is absolutely monotonic on the domain $(0, \rho)^m$ and agrees there with an analytic function $g : D(0, \rho)^m \to \mathbb{C}$. To see that $F \equiv g$ on $[0, \rho)^m$, we use induction on m, with the $m = 1$ case being left as an exercise (see [7, Proof of Proposition 7.3]).

Now suppose $m > 1$, let $\mathbf{c} = (c_1, \dots, c_m) \in [0, \rho)^m \setminus (0, \rho)^m$ and define

$$H := \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 2 \end{bmatrix} \qquad \text{and} \qquad A_i := \begin{cases} \mathbf{1}_{3\times 3} & \text{if } c_i > 0, \\ H & \text{if } c_i = 0. \end{cases}$$

Choosing $\mathbf{u}_n = (u_{1,n}, \dots, u_{m,n}) \in (0, \rho)^m$ such that $\mathbf{u}_n \to \mathbf{c}$, it follows that

$$\lim_{n\to\infty} F[u_{1,n}A_1, \dots, u_{m,n}A_m] = \begin{bmatrix} g(\mathbf{c}) & F(\mathbf{c}) & g(\mathbf{c}) \\ F(\mathbf{c}) & g(\mathbf{c}) & g(\mathbf{c}) \\ g(\mathbf{c}) & g(\mathbf{c}) & g(\mathbf{c}) \end{bmatrix} \in \mathcal{P}_3,$$

where the $(1, 2)$ and $(2, 1)$ entries are as claimed by the induction hypothesis. The determinants of the first and last principal minors now give that

$$g(\mathbf{c}) \geq 0 \qquad \text{and} \qquad -g(\mathbf{c})\big(g(\mathbf{c}) - F(\mathbf{c})\big)^2 \geq 0,$$

whence $F(\mathbf{c}) = g(\mathbf{c})$. □

Having considered functions defined on the positive orthant, we now look at the situation for functions defined over the whole of $\mathbb{R}^m$.

Theorem 3.31 ([7, Theorem 8.9]) *Suppose* $F : \mathbb{R}^m \to \mathbb{R}$ *for some integer* $m \geq 1$. *The following are equivalent.*

(1) F *maps* $\mathcal{M}\big([-1, 1]\big)^m$ *into* $\mathcal{M}(\mathbb{R})$.
(2) *The function* F *is real positivity preserving.*
(3) *The function* F *is absolutely monotonic on* $\mathbb{R}_+^m$ *and agrees with an entire function on* $\mathbb{R}^m$.

As before, the proof reveals that verifying positivity preservation for tuples of low-rank Hankel matrices suffices. The following notation and corollary make this precise.

For all $u \in (0, \infty)$, let $\mathcal{M}_u := \mathcal{M}\big(\{-1, u, 1\}\big)$ and

$$\mathcal{M}_{[u]} := \bigcup \big\{\mathcal{M}\big(\{s_1, s_2\}\big) : s_1 \in \{-1, 0, 1\},\ s_2 \in \{-u, 0, u\}\big\}.$$

Corollary 3.32 ([7, Theorem 8.10]) *The hypotheses in Theorem 3.31 are also equivalent to the following.*

(4) There exist $u_0 \in (0, 1)$ *and* $\epsilon > 0$ *such that* F *maps*

$$\mathcal{M}_{[u_0]}^m \cup \bigcup \big\{\mathcal{M}_{v_1} \times \cdots \times \mathcal{M}_{v_m} : v_1, \ldots, v_m \in (0, 1 + \epsilon)\big\}$$

into $\mathcal{M}(\mathbb{R})$.

4 Totally Non-negative Matrices and Positivity Preservers

In this chapter, we discuss variant notions of matrix positivity that are well studied in the literature, *total positivity* and *total non-negativity*, and characterize the maps which preserve these properties.

Definition 4.1 A real matrix A is said to be *totally non-negative* or *totally positive* if every minor of A is non-negative or positive, respectively. We will denote these matrices, as well as the property, by TN and TP.

In older texts, such matrices were called *totally positive* and *strictly totally positive*, respectively.

To introduce the theory of total positivity, we can do no better than quote from the preface of Karlin's magisterial book [54]: "Total positivity is a concept of considerable power that plays an important role in various domains of mathematics, statistics and mechanics." Karlin goes on to list "problems involving convexity, moment spaces, eigenvalues of integral operators, ... oscillation properties of solutions of linear differential equations ... the theory of approximations ... statistical decision procedures ... discerning uniformly most powerful tests for hypotheses ... ascertaining optimal policy for inventory and production processes ... analysis of diffusion-type stochastic processes, and ... coupled mechanical systems."

Perhaps the earliest result on total positivity is due to Fekete, in correspondence with Pólya [32] published in 1912. Schoenberg observed the variation-diminishing properties of TP matrices in 1930 [80], and published a series of papers on Pólya frequency functions, which are defined in terms of total positivity, in the 1950s [87–89]. Independently of Schoenberg, Krein's investigation of ordinary differential equations led him to the total positivity of Green's functions for certain differential operators, and in the mid-1930s his works with Gantmacher looked at spectral and other properties of totally positive matrices and kernels; see [36] and [54, Section 10.6].

