Total positivity: history, basics, and modern connections

Apoorva Khare

Indian Institute of Science, Bangalore
Analysis & Probability Research Group, Bangalore

PSD and TP matrices

Definition. A real symmetric matrix A is *positive definite* if any of the following equivalent conditions are satisfied.

1. The quadratic form of A is positive definite: $u^T A u > 0$ for all nonzero vectors $u \in \mathbb{R}^n$.
2. All principal minors of A are positive.
3. All eigenvalues of A are positive.

These matrices are ubiquitous in the applied sciences (correlation/covariance matrices), but also in analysis: Pick–Nevanlinna theory, the Hessian at local minima, . . .

PSD and TP matrices

Definition. A real symmetric matrix A is *positive definite* if any of the following equivalent conditions are satisfied.

1. The quadratic form of A is positive definite: $u^T A u > 0$ for all nonzero vectors $u \in \mathbb{R}^n$.
2. All principal minors of A are positive.
3. All eigenvalues of A are positive.

These matrices are ubiquitous in the applied sciences (correlation/covariance matrices), but also in analysis: Pick–Nevanlinna theory, the Hessian at local minima, ...

Definition. A rectangular matrix is *totally positive (TP)* if all minors are positive. (Similarly, totally non-negative (TN).)

Thus all entries > 0, all 2×2 minors > 0, ...

These matrices occur widely in mathematics: analysis, probability and statitics, differential equations, interpolation theory, matrix theory, representation theory, cluster algebras, integrable systems, combinatorics, ...

Totally positive matrices in mathematics

TP and TN matrices occur in

- analysis and differential equations (Aissen, Edrei, Schoenberg, Pólya, Loewner, Whitney)
- probability and statistics (Efron, Karlin, Pitman, Proschan, Rinott)
- interpolation theory and splines (Curry, Schoenberg)
- Gabor analysis (Gröchenig, Stöckler)
- interacting particle systems (Gantmacher, Krein)
- matrix theory (Ando, Cryer, Fallat, Garloff, Johnson, Pinkus, Sokal)

- representation theory (Lusztig, Postnikov)
- cluster algebras (Berenstein, Fomin, Zelevinsky)
- integrable systems (Kodama, Williams)
- quadratic algebras (Borger, Davydov, Grinberg, Hô Hai, Skryabin)
- combinatorics (Brenti, Lindström–Gessel–Viennot, Skandera, Sturmfels)

$\vdots$

Spectral properties of PSD and TP matrices: I

Theorem (folklore)

Suppose $A_{n \times n}$ is a real **symmetric** *matrix. The following are equivalent.*

1. *All principal minors of A are > 0 (or ≥ 0),*
   i.e. A is positive (semi)definite.
2. *All eigenvalues of (all principal submatrices of) A are > 0 (or ≥ 0).*

Theorem (folklore)

Suppose $A_{n\times n}$ is a real **symmetric** *matrix. The following are equivalent.*

1. *All principal minors of A are > 0 (or ≥ 0),*
   i.e. A is positive (semi)definite.
2. *All eigenvalues of (all principal submatrices of) A are > 0 (or ≥ 0).*

Theorem (Gantmacher–Krein, *Compos. Math.* 1937)

Suppose $A_{n\times n}$ is a real **square** *matrix. The following are equivalent.*

1. *All minors of A are > 0 (or ≥ 0),*
   i.e. A is totally positive (or totally non-negative).
2. *All eigenvalues of all square submatrices of A are > 0 (or ≥ 0).*

The proof uses Perron's theorem on positive matrices
+ Kronecker's theorem on compound matrices:

Proof for totally positive matrices: The r*th compound matrix* $C_r(A)$ is the $\binom{n}{r} \times \binom{n}{r}$ matrix, with $C_r(A)_{J,K} := \det A_{J \times K}$, where $|J| = |K| = r$.
(The subsets $J \subset [n]$ are lexicographically ordered.)

Properties of compound matrices:

- $C_0(A) := (1), \quad C_1(A) = A,$ and $C_n(A) = \det(A)$.
- If A is upper/lower triangular, diagonal, or symmetric, then so is $C_r(A)$.
- Cauchy–Binet formula: $C_r(AB) = C_r(A) C_r(B)$ for $A, B \in \mathbb{R}^{n \times n}$.

Spectral properties of PSD and TP matrices: II

Proof for totally positive matrices: The r*th compound matrix* $C_r(A)$ is the $\binom{n}{r} \times \binom{n}{r}$ matrix, with $C_r(A)_{J,K} := \det A_{J \times K}$, where $|J| = |K| = r$. (The subsets $J \subset [n]$ are lexicographically ordered.)

Properties of compound matrices:

- $C_0(A) := (1), \quad C_1(A) = A, \quad$ and $C_n(A) = \det(A)$.
- If A is upper/lower triangular, diagonal, or symmetric, then so is $C_r(A)$.
- Cauchy–Binet formula: $C_r(AB) = C_r(A)C_r(B)$ for $A, B \in \mathbb{R}^{n \times n}$.

Corollary: If $A = MJM^{-1}$ with J the Jordan canonical form, then $C_r(A) = C_r(M)C_r(J)C_r(M)^{-1}$, so the eigenvalues of $C_r(A)$ are those of $C_r(J)$, i.e. r-fold products of eigenvalues of A.

Spectral properties of PSD and TP matrices: II

Proof for totally positive matrices: The r*th compound matrix* $C_r(A)$ is the $\binom{n}{r} \times \binom{n}{r}$ matrix, with $C_r(A)_{J,K} := \det A_{J \times K}$, where $|J| = |K| = r$. (The subsets $J \subset [n]$ are lexicographically ordered.)

Properties of compound matrices:

- $C_0(A) := (1), \quad C_1(A) = A, \quad$ and $C_n(A) = \det(A)$.
- If A is upper/lower triangular, diagonal, or symmetric, then so is $C_r(A)$.
- Cauchy–Binet formula: $C_r(AB) = C_r(A)C_r(B)$ for $A, B \in \mathbb{R}^{n \times n}$.

Corollary: If $A = MJM^{-1}$ with J the Jordan canonical form, then $C_r(A) = C_r(M)C_r(J)C_r(M)^{-1}$, so the eigenvalues of $C_r(A)$ are those of $C_r(J)$, i.e. r-fold products of eigenvalues of A.

