

FUNCTIONS OPERATING ON POSITIVE DEFINITE MATRICES AND A THEOREM OF SCHOENBERG

BY

JENS PETER REUS CHRISTENSEN AND PAUL RESSEL

ABSTRACT. We prove that the set of all functions $f: [-1, 1] \rightarrow [-1, 1]$ operating on real positive definite matrices and normalized such that $f(1) = 1$, is a Bauer simplex, and we identify its extreme points. As an application we obtain Schoenberg's theorem characterising positive definite kernels on the infinite dimensional Hilbert sphere.

In his famous paper *Positive definite functions on spheres* [6], I. J. Schoenberg proved the following result: every continuous function $f: [-1, 1] \rightarrow \mathbb{R}$ which has the property that $f(\langle x, y \rangle)$ is a positive definite kernel on the infinite dimensional Hilbert sphere, has a power series representation with nonnegative coefficients. He obtained this result by first proving finite dimensional analogues involving ultraspherical polynomials and then going to the limit. As a by-product he acquired this representation for all continuous functions operating on positive definite matrices.

In the present paper we get a slightly more general result (without continuity assumptions) in a completely different way. Our main result is that the set of all functions on $[-1, 1]$ operating on positive definite matrices and normalized in a suitable way forms a Bauer simplex (i.e. a Choquet simplex whose set of extreme points is closed) and that the monomials $t \mapsto t^n$ together with their two limit points are precisely the extreme points of this simplex (Theorem 1). Using an approximation argument we obtain Schoenberg's theorem for the infinite dimensional Hilbert sphere (Theorem 2). Theorem 3 finally characterises the functions of the form $1 - f$, where $f(\langle x, y \rangle)$ is positive definite on the Hilbert sphere and $f(1) = 1$.

Some related results have been proved by Rudin and Horn. Rudin [5] showed that a function defined on the open interval $] -1, 1[$ and leaving invariant only the positive definite Toeplitz matrices can be represented as a power series with nonnegative coefficients. The idea of Rudin's proof is rather different from Schoenberg's and from ours. Horn ([3, p. 270]) proved that a function defined on $]0, \infty[$ which leaves invariant the cone of positive definite matrices with nonnegative entries must be absolutely monotonic and there-

Received by the editors August 31, 1976 and, in revised form, April 15, 1977.

AMS (MOS) subject classifications (1970). Primary 15A48; Secondary 42A88.

© American Mathematical Society 1978

fore ([8, p. 146]) analytic with nonnegative coefficients.

Let us recall some definitions: a real $n \times n$ -matrix (a_{ij}) is *positive (semi) definite* if and only if it is symmetric and

$$\sum_{i,j=1}^n c_i c_j a_{ij} \geq 0$$

for all $(c_1, \dots, c_n) \in \mathbb{R}^n$. The matrix (a_{ij}) is called *conditionally negative definite* if and only if it is symmetric and

$$\sum_{i,j=1}^n c_i c_j a_{ij} \leq 0$$

for all $(c_1, \dots, c_n) \in \mathbb{R}^n$ such that $\sum_{i=1}^n c_i = 0$. There is a very close connection between these two types of matrices: (a_{ij}) is conditionally negative definite iff $(e^{-ta_{ij}})$ is positive definite for all $t > 0$ (cf. [4]). If X is a nonempty set, then a function $\varphi: X \times X \rightarrow \mathbb{R}$ is called a *positive [conditionally negative] definite kernel* if and only if for every finite subset $\{x_1, \dots, x_n\} \subseteq X$ the $n \times n$ -matrix $(\varphi(x_i, x_j))$ is positive [conditionally negative] definite.

We introduce C_1 to be the set of all functions $f: [-1, 1] \rightarrow \mathbb{R}$ such that $(f(a_{ij}))$ is positive definite whenever (a_{ij}) is positive definite and $|a_{ij}| \leq 1$ for all $i, j = 1, \dots, n$. By the Schur Product Theorem ([7, p. 14]) C_1 is a multiplicative closed convex cone in $\mathbb{R}^{[-1,1]}$. It is immediate that each $f \in C_1$ is nonnegative on $[0, 1]$. For $t \in [-1, 1]$ the matrix $\begin{pmatrix} 1 & t \\ t & 1 \end{pmatrix}$ is positive definite, hence so is

$$\begin{pmatrix} f(1) & f(t) \\ f(t) & f(1) \end{pmatrix}$$

if $f \in C_1$, and the determinant criterion for positive definiteness gives us

$$|f(t)| \leq f(1) \quad \text{for all } t \in [-1, 1].$$

Therefore $K_1 := \{f \in C_1 | f(1) = 1\}$ is a compact convex base for the cone C_1 .