For more on these four authors, one may consult the afterwork of Pinkus's book on total positivity [69], which also contains a wealth of results on totally positive and totally non-negative matrices. For a modern collection of applications of the theory of total positivity, see the book edited by Gasca and Micchelli [37].

More recently, total positivity has had a major impact on Lie theory. Lusztig extended the theory of total positivity to the setting of linear algebraic groups; see [62] for an exposition of this work. This led Fomin and Zelevinsky to investigate the combinatorics of Lusztig's theory [34] and resulted in the invention of cluster algebras [35]. These objects have generated an enormous amount of activity in a short period of time, with connections across a wide range of areas within representation theory, combinatorics, geometry, and mathematical physics. For the latter, we will mention only the totally non-negative Grassmannian [72], its connections with scattering amplitudes for quantum field theories [1], and the work by Kodama and Williams on regular soliton solutions of the Kadomtsev–Petviashvili equation [59].

Example 4.2 Perhaps the most well-known class of totally positive matrices consists of the *(generalized) Vandermonde matrices*: for real numbers $0 < x_1 < \cdots < x_m$ and $\alpha_1 < \cdots < \alpha_n$, the $m \times n$ matrix

$$A := [x_j^{\alpha_k}]_{1 \leq j \leq m,\ 1 \leq k \leq n}$$

is totally positive. Indeed, it suffices to show the positivity of any such matrix determinant $\det A$ when $m = n$. That $\det A$ is non-zero follows from Laguerre's extension of Descartes' rule of signs (see [51]) and by fixing the x_j and considering a linear homotopy from $(0, 1, \ldots, n-1)$ to $(\alpha_1, \ldots, \alpha_n)$, one obtains a continuous non-vanishing function from the usual Vandermonde determinant $\prod_{1 \leq j < k \leq n}(x_k - x_j)$ (which is positive) to $\det A$.

Example 4.3 Another prominent class of symmetric totally positive matrices consists of the Hankel moment matrices $H_\mu := [s_{j+k}(\mu)]_{j,k \geq 0}$ corresponding to admissible measures μ; see Definition 3.16.

4.1 Totally Non-negative and Totally Positive Kernels

An important generalization of TN and TP matrices is given by the following functional form.

Definition 4.4 Let X and Y be totally ordered sets, and let $K : X \times Y \to \mathbb{R}$ be a kernel.

(1) The kernel K is *totally positive of order* r, denoted TP_r, if, for any n-tuples of points $x_1 < \cdots < x_n$ in X and $y_1 < \cdots < y_n$ in Y, where $1 \leq n \leq r$, the matrix

$$[K(x_j, y_k)]_{j,k=1}^n$$

has positive determinant.

(2) The kernel K is *totally positive* if K is TP_r for all $r \geq 1$.

(3) Similarly, one defines TN_r kernels and totally non-negative kernels by replacing the word "positive" in the above by "non-negative."

If $X = \{1, \ldots, m\}$ and $Y = \{1, \ldots, n\}$, we recover the earlier notions of totally positive and totally non-negative matrices. When X and Y are taken to be real intervals, TN and TP kernels can be thought of as continuous analogues of TN and TP matrices. In fact, one has a continuous analogue of the Cauchy–Binet formula, which generalizes its traditional version.

Theorem 4.5 (Basic Composition Lemma, See, e.g., [54, 55]) *Suppose $X, Y, Z \subset \mathbb{R}$ and let μ be a non-negative Borel measure on Y. Suppose $K : X \times Y \to \mathbb{R}$ and $L : Y \times Z \to \mathbb{R}$ are pointwise Borel measurable with respect to Y, and let*

$$M : X \times Z \to \mathbb{R}; \ (x, z) \mapsto \int_Y K(x, y)L(y, z)\, \mathrm{d}\mu(y).$$

If M is well defined on the whole of $X \times Z$, then

$$\det \begin{bmatrix} M(x_1, z_1) & \ldots & M(x_1, z_m) \\ \vdots & \ddots & \vdots \\ M(x_m, z_1) & \ldots & M(x_m, z_m) \end{bmatrix}$$

$$= \int \cdots \int_{y_1 < y_2 < \cdots < y_m \in Y} \det[K(x_i, y_j)]_{i,j=1}^m \det[L(y_j, z_k)]_{j,k=1}^m \prod_{j=1}^m \mathrm{d}\mu(y_j).$$

As an immediate consequence, we have the following corollary.

Corollary 4.6 *In the setting of Theorem 4.5, if the kernels K and L are both TN_r or TP_r for some $r \geq 1$, then M has the same property. In particular, if K and L are both TN or TP, then so is M.*

We conclude this part with an observation of Pólya that connects to a class of well-studied functions, and also implies the positive definiteness of the Gaussian kernel. Recall from the proof of Theorem 2.4 above that this latter property was crucially used by Schoenberg in characterizing metric space embeddings into Hilbert space; however, its proof above was only outlined (via the more sophisticated machinery of Fourier analysis and Bochner's theorem).