- Number the eigenvalues of $A_{n \times n}$ via: $|\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_n|$.
- Now if A is TP, every $C_r(A)$ has positive entries, so a Perron eigenvalue, which must be $\lambda_1 \cdots \lambda_r > 0$. These inequalities give: $\lambda_j \in (0, \infty)\ \forall j$.

Spectral properties of PSD and TP matrices: II

Proof for totally positive matrices: The r*th compound matrix* $C_r(A)$ is the $\binom{n}{r} \times \binom{n}{r}$ matrix, with $C_r(A)_{J,K} := \det A_{J \times K}$, where $|J| = |K| = r$.
(The subsets $J \subset [n]$ are lexicographically ordered.)

Properties of compound matrices:

- $C_0(A) := (1), \quad C_1(A) = A, \quad$ and $C_n(A) = \det(A)$.
- If A is upper/lower triangular, diagonal, or symmetric, then so is $C_r(A)$.
- Cauchy–Binet formula: $C_r(AB) = C_r(A)C_r(B)$ for $A, B \in \mathbb{R}^{n \times n}$.

Corollary: If $A = MJM^{-1}$ with J the Jordan canonical form, then $C_r(A) = C_r(M)C_r(J)C_r(M)^{-1}$, so the eigenvalues of $C_r(A)$ are those of $C_r(J)$, i.e. r-fold products of eigenvalues of A.

- Number the eigenvalues of $A_{n \times n}$ via: $|\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_n|$.
- Now if A is TP, every $C_r(A)$ has positive entries, so a Perron eigenvalue, which must be $\lambda_1 \cdots \lambda_r > 0$. These inequalities give: $\lambda_j \in (0, \infty)\ \forall j$.
- Since each Perron eigenvalue is simple, $\lambda_1 > \lambda_2 > \cdots > \lambda_n > 0$.

Examples of TP/TN matrices

1. The lower-triangular matrix $A = (\mathbf{1}_{j \geq k})_{j,k=1}^n$ is TN.
2. Generalized Vandermonde matrices are TP: if $0 < x_1 < \cdots < x_n$ and $y_1 < y_2 < \cdots < y_n$ are real, then

   $$\det(x_j^{y_k})_{j,k=1}^n > 0.$$

   This uses *Descartes' rule of signs* (1637).

Examples of TP/TN matrices

1. The lower-triangular matrix $A = (\mathbf{1}_{j \geq k})_{j,k=1}^n$ is TN.
2. Generalized Vandermonde matrices are TP: if $0 < x_1 < \cdots < x_n$ and $y_1 < y_2 < \cdots < y_n$ are real, then
$$\det(x_j^{y_k})_{j,k=1}^n > 0.$$
This uses *Descartes' rule of signs* (1637).
3. (Pólya:) The *Gaussian kernel* is TP: given $\sigma > 0$ and scalars
$$x_1 < x_2 < \cdots < x_n, \qquad y_1 < y_2 < \cdots < y_n,$$
the matrix $G[\mathbf{x}; \mathbf{y}] := (e^{-\sigma(x_j - y_k)^2})_{j,k=1}^n$ is TP.

Examples of TP/TN matrices

1. The lower-triangular matrix $A = (\mathbf{1}_{j \geq k})_{j,k=1}^n$ is TN.
2. Generalized Vandermonde matrices are TP: if $0 < x_1 < \cdots < x_n$ and $y_1 < y_2 < \cdots < y_n$ are real, then
$$\det(x_j^{y_k})_{j,k=1}^n > 0.$$
This uses *Descartes' rule of signs* (1637).
3. (Pólya:) The *Gaussian kernel* is TP: given $\sigma > 0$ and scalars
$$x_1 < x_2 < \cdots < x_n, \qquad y_1 < y_2 < \cdots < y_n,$$
the matrix $G[\mathbf{x}; \mathbf{y}] := (e^{-\sigma(x_j - y_k)^2})_{j,k=1}^n$ is TP.

   Proof: It suffices to show $\det G[\mathbf{x}; \mathbf{y}] > 0$. Now factorize:
$$G[\mathbf{x}; \mathbf{y}] = \operatorname{diag}(e^{-\sigma x_j^2})_{j=1}^n \cdot ((e^{2\sigma x_j})^{y_k})_{j,k=1}^n \cdot \operatorname{diag}(e^{-\sigma y_k^2})_{k=1}^n.$$
The middle matrix is a generalized Vandermonde matrix, so all three factors have positive determinants. □

Whitney's density theorem

Recall: every positive semidefinite matrix A can be approximated by a sequence of positive definite matrices: $A + \varepsilon \mathrm{Id}, \ \varepsilon \to 0^+$. (Consider the eigenvalues.)

In other words, positive definite matrices are dense in psd matrices.

Whitney's density theorem

Recall: every positive semidefinite matrix A can be approximated by a sequence of positive definite matrices: $A + \varepsilon \mathrm{Id}, \ \varepsilon \to 0^+$. (Consider the eigenvalues.)

In other words, positive definite matrices are dense in psd matrices.

A similar density was discovered by (Anne M.) Whitney:

Theorem (Whitney, *J. d'Analyse Math.* 1952)

The set of $m \times n$ TP_r matrices is dense in the set of $m \times n$ TN_r matrices.

Whitney's density theorem

Recall: every positive semidefinite matrix A can be approximated by a sequence of positive definite matrices: $A + \varepsilon \mathrm{Id}$, $\varepsilon \to 0^+$. (Consider the eigenvalues.)

In other words, positive definite matrices are dense in psd matrices.

A similar density was discovered by (Anne M.) Whitney:

Theorem (Whitney, *J. d'Analyse Math.* 1952)

The set of $m \times n$ TP_r matrices is dense in the set of $m \times n$ TN_r matrices.

Whitney's proof goes through verbatim for symmetric matrices.
It also goes through for 'continuous' versions, i.e. *kernels*.

Totally positive / non-negative kernels

Definition: Given totally ordered sets X, Y and a map ('kernel') $K : X \times Y \to \mathbb{R}$, we say that K is *totally positive of order* r *(TP_r)* if given any $p \leq r$ and arguments

$$x_1 < \cdots < x_p \text{ in } X, \qquad y_1 < \cdots < y_p \text{ in } Y,$$

the determinant $\det(K(x_j, y_k))_{j,k=1}^p > 0$.