In order to prove our main result, Theorem 1 below, which identifies the extreme points of K_1 , we need two lemmas which we prove first.

LEMMA 1. For all $n \in 2\mathbb{N}_0 + 1$ the kernel $(s, t) \mapsto |\cos(s - t)|^n$ on \mathbb{R}^2 is not positive definite. For all $n \in 2\mathbb{N}$ the kernel

$$(s, t) \mapsto |\cos(s - t)|^n \operatorname{sgn}(\cos(s - t))$$

is not positive definite.

PROOF. We only prove the first statement. That the second one is also true can be seen in a completely analogous way. Let $n \in 2\mathbb{N}_0 + 1$ and assume that $|\cos(s - t)|^n$ is positive definite on \mathbb{R}^2 . By Bochner's theorem (cf. [2]) there exists a symmetric probability measure μ on \mathbb{R} such that

$$|\cos t|^n = \int_{-\infty}^{\infty} e^{it\lambda} d\mu(\lambda) = \int_{-\infty}^{\infty} \cos(t\lambda) d\mu(\lambda) \quad \text{for all } t \in \mathbb{R}.$$

From $|\cos \pi|^n = 1$ we see that $\mu(2\mathbb{Z}) = 1$. Now $(\cos t)^{2n}$ is the characteristic function of $\rho := \nu^{*2n}$ where $\nu := \frac{1}{2}(\varepsilon_{-1} + \varepsilon_1)$. The measure ρ is concentrated on $[-2n, 2n] \cap \mathbb{Z}$ and $\mu * \mu = \rho$, hence $\mu([-n, n] \cap \mathbb{Z}) = 1$. But n is odd, therefore $\mu(\{-n, n\}) = 0$ and $\mu([-(n-1), n-1] \cap \mathbb{Z}) = 1$ contradicting the fact that $\mu * \mu(\{2n\}) = \rho(\{2n\}) > 0$. \square

LEMMA 2. Let $f \in C_1$ and $t \in [-1, 1]$. Put $f_1(s) := f(s) + f(st)$, $f_2(s) := f(s) - f(st)$ for $s \in [-1, 1]$. Then f_1 and f_2 also belong to C_1 .

PROOF. Let (a_{ij}) be a fixed positive definite $n \times n$ -matrix, $|a_{ij}| \leq 1$ for all $i, j = 1, \dots, n$, and let $(c_1, \dots, c_n) \in \mathbb{R}^n$ also be fixed. Define $F: [-1, 1] \rightarrow \mathbb{R}$ by

$$F(s) := \sum_{i,j=1}^n c_i c_j f(s a_{ij}).$$

We claim that $F \in C_1$. In fact for any positive definite $m \times m$ -matrix (b_{pq}) , $|b_{pq}| \leq 1$ for all $p, q = 1, \dots, m$, and for all $(d_1, \dots, d_m) \in \mathbb{R}^m$ we get

$$\sum_{p,q=1}^m d_p d_q F(b_{pq}) = \sum_{p,q=1}^m \sum_{i,j=1}^n d_p d_q c_i c_j f(b_{pq} a_{ij})$$

which is nonnegative because the tensor product $(a_{ij}) \otimes (b_{pq})$ is positive definite, too. In particular

$$|F(t)| \leq F(1) = \sum c_i c_j f(a_{ij})$$

and this shows that f_1, f_2 both belong to C_1 . \square

THEOREM 1. The set K_1 is a Bauer simplex whose extreme points consist of the monomials $t \mapsto t^n$, $n \in \mathbb{N}_0$, and the discontinuous functions $1_{\{1\}} - 1_{\{-1\}}$ and $1_{\{-1,1\}}$.