Lemma 4.7 (Pólya) *The Gaussian kernel* $K : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ *given by* $K(x, y) := \exp(-(x - y)^2)$ *is totally positive.*

Proof It suffices to show that every square matrix generated from the kernel has positive determinant. Given real numbers $x_1 < \cdots < x_n$ and $y_1 < \cdots < y_n$, we observe the following factorization:

$$[\exp(-(x_j - y_k)^2)]_{j,k=1}^n = \operatorname{diag}[\exp(-x_j^2)]_{j=1}^n [\exp(2x_j y_k)]_{j,k=1}^n \operatorname{diag}[\exp(-y_k^2)]_{k=1}^n.$$

The proof concludes by observing that all three matrices on the right-hand side have positive determinants, the second because it is a Vandermonde matrix $[p_j^{\alpha_k}]$ with $p_j = \exp(2x_j)$ and $\alpha_k = y_k$. □

Example 4.8 The Gaussian function $f(x) = \exp(-x^2)$ is thus an example of a *Pólya frequency function*, that is, one for which $f(x-y)$ is a TP kernel on $\mathbb{R} \times \mathbb{R}$. As noted above, these functions were intensively studied by Schoenberg, and continue to be much studied in mathematics and statistics; two of the classic references are [22, 27].

The case of the multivariate Gaussian kernel follows immediately from the one-dimensional version.

Corollary 4.9 *For all* $d \geq 1$, *the Gaussian kernel*

$$\mathbb{R}^d \times \mathbb{R}^d \to (0, \infty); \;\; (\mathbf{x}, \mathbf{y}) \mapsto K(\mathbf{x}, \mathbf{y}) := \exp(-\|\mathbf{x} - \mathbf{y}\|^2)$$

is positive semidefinite on $\mathbb{R}^d \times \mathbb{R}^d$. *In other words, the matrix* $[\exp(-\|\mathbf{x}_j - \mathbf{x}_k\|^2)]_{j,k=1}^n$ *is positive semidefinite for all* $\mathbf{x}_1, \ldots, \mathbf{x}_n \in \mathbb{R}^d$.

Proof The $d = 1$ case is a direct consequence of Lemma 4.7, and the case of general d follows from this by using the Schur product theorem. □

4.2 Entrywise Preservers of Totally Non-negative Matrices

The TN property is very rigid when it comes to entrywise operations, as the following result makes clear.

Theorem 4.10 ([8, Theorem 2.1]) *Let $F : \mathbb{R}_+ \to \mathbb{R}$ be a function and let $d := \min(m, n)$, where m and n are positive integers. The following are equivalent.*

(1) *F preserves TN entrywise on $m \times n$ matrices.*
(2) *F preserves TN entrywise on $d \times d$ matrices.*
(3) *F is either a non-negative constant or*

 (a) $(d = 1)$ $F(x) \geq 0$;
 (b) $(d = 2)$ $F(x) = cx^\alpha$ *for some* $c > 0$ *and some* $\alpha \geq 0$;
 (c) $(d = 3)$ $F(x) = cx^\alpha$ *for some* $c > 0$ *and some* $\alpha \geq 1$;
 (d) $(d \geq 4)$ $F(x) = cx$ *for some* $c > 0$.

Proof That (1) $\iff$ (2) is immediate, as is the equivalence of (2) and (3) when $d = 1$. For larger values of d, we sketch the implication (2) $\implies$ (3).

For $d = 2$, let the totally non-negative matrices

$$A(x, y) := \begin{bmatrix} x & xy \\ 1 & y \end{bmatrix} \qquad \text{and} \qquad B(x, y) := \begin{bmatrix} xy & x \\ y & 1 \end{bmatrix} \qquad (x, y \geq 0). \tag{4.1}$$

If the non-constant function F preserves TN entrywise for 2×2 matrices, then the non-negativity of the determinants of $F[A(x, y)]$ and $F[B(x, y)]$ gives that

$$F(xy)F(1) = F(x)F(y) \qquad \text{for all } x, y \geq 0. \tag{4.2}$$

It follows that F is strictly positive. Applying Vasudeva's argument, as set out before Proposition 3.8, now implies that F is continuous on $(0, \infty)$. Since the identity (4.2) shows that $x \mapsto F(x)/F(1)$ is multiplicative, there exists an exponent $\alpha \in \mathbb{R}_+$ such that $F(x) = F(1)x^\alpha$ for all $x > 0$. The final details are left as an exercise.

For $d = 3$, note that the 3×3 matrix $A \oplus 0$ is totally non-negative if and only if the 2×2 matrix A is. Hence the previous working gives that $F(x) = cx^\alpha$ for some $c > 0$ and $\alpha \geq 0$. Looking at $\det F[C]$ for the totally non-negative matrix

$$C := \begin{bmatrix} 1 & 1/\sqrt{2} & 0 \\ 1/\sqrt{2} & 1 & 1/\sqrt{2} \\ 0 & 1/\sqrt{2} & 1 \end{bmatrix} \tag{4.3}$$

shows that we must have $\alpha \geq 1$.

The argument to rule out the possibility that $\alpha \in [1, 2)$ when $d \geq 4$ is more involved, but makes use of an example of Fallat, Johnson, and Sokal [31, Example 5.8]. Full details are provided in [8]. □

If our totally non-negative matrices are also required to be symmetric, and so positive semidefinite, then the classes of preservers are enlarged somewhat, but still fairly restrictive.