Totally positive / non-negative kernels

Definition: Given totally ordered sets X, Y and a map ('kernel') $K : X \times Y \to \mathbb{R}$, we say that K is *totally positive of order r (TP_r)* if given any $p \leq r$ and arguments

$$x_1 < \cdots < x_p \text{ in } X, \qquad y_1 < \cdots < y_p \text{ in } Y,$$

the determinant $\det(K(x_j, y_k))_{j,k=1}^p > 0$.

- We say K is *totally positive* if K is TP_r for all r.
- Similarly for TN_r and TN kernels.

Examples:
(1) X, Y are finite $\rightsquigarrow$ TP_r and TN_r matrices.

Definition: Given totally ordered sets X, Y and a map ('kernel') $K : X \times Y \to \mathbb{R}$, we say that K is *totally positive of order* r *(TP_r)* if given any $p \leq r$ and arguments

$$x_1 < \cdots < x_p \text{ in } X, \qquad y_1 < \cdots < y_p \text{ in } Y,$$

the determinant $\det(K(x_j, y_k))_{j,k=1}^p > 0$.

- We say K is *totally positive* if K is TP_r for all r.
- Similarly for TN_r and TN kernels.

Examples:

(1) X, Y are finite $\rightsquigarrow$ TP_r and TN_r matrices.

(2) Say $X = Y = \mathbb{R}$. The kernel $K(x, y) := e^{xy}$ is TP. (Why?)

(3) Say $X = Y = \mathbb{R}$. The kernels $K_\pm(x, y) := \exp(\pm(x \pm y)^2)$ are TP.

Structured kernels: I. Hankel kernels

The kernel $K_+(x,y) := \exp((x+y)^2)$ is an example of a continuous *Hankel* kernel on $\mathbb{R} \times \mathbb{R}$,
i.e. there exists f such that $K(x,y) = f(x+y)$.

Structured kernels: I. Hankel kernels

The kernel $K_+(x,y) := \exp((x+y)^2)$ is an example of a continuous *Hankel* kernel on $\mathbb{R} \times \mathbb{R}$,
i.e. there exists f such that $K(x,y) = f(x+y)$.

Theorem (Widder, *Bull. Amer. Math. Soc.* 1934)

Suppose $X \subset \mathbb{R}$ is an open interval. Every continuous Hankel TN kernel on $X \times X$ is of the form

$$K(x,y) = \int_{\mathbb{R}} \exp(-(x+y)u) \; d\sigma(u), \qquad x, y \in X$$

where $\sigma : \mathbb{R} \to \mathbb{R}$ is a non-decreasing function.
Moreover, K is TP if and only if the measure $d\sigma$ has infinite support.

Structured kernels: I. Hankel kernels

The kernel $K_+(x,y) := \exp((x+y)^2)$ is an example of a continuous *Hankel* kernel on $\mathbb{R} \times \mathbb{R}$,
i.e. there exists f such that $K(x,y) = f(x+y)$.

Theorem (Widder, *Bull. Amer. Math. Soc.* 1934)

Suppose $X \subset \mathbb{R}$ is an open interval. Every continuous Hankel TN kernel on $X \times X$ is of the form

$$K(x,y) = \int_{\mathbb{R}} \exp(-(x+y)u)\ d\sigma(u), \qquad x, y \in X$$

where $\sigma : \mathbb{R} \to \mathbb{R}$ is a non-decreasing function.
Moreover, K is TP if and only if the measure $d\sigma$ has infinite support.

Thus, 'Whitney' density also holds for such kernels: $K(x,y)$ is the limit of

$$K_\varepsilon(x,y) := K(x,y) + \varepsilon \int_0^1 e^{-(x+y)u}\ du, \qquad \text{as} \quad \varepsilon \to 0^+.$$

Structured kernels: II. Toeplitz kernels

Similarly, the Gaussian kernel $K_-(x, y) := \exp(-(x-y)^2)$ is TP from above.
More generally, a *totally non-negative function* is $\Lambda : \mathbb{R} \to \mathbb{R}$ such that the Toeplitz kernel

$$T_\Lambda(x, y) := \Lambda(x - y), \qquad x, y \in \mathbb{R}$$

is totally non-negative.

'Representative' examples:

1. $\Lambda(x) = e^{-x^2}$.

Structured kernels: II. Toeplitz kernels

Similarly, the Gaussian kernel $K_-(x,y) := \exp(-(x-y)^2)$ is TP from above.
More generally, a *totally non-negative function* is $\Lambda : \mathbb{R} \to \mathbb{R}$ such that the Toeplitz kernel

$$T_\Lambda(x,y) := \Lambda(x-y), \qquad x,y \in \mathbb{R}$$

is totally non-negative.

'Representative' examples:

1. $\Lambda(x) = e^{-x^2}$.
2. $\Lambda(x) = c\mathbf{1}(x=a)$ for $a \in \mathbb{R}$ and $c \geq 0$. (Draw any submatrix from T_Λ; it is either a diagonal matrix or has a zero row/column.)

Structured kernels: II. Toeplitz kernels

Similarly, the Gaussian kernel $K_-(x,y) := \exp(-(x-y)^2)$ is TP from above. More generally, a *totally non-negative function* is $\Lambda : \mathbb{R} \to \mathbb{R}$ such that the Toeplitz kernel

$$T_\Lambda(x,y) := \Lambda(x-y), \qquad x,y \in \mathbb{R}$$

is totally non-negative.

'Representative' examples:

1. $\Lambda(x) = e^{-x^2}$.
2. $\Lambda(x) = c\mathbf{1}(x=a)$ for $a \in \mathbb{R}$ and $c \geq 0$. (Draw any submatrix from T_Λ; it is either a diagonal matrix or has a zero row/column.)
3. $\Lambda(x) = e^{ax+b}$ is TN. Indeed,

$$T_\Lambda((x_j, y_k)) = (e^{ax_j - ay_k + b})_{j,k \geq 1}$$

and this has rank-one, so all 'larger' minors vanish (hence are ≥ 0).

Structured kernels: II. Toeplitz kernels

Similarly, the Gaussian kernel $K_-(x,y) := \exp(-(x-y)^2)$ is TP from above. More generally, a *totally non-negative function* is $\Lambda : \mathbb{R} \to \mathbb{R}$ such that the Toeplitz kernel

$$T_\Lambda(x,y) := \Lambda(x-y), \qquad x, y \in \mathbb{R}$$

is totally non-negative.