PROOF. Let $f \in \text{ex}(K_1)$ and $t \in [-1, 1]$. By Lemma 2 there exists a constant $\lambda = \lambda(t) \geq 0$ such that

$$f(s) + f(st) = \lambda f(s) \quad \text{for all } s \in [-1, 1],$$

hence $\lambda = 1 + f(t)$ and therefore

$$f(st) = f(s)f(t) \quad \text{for all } s, t \in [-1, 1],$$

i.e. f is multiplicative. It is well known that there exists $\alpha \in [0, \infty]$ such that for all $s \in]0, 1]$, $f(s) = s^\alpha$ where $1^\infty := 1$, $s^\infty := 0 \quad \forall s \in]0, 1]$. If $\alpha < \infty$ is not integer then there exists a positive definite matrix (a_{ij}) with $a_{ij} > 0$ such that (a_{ij}^α) is not positive definite; see [3, pp. 270–271] for an elementary proof of this fact. We conclude that $\alpha \in \overline{\mathbb{N}}_0 = \{0, 1, 2, \dots, \infty\}$. If $\alpha = \infty$ then

$f = 1_{\{-1,1\}}$ or $f = 1_{\{1\}} - 1_{\{-1\}}$. Let $\alpha = n \in \mathbb{N}_0$ be finite. If $f(-1) = 1$ then $f(t) = |t|^n$; by Lemma 1, n has to be even, therefore

$$f(t) = t^n \quad \text{for all } t \in [-1, 1],$$

except for the special case that $n = 0$ and $f(t) = 1_{[-1,1] \setminus \{0\}}(t)$. If $f(-1) = -1$ then $f(t) = |t|^n \cdot \text{sgn}(t)$ and for $n > 0$ we get again from Lemma 1 that n must be odd, i.e.

$$f(t) = t^n \quad \text{for all } t \in [-1, 1]$$

holds again.

Two cases are left:

$$f(t) = \text{sgn}(t) \quad \text{and} \quad f(t) = \text{sgn}^2(t) = 1_{[-1,1] \setminus \{0\}}(t).$$

We shall show that these two functions do not belong to K_1 . Let $t_1 = 0$, $t_2 = \pi/2$ and $t_3 = \pi$ and put $a_{ij} := \frac{1}{2} + \cos(t_i - t_j)$, $i, j = 1, 2, 3$. The matrix

$$(a_{ij}) = \begin{pmatrix} 1 & \frac{1}{2} & 0 \\ \frac{1}{2} & 1 & \frac{1}{2} \\ 0 & \frac{1}{2} & 1 \end{pmatrix}$$

is positive definite, but

$$A := (1_{[0,1]}(a_{ij})) = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix}$$

has not this property, because $(1, -1, 1)A(1, -1, 1)' = -1 < 0$. This finishes the first half of the proof.

Let $T := \{t \mapsto t^n | n \in \mathbb{N}_0\} \cup \{1_{\{1\}} - 1_{\{-1\}}, 1_{[-1,1]}\}$. By the Schur Product Theorem the functions $t \mapsto t^n$ belong to K_1 and so do their two limit points, hence $T \subseteq K_1$. Let $K_0 := \{f: [-1, 1] \rightarrow [-1, 1] | f(st) = f(s)f(t) \ \forall s, t \in [-1, 1], f(1) = 1\}$. In Theorem 2.3 of [1] we have shown that $K_0 = \text{ex}(P)$, where P denotes the set of all $f: [-1, 1] \rightarrow [-1, 1]$ which are positive definite in the semigroup sense, i.e. which fulfill

$$\sum_{i,j=1}^n c_i c_j f(s_i s_j) \geq 0$$

for all $n \in \mathbb{N}$ and for all $\{s_1, \dots, s_n\} \subseteq [-1, 1]$, $(c_1, \dots, c_n) \in \mathbb{R}^n$, normalized such that $f(1) = 1$. Thus we have

$$\text{ex}(K_1) \subseteq T \subseteq K_1 \cap \text{ex}(P)$$

and finally $\text{ex}(K_1) = T$.

It is very easy to see that every function in K_1 has a unique integral (= series) representation over T . Therefore K_1 is a Bauer simplex. \square

Let H denote a fixed infinite dimensional real Hilbert space with scalar

product $\langle \cdot, \cdot \rangle$ and denote by $S := \{x \in H \mid \|x\| = 1\}$ the unit sphere of H . Restricted to S the scalar product is a mapping into the interval $[-1, 1]$ and plainly for every $f \in K_1$ the composed map $f(\langle x, y \rangle)$ is a positive definite kernel on $S \times S$. The next theorem shows that the converse of this result also holds.