Theorem 4.11 ([8, Theorem 2.3]) *Let $F : \mathbb{R}_+ \to \mathbb{R}$ and let d be a positive integer. The following are equivalent.*

(1) *F preserves TN entrywise on symmetric $d \times d$ matrices.*
(2) *F is either a non-negative constant or*

  (a) *($d = 1$) $F \geq 0$;*
  (b) *($d = 2$) F is non-negative, non-decreasing, and multiplicatively mid-convex, that is, $F(\sqrt{xy})^2 \leq F(x)F(y)$ for all $x, y \in [0, \infty)$, so continuous;*
  (c) *($d = 3$) $F(x) = cx^\alpha$ for some $c > 0$ and some $\alpha \geq 1$;*
  (d) *($d = 4$) $F(x) = cx^\alpha$ for some $c > 0$ and some $\alpha \in \{1\} \cup [2, \infty)$;*
  (e) *($d \geq 5$) $F(x) = cx$ for some $c > 0$.*

4.3 Entrywise Preservers of Totally Positive Matrices

In moving from total non-negativity to total positivity, we face two significant technical challenges. Firstly, the idea of realizing totally non-negative $d \times d$ matrices as submatrices of totally non-negative $(d+1) \times (d+1)$ matrices, by padding with zeros, does not transfer to the TP setting. Secondly, it is no longer possible to use Vasudeva's idea to establish multiplicative mid-point convexity, since the test matrices used for this are not always totally positive.

The first issue leads us into the domain of *totally positive completion problems* [30]. It is possible to do this generality, using parametrizations of TP matrices [34] or exterior bordering [29, Chapter 9], but the following result has the advantage of providing an explicit embedding into a well-known class of matrices.

Lemma 4.12 ([8, Lemma 3.2]) *Any totally positive 2×2 matrix may be realized as the leading principal submatrix of a positive multiple of a rectangular totally positive generalized Vandermonde matrix of any larger size.*

Remark 4.13 ([8, Remark 3.4]) Lemma 4.12 can be strengthened to the following completion result: given integers $m, n \geq 2$, an arbitrary 2×2 matrix A occurs as a minor in a totally positive $m \times n$ matrix at any given position (that is, in a specified pair of rows and pair of columns) if and only if A is totally positive.

The other tool which will be vital to our deliberations is the following result of Whitney.

Theorem 4.14 ([97, Theorem 1]) *The set of totally positive $m \times n$ matrices is dense in the set of totally non-negative $m \times n$ matrices.*

With these tools in hand, we are able to provide a complete classification of the entrywise TP preservers of each fixed size, akin to the results in the preceding section.

Theorem 4.15 ([8, Theorem 3.1]) *Let $F : (0, \infty) \to \mathbb{R}$ be a function and let $d := \min(m, n)$, where m and n are positive integers. The following are equivalent.*

(1) *F preserves total positivity entrywise on $m \times n$ matrices.*
(2) *F preserves total positivity entrywise on $d \times d$ matrices.*
(3) *The function F satisfies*

   (a) $(d = 1)$ $F(x) > 0$;
   (b) $(d = 2)$ $F(x) = cx^\alpha$ *for some* $c > 0$ *and some* $\alpha > 0$;
   (c) $(d = 3)$ $F(x) = cx^\alpha$ *for some* $c > 0$ *and some* $\alpha \geq 1$;
   (d) $(d \geq 4)$ $F(x) = cx$ *for some* $c > 0$.

Proof We sketch the proof that (2) $\implies$ (3) when $d = 2$ and $d \geq 3$. For the first case, working with the matrix

$$\begin{bmatrix} y & x \\ x & y \end{bmatrix} \qquad (y > x > 0)$$

shows that F takes positive values and is increasing, so is Borel measurable and continuous except on a countable set. We now fix a point of continuity a and use the totally positive matrices

$$A(x, y, \epsilon) := \begin{bmatrix} ax & axy \\ a - \epsilon & ay \end{bmatrix} \quad \text{and} \quad B(x, y, \epsilon) := \begin{bmatrix} axy & ax \\ ay & a + \epsilon \end{bmatrix}$$

to show that

$$0 \leq \lim_{\epsilon \to 0^+} \det F[A(x, y, \epsilon)] = F(ax)F(ay) - F(axy)F(a)$$
$$\text{and} \quad 0 \leq \lim_{\epsilon \to 0^+} \det F[B(x, y, \epsilon)] = F(a)F(axy) - F(ax)F(ay)$$

for all $x, y > 0$. Hence $G : x \mapsto F(ax)/F(a)$ is such that

$$G(xy) = G(x)G(y) \qquad \text{for all } x, y > 0,$$

so G is a measurable solution of the Cauchy functional equation. It follows that $G(x) = x^\alpha$ for some $\alpha \in \mathbb{R}$. As F, and so G, is increasing, we must have $\alpha > 0$.