'Representative' examples:

1. $\Lambda(x) = e^{-x^2}$.
2. $\Lambda(x) = c\mathbf{1}(x=a)$ for $a \in \mathbb{R}$ and $c \geq 0$. (Draw any submatrix from T_Λ; it is either a diagonal matrix or has a zero row/column.)
3. $\Lambda(x) = e^{ax+b}$ is TN. Indeed,

   $$T_\Lambda((x_j, y_k)) = (e^{ax_j - ay_k + b})_{j,k \geq 1}$$

   and this has rank-one, so all 'larger' minors vanish (hence are ≥ 0).
4. $\Lambda(x) = \mathbf{1}_{x \geq 0}$. (Can be verified to be TN by explicit computation.)

Note: the last two examples are not integrable functions.

Pólya frequency functions

A function $\Lambda : \mathbb{R} \to \mathbb{R}$ is a *Pólya frequency function* if (a) it is integrable, (b) it is nonzero at two points, and (c) the associated Toeplitz kernel T_Λ is TN.

Pólya frequency functions

A function $\Lambda : \mathbb{R} \to \mathbb{R}$ is a *Pólya frequency function* if (a) it is integrable, (b) it is nonzero at two points, and (c) the associated Toeplitz kernel T_Λ is TN.

Pólya frequency functions have a beautiful structure theory, developed by Schoenberg and others. They connect to real function theory, PDEs, Gabor analysis, . . .

Examples:

1. The Gaussian kernel e^{-x^2}.
2. While $\mathbf{1}_{x \geq 0}$ is not integrable, $e^{-x}\mathbf{1}_{x \geq 0}$ is a Pólya frequency function.

Pólya frequency functions

A function $\Lambda : \mathbb{R} \to \mathbb{R}$ is a *Pólya frequency function* if (a) it is integrable, (b) it is nonzero at two points, and (c) the associated Toeplitz kernel T_Λ is TN.

Pólya frequency functions have a beautiful structure theory, developed by Schoenberg and others. They connect to real function theory, PDEs, Gabor analysis, ...

Examples:

1. The Gaussian kernel e^{-x^2}.
2. While $\mathbf{1}_{x \geq 0}$ is not integrable, $e^{-x}\mathbf{1}_{x \geq 0}$ is a Pólya frequency function. This is because if $\Lambda(x)$ is a TN function, then so is $\Delta(x) := e^{ax+b}\Lambda(x)$, since

$$(T_\Delta(x_j, y_k))_{j,k=1}^n = \mathrm{diag}(e^{ax_j+b})_{j=1}^n (T_\Lambda(x_j, y_k))_{j,k=1}^n \mathrm{diag}(e^{-ay_k})_{k=1}^n,$$

and so the left-side has determinant ≥ 0.

Variation diminishing property

Totally non-negative matrices/kernels have the variation diminishing property:

Theorem (Schoenberg, *Math. Z.* 1930, *J. d'Analyse Math.* 1951)

1. *If a matrix* $A_{n \times m}$ *is TN, then* $S^-(A\mathbf{x}) \leq S^-(\mathbf{x})$ *for all* $\mathbf{x} \in \mathbb{R}^m$*.*
   ('Variation' denotes the number of sign changes in the values taken by the vector.)

Variation diminishing property

Totally non-negative matrices/kernels have the variation diminishing property:

Theorem (Schoenberg, *Math. Z.* 1930, *J. d'Analyse Math.* 1951)

1. *If a matrix $A_{n \times m}$ is TN, then $S^-(A\mathbf{x}) \leq S^-(\mathbf{x})$ for all $\mathbf{x} \in \mathbb{R}^m$. ('Variation' denotes the number of sign changes in the values taken by the vector.)*

2. (Continuous version: Now a Pólya frequency function is thought of as a *kernel*, to integrate against. And the 'variation' denotes the (possibly infinite) number of sign changes in the values of a function.)
   If $f : \mathbb{R} \to \mathbb{R}$ is integrable on all finite intervals, and Λ is a PF function,

$$S^-(g) \leq S^-(f), \qquad \textit{where} \qquad g(x) := \int_{\mathbb{R}} \Lambda(x-t) f(t) \ dt.$$

Variation diminishing property

Totally non-negative matrices/kernels have the variation diminishing property:

Theorem (Schoenberg, *Math. Z.* 1930, *J. d'Analyse Math.* 1951)

1. *If a matrix* $A_{n \times m}$ *is TN, then* $S^-(A\mathbf{x}) \leq S^-(\mathbf{x})$ *for all* $\mathbf{x} \in \mathbb{R}^m$*.*
('Variation' denotes the number of sign changes in the values taken by the vector.)

2. (Continuous version: Now a Pólya frequency function is thought of as a *kernel*, to integrate against. And the 'variation' denotes the (possibly infinite) number of sign changes in the values of a function.)
If $f : \mathbb{R} \to \mathbb{R}$ *is integrable on all finite intervals, and* Λ *is a PF function,*

$$S^-(g) \leq S^-(f), \qquad \textit{where} \qquad g(x) := \int_{\mathbb{R}} \Lambda(x-t) f(t) \; dt.$$

This has a 'discrete' variant for infinite sequences – discussed next.

Pólya frequency sequences

Discrete version: Pólya frequency *sequences*. These are sequences $\mathbf{a} = (a_n)_{n \in \mathbb{Z}}$, such that for any **integers**

$$l_1 < l_2 < \cdots < l_n, \qquad m_1 < m_2 < \cdots < m_n,$$

the determinant $\det(a_{l_j - m_k})_{j,k=1}^n \geq 0$.

Discrete version: Pólya frequency *sequences*. These are sequences $\mathbf{a} = (a_n)_{n \in \mathbb{Z}}$, such that for any **integers**

$$l_1 < l_2 < \cdots < l_n, \qquad m_1 < m_2 < \cdots < m_n,$$

the determinant $\det(a_{l_j - m_k})_{j,k=1}^n \geq 0$.

In other words, these are bi-infinite Toeplitz matrices

$$\begin{pmatrix} \ddots & \vdots & \vdots & \vdots & \vdots & \\ \cdots & a_0 & a_{-1} & a_{-2} & a_{-3} & \cdots \\ \cdots & a_1 & a_0 & a_{-1} & a_{-2} & \cdots \\ \cdots & a_2 & a_1 & a_0 & a_{-1} & \cdots \\ \cdots & a_3 & a_2 & a_1 & a_0 & \cdots \\ & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

which are totally non-negative.