THEOREM 2 (cf. SCHOENBERG [6, Theorem 2]). *Let $f: [-1, 1] \rightarrow [-1, 1]$ be a function such that*

$$(1) f(1) = 1,$$

$$(2) f(\langle x, y \rangle) \text{ is a positive definite kernel on } S \times S.$$

Then f belongs to K_1 .

PROOF. Let (a_{ij}) be a positive definite $k \times k$ -matrix, $|a_{ij}| \leq 1$ for all $i, j = 1, \dots, k$ and let $(c_1, \dots, c_k) \in \mathbb{R}^k$. We have to show that $\sum_{i,j=1}^k c_i c_j f(a_{ij}) \geq 0$. There exist vectors $x_1, \dots, x_k \in \mathbb{R}^k$ such that $a_{ij} = \langle x_i, x_j \rangle \forall i, j$. Without loss of generality we assume that $H = l^2(\mathbb{N})$ and define $x_{i\alpha} = (x_{i\alpha}(1), x_{i\alpha}(2), \dots) \in H$, $1 \leq i \leq k$, $1 \leq \alpha \leq n$ (n being some fixed natural number) by

$$x_{i\alpha}(p) := \begin{cases} x_i(p) & \text{if } 1 \leq p \leq k, \\ \sqrt{1 - \|x_i\|^2} & \text{if } p = k + (i-1)n + \alpha, \\ 0 & \text{else.} \end{cases}$$

(Note that $\|x_i\|^2 = a_{ii} \leq 1$.) Then

$$\langle x_{i\alpha}, x_{j\beta} \rangle = \begin{cases} \langle x_i, x_j \rangle = a_{ij} & \text{if } i \neq j, \\ \|x_i\|^2 = a_{ii} & \text{if } i = j \text{ and } \alpha \neq \beta, \\ 1 & \text{if } i = j \text{ and } \alpha = \beta, \end{cases}$$

and putting $d_{i\alpha} := c_i/n$ for $i = 1, \dots, k$ and $\alpha = 1, \dots, n$ we have

$$\begin{aligned} 0 &\leq \sum_{i,j=1}^k \sum_{\alpha,\beta=1}^n d_{i\alpha} d_{j\beta} f(\langle x_{i\alpha}, x_{j\beta} \rangle) \\ &= \sum_{i \neq j} c_i c_j f(a_{ij}) + \frac{n^2 - n}{n^2} \sum_{i=1}^k c_i^2 f(a_{ii}) + \frac{1}{n} f(1) \sum_{i=1}^k c_i^2 \\ &= \sum_{i,j=1}^k c_i c_j f(a_{ij}) + \frac{1}{n} \left[f(1) \sum_{i=1}^k c_i^2 - \sum_{i=1}^k c_i^2 f(a_{ii}) \right]. \end{aligned}$$

This holds for all $n \in \mathbb{N}$, hence $0 \leq \sum_{i,j=1}^k c_i c_j f(a_{ij})$ showing that $(f(a_{ij}))$ is positive definite. \square

Let (a_{ij}) be positive definite, $|a_{ij}| \leq 1$ for all i, j , and let $f \in K_1$. Consider the new matrix $(1 - f(a_{ij}))$; we claim that it is conditionally negative definite. Indeed, for $(c_1, \dots, c_n) \in \mathbb{R}^n$, $\sum_{i=1}^n c_i = 0$, we have

$$\sum_{i,j=1}^n c_i c_j (1 - f(a_{ij})) = - \sum_{i,j=1}^n c_i c_j f(a_{ij}) < 0.$$

This shows that $1 - f$ belongs to the closed convex cone C_2 of all functions $h: [-1, 1] \rightarrow [0, \infty[$ which satisfy the two conditions

$$h(1) = 0 \text{ and}$$

$(h(a_{ij}))$ is conditionally negative definite whenever (a_{ij}) is positive definite, $|a_{ij}| < 1 \forall i, j$.