Finally, if $d \geq 3$, then the embedding of Lemma 4.12 and the previous working give positive constants c and α such that $F(x) = cx^\alpha$. In particular, the function F admits a continuous extension $\tilde{F}$ to $\mathbb{R}_+$. The density of TP in TN, that is, Theorem 4.14, implies that $\tilde{F}$ preserves TN entrywise on $d \times d$ matrices. Theorem 4.10 now establishes the form of $\tilde{F}$, and so of F. □

We may consider a version of the previous theorem which restricts to the case of totally positive matrices which are symmetric. A moment's thought leads to the consideration of a symmetric version of the matrix completion problem.

Lemma 4.16 ([8, Lemma 3.7]) *Any symmetric totally positive 2×2 matrix occurs as the leading principal submatrix of a totally positive $d \times d$ Hankel matrix, where $d \geq 2$ can be taken arbitrary large.*

Proof It suffices to embed the matrix

$$\begin{bmatrix} 1 & a \\ a & b \end{bmatrix} \qquad (0 < a < \sqrt{b})$$

into such a Hankel matrix. It is an exercise to prove the existence of a continuous function $f : [0, 1] \to \mathbb{R}_+;\ x \mapsto cx^s$ such that

$$\int_0^1 f(x)\,\mathrm{d}x = a \qquad \text{and} \qquad \int_0^1 f(x)^2\,\mathrm{d}x = b,$$

and then setting

$$a_{jk} := \int_0^1 f(x)^{j+k}\,\mathrm{d}x \qquad (j, k \geq 0)$$

gives a Hankel matrix A as required. The verification of total positivity may be made with the help of Andréief's identity,

$$\det\left[\int \phi_i(x)\psi_j(x)\,\mathrm{d}x\right]_{i,j=1}^k = \frac{1}{k!}\int\cdots\int \det(\phi_i(x_j))_{i,j=1}^k \det(\psi_i(x_j))_{i,j=1}^k \,\mathrm{d}x_1\cdots\mathrm{d}x_k,$$

where $\phi_i(x) = f(x)^{\alpha_i - 1}$ and $\psi_j(x) = f(x)^{\beta_j - 1}$, with

$$1 \leq \alpha_1 < \cdots < \alpha_k \leq d \qquad \text{and} \qquad 1 \leq \beta_1 < \cdots < \beta_k \leq d,$$

together with the total positivity of generalized Vandermonde matrices. □

We remark here that the preceding result can be further strengthened to have the symmetric TP 2×2 matrix occur in any "symmetric" position inside a larger square symmetric TP Hankel matrix, in the spirit of Remark 4.13. See [8, Theorem 3.9] for details.

We now state the symmetric version of Theorem 4.15.

Theorem 4.17 ([8, Theorem 3.6]) *Let $F : (0, \infty) \to \mathbb{R}$ and let d be a positive integer. The following are equivalent.*

(1) *F preserves total positivity entrywise on symmetric $d \times d$ matrices.*
(2) *The function F satisfies*

(a) $(d = 1)$ $F(x) > 0$*;*
(b) $(d = 2)$ F *is positive, increasing, and multiplicatively mid-convex, that is,* $F(\sqrt{xy})^2 \leq F(x)F(y)$ *for all* $x, y \in (0, \infty)$*, so continuous;*
(c) $(d = 3)$ $F(x) = cx^\alpha$ *for some* $c > 0$ *and some* $\alpha \geq 1$*;*
(d) $(d = 4)$ $F(x) = cx^\alpha$ *for some* $c > 0$ *and some* $\alpha \in \{1\} \cup [2, \infty)$*.*
(e) $(d \geq 5)$ $F(x) = cx$ *for some* $c > 0$*.*

Although we have developed the key ingredients to prove this theorem, we content ourselves with referring the interested reader to [8].