History of total positivity / variation diminution: I

- The term ‘variation diminishing property’ was coined by Pólya, with ‘variation’ denoting the number of sign changes in a vector/function.

History of total positivity / variation diminution: I

- The term 'variation diminishing property' was coined by Pólya, with 'variation' denoting the number of sign changes in a vector/function.
- This was in use since Descartes' rule of signs. E.g. in 1883, Laguerre proved several variants of the rule of signs. A sample result is:

  Suppose $f(x)$ is a real polynomial. Given $\gamma \geq 0$, the number of variations of the power series $e^{\gamma x} f(x)$ is non-increasing in γ, and is always bounded below by the number of positive roots of f.

History of total positivity / variation diminution: I

- The term 'variation diminishing property' was coined by Pólya, with 'variation' denoting the number of sign changes in a vector/function.

- This was in use since Descartes' rule of signs. E.g. in 1883, Laguerre proved several variants of the rule of signs. A sample result is:
  Suppose $f(x)$ is a real polynomial. Given $\gamma \geq 0$, the number of variations of the power series $e^{\gamma x} f(x)$ is non-increasing in γ, and is always bounded below by the number of positive roots of f.

- Laguerre did not fully prove this; completed by Fekete (1912) – in correspondence with Pólya – as a corollary of the following result:
  Let $p(t) = \sum_{k \geq 0} c_k t^k$ be a formal power series with all $c_k \in [0, \infty)$. For the standard monomial basis x^k, the operator of multiplication by $p(t)$ is:

$$T_{\mathbf{c}} := \begin{pmatrix} c_0 & 0 & 0 & \cdots \\ c_1 & c_0 & 0 & \cdots \\ c_2 & c_1 & c_0 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

  If $T_{\mathbf{c}}$ is TN, $S^-(T_{\mathbf{c}}\mathbf{x}) \leq S^-(\mathbf{x})$ for all 'finite' $\mathbf{x}$. ($T_{\mathbf{c}}\mathbf{x}$ can be infinite!)

History of total positivity / variation diminution: II

Thus we go from Descartes (1637) to Laguerre (1883) to Fekete–Pólya (1912).

- Next came Schoenberg (1930), who showed that if $A_{n \times m}$ is TN then it has the variation diminishing property: $S^-(A\mathbf{x}) \leq S^-(\mathbf{x})$.
- Clearly, so does $-A$. Characterize the matrices having this property?

History of total positivity / variation diminution: II

Thus we go from Descartes (1637) to Laguerre (1883) to Fekete–Pólya (1912).

- Next came Schoenberg (1930), who showed that if $A_{n \times m}$ is TN then it has the variation diminishing property: $S^-(A\mathbf{x}) \leq S^-(\mathbf{x})$.
- Clearly, so does $-A$. Characterize the matrices having this property?
- Achieved by Motzkin in PhD thesis (1936) $\rightsquigarrow$ sign-regular matrices.

Remarkably, Motzkin's thesis also contained:

2. Motzkin transposition theorem
3. Fourier–Motzkin Elimination (FME) algorithm
4. A convex polyhedral set is the Minkowski sum of a compact (convex) polytope and a polyhedral cone.

History of total positivity / variation diminution: II

Thus we go from Descartes (1637) to Laguerre (1883) to Fekete–Pólya (1912).

- Next came Schoenberg (1930), who showed that if $A_{n \times m}$ is TN then it has the variation diminishing property: $S^-(A\mathbf{x}) \leq S^-(\mathbf{x})$.
- Clearly, so does $-A$. Characterize the matrices having this property?
- Achieved by Motzkin in PhD thesis (1936) $\rightsquigarrow$ sign-regular matrices.

Remarkably, Motzkin's thesis also contained:

2. Motzkin transposition theorem
3. Fourier–Motzkin Elimination (FME) algorithm
4. A convex polyhedral set is the Minkowski sum of a compact (convex) polytope and a polyhedral cone.

5. Beyond his PhD: studied what is now called the 'Motzkin number'.
6. First example of PID that is not a Euclidean domain: $\mathbb{Z}[(1+\sqrt{-19})/2]$.
7. Provided the first explicit polynomial that is positive on $\mathbb{R}^2$, yet not a sum-of-squares (Hilbert's 17th problem): $x^4y^2 + x^2y^4 - 3x^2y^2 + 1$.

Sign non-reversal property

In fact, TN and TP matrices are *characterized* by two properties together:

- Variation diminishing property: $S^-(A\mathbf{x}) \leq S^-(\mathbf{x})$ for all $\mathbf{x} \in \mathbb{R}^m$;
- *Sign non-reversal property:* If $\mathbf{x} \neq 0, \exists j$ such that $x_j \neq 0, \; x_j(A\mathbf{x})_j > 0$.

Sign non-reversal property

In fact, TN and TP matrices are *characterized* by two properties together:

- Variation diminishing property: $S^-(A\mathbf{x}) \leq S^-(\mathbf{x})$ for all $\mathbf{x} \in \mathbb{R}^m$;
- *Sign non-reversal property:* If $\mathbf{x} \neq 0, \exists j$ such that $x_j \neq 0,\ x_j(A\mathbf{x})_j > 0$.

It turns out that the latter property alone is sufficient!

Theorem (Choudhury–Kannan–K., *Bull. London Math. Soc.*, in press)

Given $m, n \geq k \geq 1$, and $A \in \mathbb{R}^{n \times m}$, the following are equivalent:

1. *The matrix A is TP_k.*
2. *Every square submatrix of size $r \leq k$ has the sign non-reversal property.*

Sign non-reversal property

In fact, TN and TP matrices are *characterized* by two properties together:

- Variation diminishing property: $S^-(A\mathbf{x}) \leq S^-(\mathbf{x})$ for all $\mathbf{x} \in \mathbb{R}^m$;
- *Sign non-reversal property:* If $\mathbf{x} \neq 0, \exists j$ such that $x_j \neq 0,\ x_j(A\mathbf{x})_j > 0$.

It turns out that the latter property alone is sufficient!

Theorem (Choudhury–Kannan–K., *Bull. London Math. Soc.*, in press)

Given $m, n \geq k \geq 1$, and $A \in \mathbb{R}^{n \times m}$, the following are equivalent:

1. *The matrix A is TP_k.*
2. *Every square submatrix of size $r \leq k$ has the sign non-reversal property.*

Theorem (Choudhury–Kannan–K., *Bull. London Math. Soc.*, in press)

Given $m, n \geq k \geq 1$, and $A \in \mathbb{R}^{n \times m}$, the following are equivalent:

1. *The matrix A is TN_k.*
2. *Every square submatrix of size $r \leq k$ has the non-strict sign non-reversal property – for a **single** vector in $\mathbb{R}^r$.*

Convolution

How to construct new examples of TP/TN kernels from old ones?

Convolution

How to construct new examples of TP/TN kernels from old ones?

First consider the matrix/'discrete' case: given two matrices $A_{m\times n}$ and $B_{n\times p}$ which are both TN_r, *their product is also* TN_r.

Proof: Given sets I, J of rows/columns of size $p \leq r$, use the Cauchy–Binet identity:

$$\det(AB)_{I\times J} = \det(A_{I\times[n]}B_{[n]\times J}) = \sum_{K\subset[n],\ |K|=p} \det A_{I\times K} \det B_{K\times J} \geq 0.$$

Convolution

How to construct new examples of TP/TN kernels from old ones?

First consider the matrix/'discrete' case: given two matrices $A_{m \times n}$ and $B_{n \times p}$ which are both TN_r, *their product is also* TN_r.

Proof: Given sets I, J of rows/columns of size $p \leq r$, use the Cauchy–Binet identity:

$$\det(AB)_{I \times J} = \det(A_{I \times [n]} B_{[n] \times J}) = \sum_{K \subset [n],\ |K| = p} \det A_{I \times K} \det B_{K \times J} \geq 0.$$

The Cauchy–Binet formula has a *continuous* version $\rightsquigarrow$ *Basic composition formula* (Pólya–Szegő). This implies:

Corollary: If $\Lambda_1, \Lambda_2 : \mathbb{R} \to [0, \infty)$ are integrable TN_r functions, then so is their *convolution*

$$(\Lambda_1 * \Lambda_2)(x) := \int_{\mathbb{R}} \Lambda_1(t) \Lambda_2(x - t)\ dt, \qquad x \in \mathbb{R}.$$

This will help construct additional examples of Pólya frequency functions.

More examples:

1. If $\Lambda(x)$ is a PF function, then so is $\Lambda(ax+b)$ for $a \neq 0$.
2. If $\Lambda_n(x)$ is a PF function, and $\Lambda_n \to \Lambda$ pointwise, with Λ integrable and nonzero on two points $\leadsto$ Λ is a PF function.

Pólya frequency functions and Laplace transforms

More examples:

1. If $\Lambda(x)$ is a PF function, then so is $\Lambda(ax+b)$ for $a \neq 0$.
2. If $\Lambda_n(x)$ is a PF function, and $\Lambda_n \to \Lambda$ pointwise, with Λ integrable and nonzero on two points $\leadsto \Lambda$ is a PF function.

The bilateral Laplace transform of a PF function Λ is

$$\mathcal{B}(\Lambda)(s) := \int_{\mathbb{R}} e^{-sx} \Lambda(x) \, dx, \qquad s \in \mathbb{C}.$$

Fact: $\mathcal{B}$ is an algebra map: $\mathcal{B}(\Lambda_1 * \Lambda_2) = \mathcal{B}(\Lambda_1)\mathcal{B}(\Lambda_2)$.

More examples:

1. If $\Lambda(x)$ is a PF function, then so is $\Lambda(ax+b)$ for $a \neq 0$.
2. If $\Lambda_n(x)$ is a PF function, and $\Lambda_n \to \Lambda$ pointwise, with Λ integrable and nonzero on two points $\rightsquigarrow$ Λ is a PF function.

The bilateral Laplace transform of a PF function Λ is

$$\mathcal{B}(\Lambda)(s) := \int_{\mathbb{R}} e^{-sx} \Lambda(x)\, dx, \qquad s \in \mathbb{C}.$$

Fact: $\mathcal{B}$ is an algebra map: $\mathcal{B}(\Lambda_1 * \Lambda_2) = \mathcal{B}(\Lambda_1)\mathcal{B}(\Lambda_2)$.

Now consider *one-sided* PF functions: $\varphi_a(x) := \frac{1}{a} e^{-x/a} \mathbf{1}_{x \geq 0}$ $\rightsquigarrow$ Laplace transform $\mathcal{B}(\varphi_a)(s) = 1/(1+as)$.

- Let $a_j > 0$ with $\sum_{j=1}^{\infty} a_j < \infty$. Then for each n, the convolution $\varphi_{a_1} * \cdots * \varphi_{a_n}$ is a one-sided PF function, with Laplace transform

$$\mathcal{B}(\varphi_{a_1} * \cdots * \varphi_{a_n})(s) = \frac{1}{\prod_{j=1}^{n} (1 + a_j s)}.$$

- Shifting the origin of $\varphi_{a_1} * \cdots * \varphi_{a_n}$ to $\delta \geq 0$ yields a one-sided PF function with Laplace transform $e^{-\delta s} / \prod_{j=1}^n (1 + a_j s)$.

Laguerre–Pólya class and Schoenberg's results: I. One-sided

- Shifting the origin of $\varphi_{a_1} * \cdots * \varphi_{a_n}$ to $\delta \geq 0$ yields a one-sided PF function with Laplace transform $e^{-\delta s} / \prod_{j=1}^{n} (1 + a_j s)$.

- Taking limits of PF functions gives a PF function $\rightsquigarrow$ a PF function with Laplace transform $e^{-\delta s} / \prod_{j=1}^{\infty} (1 + a_j s)$.

- Its reciprocal is the analytic (entire) function $e^{\delta s} \prod_{j=1}^{\infty} (1 + a_j s)$.

- Shifting the origin of $\varphi_{a_1} * \cdots * \varphi_{a_n}$ to $\delta \geq 0$ yields a one-sided PF function with Laplace transform $e^{-\delta s} / \prod_{j=1}^{n} (1 + a_j s)$.
- Taking limits of PF functions gives a PF function $\rightsquigarrow$ a PF function with Laplace transform $e^{-\delta s} / \prod_{j=1}^{\infty} (1 + a_j s)$.
- Its reciprocal is the analytic (entire) function $e^{\delta s} \prod_{j=1}^{\infty} (1 + a_j s)$.