References

1. N. Arkani-Hamed, J.L. Bourjaily, F. Cachazo, A.B. Goncharov, A. Postnikov, J. Trnka. Scattering amplitudes and the positive Grassmannian (2012). Preprint. Available at http://arxiv.org/abs/1212.5605
2. V.S. Barbosa, V.A. Menegatto, Strictly positive definite kernels on compact two-point homogeneous spaces. Math. Inequal. Appl. **19**(2), 743–756 (2016)
3. V.S. Barbosa, V.A. Menegatto, Strict positive definiteness on products of compact two-point homogeneous spaces. Integr. Transf. Spec. Funct. **28**(1), 56–73 (2017)
4. R.K. Beatson, W. zu Castell, Dimension hopping and families of strictly positive definite zonal basis functions on spheres. J. Approx. Theory **221**(C), 22–37 (2017)
5. R.K. Beatson, W. zu Castell, Y. Xu, A Pólya criterion for (strict) positive-definiteness on the sphere. IMA J. Numer. Anal. **34**(2), 550–568 (2014)
6. A. Belton, D. Guillot, A. Khare, M. Putinar, Matrix positivity preservers in fixed dimension. I. Adv. Math. **298**, 325–368 (2016)
7. A. Belton, D. Guillot, A. Khare, M. Putinar, Moment-sequence transforms (2016). Preprint. Available at http://arxiv.org/abs/1610.05740
8. A. Belton, D. Guillot, A. Khare, M. Putinar, Total-positivity preservers (2017). Preprint. Available at http://arxiv.org/abs/1711.10468
9. A. Belton, D. Guillot, A. Khare, M. Putinar, A panorama of positivity (2018). Preprint. Available at http://arxiv.org/abs/1812.05482
10. A. Belton, D. Guillot, A. Khare, M. Putinar, A panorama of positivity. II: fixed dimension, in *Complex Analysis and Spectral Theory: Thomas Ransford Festschrift*, ed. by J. Mashreghi. CRM Proceedings and Lecture Notes Series (American Mathematical Society, Providence, to appear)
11. C. Berg, E. Porcu, From Schoenberg coefficients to Schoenberg functions. Constr. Approx. **45**(2), 217–241 (2017)
12. C. Berg, J.P.R. Christensen, P. Ressel, Positive definite functions on abelian semigroups. Math. Ann. **223**(3), 253–274 (1976)
13. C. Berg, A.P. Peron, E. Porcu, Schoenberg's theorem for real and complex Hilbert spheres revisited. J. Approx. Theory **228**, 58–78 (2018)
14. S. Bernstein, Sur les fonctions absolument monotones. Acta Math. **52**(1), 1–66 (1929)
15. G. Birkhoff, Metric foundations of geometry. I. Trans. Am. Math. Soc. **55**, 465–492 (1944)
16. L.M. Blumenthal, *Theory and Applications of Distance Geometry*, 2nd edn. (Chelsea Publishing Co., Bronx, 1970)
17. R.P. Boas, Jr., D.V. Widder, Functions with positive differences. Duke Math. J. **7**(1), 496–503 (1940)
18. S. Bochner, Monotone Funktionen, Stieltjessche Integrale und harmonische Analyse. Math. Ann. **108**(1), 378–410 (1933)
19. S. Bochner, Hilbert distances and positive definite functions. Ann. Math. **42**(3), 647–656 (1941)

20. S. Bochner, *Harmonic Analysis and the Theory of Probability* (University of California Press, Berkeley, 1955)
21. R.N. Bonfim, J.C. Guella, V.A. Menegatto, Strictly positive definite functions on compact two-point homogeneous spaces: the product alternative. SIGMA Symmetry Integrability Geom. Methods Appl. **14**, 112 (2018)
22. F. Brenti, Unimodal, log-concave and Pólya frequency sequences in combinatorics. Mem. Am. Math. Soc. **81**(413), viii+106 (1989)
23. D. Chen, V.A. Menegatto, X. Sun, A necessary and sufficient condition for strictly positive definite functions on spheres. Proc. Am. Math. Soc. **131**(9), 2733–2740 (2003)
24. J.P.R. Christensen, P. Ressel, Functions operating on positive definite matrices and a theorem of Schoenberg. Trans. Am. Math. Soc. **243**, 89–95 (1978)
25. H. Cohn, Y. Zhao, Sphere packing bounds via spherical codes. Duke Math. J. **163**(10), 1965–2002 (2014)
26. J. Dattorro, Equality relating Euclidean distance cone to positive semidefinite cone. Linear Algebra Appl. **428**(11–12), 2597–2600 (2008)
27. B. Efron, Increasing properties of Pólya frequency functions. Ann. Math. Stat. **36**(1), 272–279 (1965)
28. J. Emonds, H. Führ, Strictly positive definite functions on compact abelian groups. Proc. Am. Math. Soc. **139**(3), 1105–1113 (2011)
29. S.M. Fallat, C.R. Johnson, *Totally Nonnegative Matrices*. Princeton Series in Applied Mathematics (Princeton University Press, Princeton, 2011)
30. S.M. Fallat, C.R. Johnson, R.L. Smith, The general totally positive matrix completion problem with few unspecified entries. Electron. J. Linear Algebra **7**, 1–20 (2000)
31. S.M. Fallat, C.R. Johnson, A.D. Sokal, Total positivity of sums, Hadamard products and Hadamard powers: results and counterexamples. Linear Algebra Appl. **520**, 242–259 (2017)
32. M. Fekete, Über ein problem von Laguerre. Rend. Circ. Math. Palermo **34**, 89–120 (1912)
33. C.H. FitzGerald, C.A. Micchelli, A. Pinkus, Functions that preserve families of positive semidefinite matrices. Linear Algebra Appl. **221**, 83–102 (1995)
34. S. Fomin, A. Zelevinsky, Total positivity: tests and parametrizations. Math. Intell. **22**(1), 23–33 (2000)
35. S. Fomin, A. Zelevinsky, Cluster algebras. I. Foundations. J. Am. Math. Soc. **15**(2), 497–529 (2002)
36. F.R. Gantmacher, M.G. Krein, *Oscillation Matrices and Kernels and Small Vibrations of Mechanical Systems*, ed. by A. Eremenko, revised edn. (AMS Chelsea Publishing, New York, 2002)
37. M. Gasca, C.A. Micchelli (eds.), *Total Positivity and its Applications*. Mathematics and Its Applications, vol. 359 (Springer, Utrecht, 1996)
38. I.M. Gelfand, Normierte Ringe. Rec. Math. [Mat. Sbornik] N. S. **9**(51), 3–24 (1941)
39. T. Gneiting, Strictly and non-strictly positive definite functions on spheres. Bernoulli **19**(4), 1327–1349 (2013)
40. J.C. Guella, V.A. Menegatto, Strictly positive definite kernels on a product of spheres. J. Math. Anal. Appl. **435**(1), 286–301 (2016)
41. J.C. Guella, V.A. Menegatto, Strictly positive definite kernels on the torus. Constr. Approx. **46**(2), 271–284 (2017)
42. J.C. Guella, V.A. Menegatto, A.P. Peron, An extension of a theorem of Schoenberg to products of spheres. Banach J. Math. Anal. **10**(4), 671–685 (2016)
43. J.C. Guella, V.A. Menegatto, A.P. Peron, Strictly positive definite kernels on a product of spheres II. SIGMA Symmetry Integrability Geom. Methods Appl. **12**, 15 (2016)
44. J.C. Guella, V.A. Menegatto, A.P. Peron, Strictly positive definite kernels on a product of circles. Positivity **21**(1), 329–342 (2017)
45. D. Guillot, B. Rajaratnam, Retaining positive definiteness in thresholded matrices. Linear Algebra Appl. **436**(11), 4143–4160 (2012)
46. D. Guillot, B. Rajaratnam, Functions preserving positive definiteness for sparse matrices. Trans. Am. Math. Soc. **367**(1), 627–649 (2015)

47. D. Guillot, A. Khare, B. Rajaratnam, Preserving positivity for rank-constrained matrices. Trans. Am. Math. Soc. **369**(9), 6105–6145 (2017)
48. H. Helson, J.-P. Kahane, Y. Katznelson, W. Rudin, The functions which operate on Fourier transforms. Acta Math. **102**(1–2), 135–157 (1959)
49. C.S. Herz, Fonctions opérant sur les fonctions définies-positives. Ann. Inst. Fourier (Grenoble) **13**(1), 161–180 (1963)
50. R.A. Horn, The theory of infinitely divisible matrices and kernels. Trans. Am. Math. Soc. **136**, 269–286 (1969)
51. G.J.O. Jameson, Counting zeros of generalised polynomials: descartes' rule of signs and Laguerre's extensions. Math. Gaz. **90**(518), 223–234 (2006)
52. J.-P. Kahane, Sur un théorème de Wiener–Lévy. C. R. Acad. Sci. Paris **246**, 1949–1951 (1958)
53. J.-P. Kahane, W. Rudin, Caractérisation des fonctions qui opèrent sur les coefficients de Fourier-Stieltjes. C. R. Acad. Sci. Paris **247**, 773–775 (1958)
54. S. Karlin, *Total Positivity*, vol. I (Stanford University Press, Palo Alto, 1968)
55. S. Karlin, Y. Rinott, A generalized Cauchy–Binet formula and applications to total positivity and majorization. J. Multivar. Anal. **27**(1), 284–299 (1988)
56. Y. Katznelson, Sur les fonctions opérant sur l'algèbre des séries de Fourier absolument convergentes. C. R. Acad. Sci. Paris **247**, 404–406 (1958)
57. A. Khare, Smooth entrywise positivity preservers, a Horn–Loewner master theorem, and Schur polynomials (2018). Preprint. Available at http://arxiv.org/abs/1809.01823
58. A. Khare, T. Tao, On the sign patterns of entrywise positivity preservers in fixed dimension (2017). Preprint. Available at http://arxiv.org/abs/1708.05197
59. Y. Kodama, L. Williams, KP solitons and total positivity for the Grassmannian. Invent. Math. **198**(3), 637–699 (2014)
60. L. Liberti, C. Lavor, N. Maculan, A. Mucherino, Euclidean distance geometry and applications. SIAM Rev. **56**(1), 3–69 (2014)
61. L. Lorch, D.J. Newman, On the composition of completely monotonic functions and completely monotonic sequences and related questions. J. Lond. Math. Soc. **28**(1), 31–45 (1983)
62. G. Lusztig, Introduction to total positivity, in *Positivity in Lie Theory: Open Problems*, ed. by J. Hilgert, J.D. Lawson, K.-H. Neeb, E.B. Vinberg. De Gruyter Expositions in Mathematics, vol. 26 (Walter de Gruyter & Co., Berlin, 1998), pp. 133–145
63. K. Menger, Untersuchungen über allgemeine Metrik. Math. Ann. **100**(1), 75–163 (1928)
64. K. Menger, New foundation of euclidean geometry. Am. J. Math. **53**(4), 721–745 (1931)
65. J. Møller, M. Nielsen, E. Porcu, E. Rubak, Determinantal point process models on the sphere. Bernoulli **24**(2), 1171–1201 (2018)
66. O.R. Musin, The kissing number in four dimensions. Ann. Math. **168**(1), 1–32 (2008)
67. O.R. Musin, Multivariate positive definite functions on spheres, in *Discrete Geometry and Algebraic Combinatorics*. Contemporary Mathematics, vol. 625 (American Mathematical Society, Providence, 2014), pp. 177–190
68. A. Pinkus, Strictly positive definite functions on a real inner product space. Adv. Comput. Math. **20**(4), 263–271 (2004)
69. A. Pinkus, *Totally Positive Matrices*. Cambridge Tracts in Mathematics, vol. 181 (Cambridge University Press, Cambridge, 2010)
70. G. Pólya, G. Szegő, *Aufgaben und Lehrsätze aus der Analysis. Band II: Funktionentheorie, Nullstellen, Polynome Determinanten, Zahlentheorie* (Springer, Berlin, 1971)
71. E. Porcu, A. Alegria, R. Furrer, Modeling temporally evolving and spatially globally dependent data. Int. Stat. Rev. **86**(2), 344–377 (2018)
72. A. Postnikov, D. Speyer, L. Williams, Matching polytopes, toric geometry, and the totally non-negative Grassmannian. J. Algebraic Combin. **30**(2), 173–191 (2009)
73. M. Putinar, Positive polynomials on compact semi-algebraic sets. Indiana Univ. Math. J. **42**(3), 969–984 (1993)