Remarkably, *every* one-sided PF function shares this property:

Theorem (Schoenberg, *J. d'Analyse Math.* 1951)

A function $\Lambda : \mathbb{R} \to \mathbb{R}$, continuous on $(0, \infty)$ and with $\int_{\mathbb{R}} \Lambda(x)\, dx = 1$, is a one-sided PF function vanishing on $(-\infty, 0)$, if and only if

$$\frac{1}{\mathcal{B}(\Lambda)(s)} = e^{\delta s} \prod_{j=1}^{\infty} (1 + a_j s), \quad \textit{where} \quad \delta, a_j \geq 0, \ \sum_j a_j < \infty.$$

- Shifting the origin of $\varphi_{a_1} * \cdots * \varphi_{a_n}$ to $\delta \geq 0$ yields a one-sided PF function with Laplace transform $e^{-\delta s} / \prod_{j=1}^n (1 + a_j s)$.
- Taking limits of PF functions gives a PF function $\rightsquigarrow$ a PF function with Laplace transform $e^{-\delta s} / \prod_{j=1}^\infty (1 + a_j s)$.
- Its reciprocal is the analytic (entire) function $e^{\delta s} \prod_{j=1}^\infty (1 + a_j s)$.

Remarkably, *every* one-sided PF function shares this property:

Theorem (Schoenberg, *J. d'Analyse Math.* 1951)

A function $\Lambda : \mathbb{R} \to \mathbb{R}$, continuous on $(0, \infty)$ and with $\int_{\mathbb{R}} \Lambda(x)\, dx = 1$, is a one-sided PF function vanishing on $(-\infty, 0)$, if and only if

$$\frac{1}{\mathcal{B}(\Lambda)(s)} = e^{\delta s} \prod_{j=1}^{\infty} (1 + a_j s), \quad \textit{where} \quad \delta, a_j \geq 0, \ \sum_j a_j < \infty.$$

This is the limit of the polynomials $(1 + \dfrac{\delta s}{n})^n \displaystyle\prod_{j=1}^{n} (1 + a_j s)$, with negative roots.

Laguerre–Pólya class and Schoenberg's results: II. Two-sided

Similarly, using the Gaussian kernel and 'oppositely directed' variants of $e^{-x}\mathbf{1}_{x\geq 0}$, Schoenberg proved:

Theorem (Schoenberg, *J. d'Analyse Math.* 1951)

A function $\Lambda : \mathbb{R} \to \mathbb{R}$ *with* $\int_{\mathbb{R}} \Lambda(x)\ dx = 1$ *is a PF function, if and only if*

$$\frac{1}{\mathcal{B}(\Lambda)(s)} = e^{-\gamma s^2 + \delta s} \prod_{j=1}^{\infty} (1 + a_j s) e^{-a_j s},$$

where $\gamma \geq 0$ *and* $\delta, a_j \in \mathbb{R}$ *are such that* $0 < \gamma + \sum_j a_j^2 < \infty$.

Similarly, using the Gaussian kernel and 'oppositely directed' variants of $e^{-x}\mathbf{1}_{x\geq 0}$, Schoenberg proved:

Theorem (Schoenberg, *J. d'Analyse Math.* 1951)

A function $\Lambda : \mathbb{R} \to \mathbb{R}$ *with* $\int_{\mathbb{R}} \Lambda(x)\ dx = 1$ *is a PF function, if and only if*

$$\frac{1}{\mathcal{B}(\Lambda)(s)} = e^{-\gamma s^2 + \delta s} \prod_{j=1}^{\infty} (1 + a_j s) e^{-a_j s},$$

where $\gamma \geq 0$ *and* $\delta, a_j \in \mathbb{R}$ *are such that* $0 < \gamma + \sum_j a_j^2 < \infty$.

These two classes of entire functions were very well-studied by Laguerre, Pólya, and Schur in the early 20th century:

1. The first class of entire functions are limits – uniform on compact sets – of real polynomials with real non-positive roots. ('One-sided')

Laguerre–Pólya class and Schoenberg's results: II. Two-sided

Similarly, using the Gaussian kernel and 'oppositely directed' variants of $e^{-x}\mathbf{1}_{x\geq 0}$, Schoenberg proved:

Theorem (Schoenberg, *J. d'Analyse Math.* 1951)

A function $\Lambda : \mathbb{R} \to \mathbb{R}$ with $\int_{\mathbb{R}} \Lambda(x)\ dx = 1$ is a PF function, if and only if

$$\frac{1}{\mathcal{B}(\Lambda)(s)} = e^{-\gamma s^2 + \delta s} \prod_{j=1}^{\infty} (1 + a_j s) e^{-a_j s},$$

where $\gamma \geq 0$ and $\delta, a_j \in \mathbb{R}$ are such that $0 < \gamma + \sum_j a_j^2 < \infty$.

These two classes of entire functions were very well-studied by Laguerre, Pólya, and Schur in the early 20th century:

1. The first class of entire functions are limits – uniform on compact sets – of real polynomials with real non-positive roots. ('One-sided')
2. The second class $\rightsquigarrow$ limits of real polynomials with real roots.

$\rightsquigarrow$ *Laguerre–Pólya functions* (allowing for a factor of cs^m, $c \geq 0, m \in \mathbb{Z}^{\geq 0}$).

From the Laguerre–Pólya class to the Riemann Hypothesis

Pólya (1927) initiated the study of functions $\Lambda(t)$ such that $\mathcal{B}(\Lambda)(s)$ has only real zeros. His work contains the following result:

Theorem (Pólya, 1927)

The following statements are equivalent:

1. *The Riemann Xi-function* $\Xi(s) = \xi(1/2 + iz)$ *is in the Laguerre–Pólya class, where* $\xi(s) := \binom{s}{2}\pi^{-s/2}\Gamma(s/2)\zeta(s)$.
2. *The Riemann Hypothesis is true.*

Pólya (1927) initiated the study of functions $\Lambda(t)$ such that $\mathcal{B}(\Lambda)(s)$ has only real zeros. His work contains the following result:

Theorem (Pólya, 1927)

The following statements are equivalent:

1. *The Riemann Xi-function* $\Xi(s) = \xi(1/2 + iz)$ *is in the Laguerre–Pólya class, where* $\xi(s) := \binom{s}{2}\pi^{-s/2}\Gamma(s/2)\zeta(s)$.
2. *The Riemann Hypothesis is true.*

Combined with Schoenberg's result above, this yields:

Theorem (Gröchenig, 2020)

Let $\xi(s) = \binom{s}{2}\pi^{-s/2}\Gamma(s/2)\zeta(s)$ *be the Riemann xi-function. If*

$$\Lambda(x) := \int_{\mathbb{R}} \xi(u + 1/2)^{-1} e^{-ixu}\, du$$

is a Pólya frequency function, then the Riemann Hypothesis is true.