74. A.W. Roberts, D.E. Varberg, *Convex Functions*. Pure and Applied Mathematics, vol. 57 (Academic, New York, 1973)
75. A.J. Rothman, E. Levina, J. Zhu, Generalized thresholding of large covariance matrices. J. Am. Stat. Assoc. **104**(485), 177–186 (2009)
76. W. Rudin, Transformations des coefficients de Fourier. C. R. Acad. Sci. Paris **243**, 638–640 (1956)
77. W. Rudin, Positive definite sequences and absolutely monotonic functions. Duke Math. J. **26**(4), 617–622 (1959)
78. W. Rudin, Some theorems on Fourier coefficients. Proc. Am. Math. Soc. **10**(6), 855–859 (1959)
79. W. Rudin, A strong converse of the Wiener-Levy theorem. Can. J. Math. **14**(4), 694–701 (1962)
80. I.J. Schoenberg, Über variationsvermindernde lineare Transformationen. Math. Z. **32**(1), 321–328 (1930)
81. I.J. Schoenberg, Remarks to Maurice Fréchet's article "Sur la définition axiomatique d'une classe d'espace distanciés vectoriellement applicable sur l'espace de Hilbert". Ann. Math. **36**(3), 724–732 (1935)
82. I.J. Schoenberg, On certain metric spaces arising from Euclidean spaces by a change of metric and their imbedding in Hilbert space. Ann. Math. **38**(4), 787–793 (1937)
83. I.J. Schoenberg, Metric spaces and completely monotone functions. Ann. Math. **39**(4), 811–841 (1938)
84. I.J. Schoenberg, Metric spaces and positive definite functions. Trans. Am. Math. Soc. **44**(3), 522–536 (1938)
85. I.J. Schoenberg, On metric arcs of vanishing Menger curvature. Ann. Math. **41**(4), 715–726 (1940)
86. I.J. Schoenberg, Positive definite functions on spheres. Duke Math. J. **9**(1), 96–108 (1942)
87. I.J. Schoenberg, On Pólya frequency functions. II. Variation-diminishing integral operators of the convolution type. Acta Sci. Math. Szeged **12**, 97–106 (1950)
88. I.J. Schoenberg, On Pólya frequency functions. I. The totally positive functions and their Laplace transforms. J. Anal. Math. **1**(1), 331–374 (1951)
89. I.J. Schoenberg, A.M. Whitney, On Pólya frequency functions. III. The positivity of translation determinants with an application to the interpolation problem by spline curves. Trans. Am. Math. Soc. **74**(2), 246–259 (1953)
90. I. Schur, Bemerkungen zur Theorie der beschränkten Bilinearformen mit unendlich vielen Veränderlichen. J. Reine Angew. Math. **140**, 1–28 (1911)
91. I. Steinwart, On the influence of the kernel on the consistency of support vector machines. J. Mach. Learn. Res. **2**(1), 67–93 (2002)
92. V.N. Vapnik, The nature of statistical learning theory, in *Statistics for Engineering and Information Science*, 2nd edn. (Springer, New York, 2000)
93. H.L. Vasudeva, Positive definite matrices and absolutely monotonic functions. Indian J. Pure Appl. Math. **10**(7), 854–858 (1979)
94. J. von Neumann, I.J. Schoenberg, Fourier integrals and metric geometry. Trans. Am. Math. Soc. **50**(2), 226–251 (1941)
95. H.-C. Wang, Two-point homogeneous spaces. Ann. Math. **55**(1), 177–191 (1952)
96. H. Weyl, Harmonics on homogeneous manifolds. Ann. Math. **35**(3), 486–499 (1934)
97. A.M. Whitney, A reduction theorem for totally positive matrices. J. Anal. Math. **2**(1), 88–92 (1952)
98. Y. Xu, Positive definite functions on the unit sphere and integrals of Jacobi polynomials. Proc. Am. Math. Soc. **146**(5), 2039–2048 (2018)
99. Y. Xu, E.W. Cheney, Strictly positive definite functions on spheres. Proc. Am. Math. Soc. **116**(4), 977–981 (1992)
100. J. Ziegel, Convolution roots and differentiability of isotropic positive definite functions on spheres. Proc. Am. Math. Soc. **142**(6), 2063–2077 (2014)

Table of Contents from Part II of the Survey