The Laguerre–Pólya class is thus a distinguished one in several areas.

From Pólya–Schur multipliers to Laguerre–Pólya

The Laguerre–Pólya class is also related to *Pólya–Schur multipliers*. Given a *multiplier sequence*

$$\Gamma = (\gamma_0, \gamma_1, \dots) \in \mathbb{R}^{\mathbb{N}}, \qquad \gamma_n \geq 0,$$

consider linear transformations of polynomials:

$$P(t) := p_0 + p_1 t + \cdots + p_n t^n \quad \mapsto \quad \Gamma[P(t)] := \gamma_0 p_0 + \gamma_1 p_1 t + \cdots + \gamma_n p_n t^n.$$

From Pólya–Schur multipliers to Laguerre–Pólya

The Laguerre–Pólya class is also related to *Pólya–Schur multipliers*. Given a *multiplier sequence*

$$\Gamma = (\gamma_0, \gamma_1, \dots) \in \mathbb{R}^{\mathbb{N}}, \qquad \gamma_n \geq 0,$$

consider linear transformations of polynomials:

$$P(t) := p_0 + p_1 t + \cdots + p_n t^n \quad \mapsto \quad \Gamma[P(t)] := \gamma_0 p_0 + \gamma_1 p_1 t + \cdots + \gamma_n p_n t^n.$$

We say that Γ is a multiplier sequence of the *first kind* (or *second kind*) if for all polynomials with real (or negative) roots, $\Gamma[P]$ has all real roots.

From Pólya–Schur multipliers to Laguerre–Pólya

The Laguerre–Pólya class is also related to *Pólya–Schur multipliers*. Given a *multiplier sequence*

$$\Gamma = (\gamma_0, \gamma_1, \dots) \in \mathbb{R}^{\mathbb{N}}, \qquad \gamma_n \geq 0,$$

consider linear transformations of polynomials:

$$P(t) := p_0 + p_1 t + \cdots + p_n t^n \quad \mapsto \quad \Gamma[P(t)] := \gamma_0 p_0 + \gamma_1 p_1 t + \cdots + \gamma_n p_n t^n.$$

We say that Γ is a multiplier sequence of the *first kind* (or *second kind*) if for all polynomials with real (or negative) roots, $\Gamma[P]$ has all real roots.

Theorem (Pólya–Schur, *J. reine angew. Math.* 1914)

A sequence Γ is a multiplier sequence of the first (or second) kind, if and only if its generating function $\Psi_\Gamma(t) := \sum_{j \geq 0} \gamma_j t^j / j!$ is in the first (or second) Laguerre–Pólya class.

Now by Schoenberg's results:
if and only if $1/\Psi_\Gamma(s)$ is the Laplace transform of a Pólya frequency function.

Modern incarnations of the Laguerre–Pólya–Schur program

Recently, Borcea–Branden (late 2000s) studied

- *higher-dimensional* versions of Pólya–Schur multipliers,
- linear operators on spaces of (multivariate) polynomials that preserved (higher dimensional) versions of stability / hyperbolicity.

This enabled them to characterize linear operators preserving Ω-stability for domains $\Omega \subset \mathbb{C}$ (studied+open since late 1800s at least); prove conjectures of C.R. Johnson; prove many Lee–Yang type theorems...

Modern incarnations of the Laguerre–Pólya–Schur program

Recently, Borcea–Branden (late 2000s) studied

- *higher-dimensional* versions of Pólya–Schur multipliers,
- linear operators on spaces of (multivariate) polynomials that preserved (higher dimensional) versions of stability / hyperbolicity.

This enabled them to characterize linear operators preserving Ω-stability for domains $\Omega \subset \mathbb{C}$ (studied+open since late 1800s at least); prove conjectures of C.R. Johnson; prove many Lee–Yang type theorems...

- Borcea–Branden–Liggett: developed the theory of negative dependence for 'strongly Rayleigh' (probability) measures $\rightsquigarrow$ proving conjectures of Liggett, Pemantle, Wagner.

Modern incarnations of the Laguerre–Pólya–Schur program

Recently, Borcea–Branden (late 2000s) studied

- *higher-dimensional* versions of Pólya–Schur multipliers,
- linear operators on spaces of (multivariate) polynomials that preserved (higher dimensional) versions of stability / hyperbolicity.

This enabled them to characterize linear operators preserving Ω-stability for domains $\Omega \subset \mathbb{C}$ (studied+open since late 1800s at least); prove conjectures of C.R. Johnson; prove many Lee–Yang type theorems...

- Borcea–Branden–Liggett: developed the theory of negative dependence for 'strongly Rayleigh' (probability) measures $\rightsquigarrow$ proving conjectures of Liggett, Pemantle, Wagner.
- Taken forward by Marcus–Spielman–Srivastava (2010s):
  - Kadison–Singer conjecture.
  - Existence of bipartite Ramanujan (expander) graphs of every degree and every order.

References

[1] A.K.
Matrix analysis and preservers of (total) positivity.
Lecture notes (website); forthcoming book with Cambridge Univ. Press, 2021.

[2] P.N. Choudhury, M.R. Kannan, and A.K.
Sign non-reversal property for TN/TP matrices; testing TP on their interval hull.
Bull. London Math. Soc., in press.

---

[3] E. Laguerre.
Sur les fonctions du genre zéro et du genre un. *C. R. Acad. Sci. Paris*, 1882.

[4] E. Laguerre.
Mémoire sur la théorie des équations numériques. *J. Math. Pures. Appl.*, 1883.

[5] M. Fekete and G. Pólya.
Über ein Problem von Laguerre. *Rend. Circ. Mat. Palermo.*, 1912.

[6] G. Pólya.
Über Annäherung durch Polynome mit lauter reellen Wurzeln. *Rend. Circ. Mat. Palermo*, 1913.

[7] G. Pólya and I. Schur.
Über zwei Arten von Faktorenfolgen in der Theorie der algebraischen Gleichungen. *J. reine angew. Math.*, 1914.

[8] I.J. Schoenberg.
On Pólya frequency functions. I. *J. d'Analyse Math.*, 1951.