The function $g := 1 - f$ belongs even to the subset K_2 of C_2 consisting of all $h \in C_2$ such that $h(0) < 1$, and furthermore g is a negative definite function on the multiplicative semigroup $[-1, 1]$ in the sense that $\sum_{i,j=1}^n c_i c_j g(s_i s_j) \leq 0$ for all $n \in \mathbb{N}$, $(s_1, \dots, s_n) \in [-1, 1]^n$ and $(c_1, \dots, c_n) \in \mathbb{R}^n$, $\sum_{i=1}^n c_i = 0$, cf. §3 in [1]. By Proposition 3.5 in [1] we have

$$|g(s) - g(0)| \leq g(s \cdot 0) = g(0) \quad \forall s \in [-1, 1]$$

implying that K_2 is a compact convex base for the cone C_2 .

THEOREM 3. *The transformation $\Phi: K_1 \rightarrow K_2$ defined by $\Phi(f) := 1 - f$ is an affine homeomorphism.*

PROOF. The only nontrivial point which remains to be proved is that $\Phi(K_1) = K_2$. Let $g \in K_2$, then e^{-tg} belongs to K_1 for all $t > 0$ and by Theorem 1

$$e^{-tg(s)} = \hat{\mu}_t(s) \quad \forall t > 0 \quad \forall s \in [-1, 1]$$

where the μ_t 's are Radon probability measures on K_0 , concentrated on the closed subset T (see the proof of Theorem 1 above for the definition of K_0 and T) and where $\hat{\mu}_t$ denotes the transform of μ_t in the sense of [1, §1]. From Remark 3.9 and Proposition 3.10 in [1] we conclude that g is negative definite on the semigroup $[-1, 1]$ and

$$g(s) = \int_{K_0 \setminus \{1\}} (1 - \rho(s)) d\mu(\rho) \quad \forall s \in [-1, 1]$$

where the finite positive Radon measure μ is the vague limit of $(\mu_t|_{K_0 \setminus \{1\}})/t$, $t \rightarrow 0$. This shows that μ is also concentrated on T . Note that $g(0) = \mu(T \setminus \{1\}) < 1$ and that therefore

$$f(s) := \int_{T \setminus \{1\}} \rho(s) d\mu(\rho) + 1 - \mu(T \setminus \{1\})$$

belongs to K_1 . Finally $1 - f = \Phi(f) = g$. \square

COROLLARY 1. *For every $g \in K_2$ there exist uniquely determined $a_{-2}, a_{-1}, a_1, a_2, \dots \geq 0$, $\sum a_n < 1$, such that*

$$g(s) = \sum_{n=1}^{\infty} a_n(1 - s^n) + a_{-1}(2 \cdot 1_{\{-1\}}(s) + 1_{]-1,1[}(s)) + a_{-2}1_{]-1,1[}(s)$$

for all $s \in [-1, 1]$ (cf. Theorem 3 in [6]).

COROLLARY 2. If $g: [-1, 1] \rightarrow [0, \infty]$, $g(1) = 0$, $g(0) < 1$ has the property that $g(\langle \cdot, \cdot \rangle)$ is a conditionally negative definite kernel on the infinite dimensional Hilbert sphere, then $g \in K_2$.

(This follows immediately from Theorem 2.)

REMARK. The functions operating on positive definite matrices we have considered in this paper were defined on the interval $[-1, 1]$. The analogous problems for functions on $[0, 1]$ have similar answers (and their proofs become a bit easier). It turns out that the set of all $f: [0, 1] \rightarrow [0, 1]$ such that $f(1) = 1$ and such that $(f(a_{ij}))$ is positive definite whenever (a_{ij}) is positive definite and $0 < a_{ij} < 1$ for all i, j , is a Bauer simplex whose extreme points consist of the monomials and their limit $1_{\{1\}}$. Theorems 2 and 3 also have their obvious counterparts.

We thank the referee for suggesting detailed improvements and for his valuable literary advice.

REFERENCES

1. C. Berg, J. P. R. Christensen and P. Ressel, *Positive definite functions on abelian semigroups*, Math. Ann. **223** (1976), 253–274.
2. S. Bochner, *Vorlesungen über Fouriersche Integrale*, Akademische Verlagsgesellschaft, Leipzig, 1932.
3. R. A. Horn, *The theory of infinitely divisible matrices and kernels*, Trans. Amer. Math. Soc. **136** (1969), 269–286.
4. C. Loewner, *On schlicht-monotonic functions of higher order*, J. Math. Anal. Appl. **14** (1966), 320–325.
5. W. Rudin, *Positive definite sequences and absolutely monotonic functions*, Duke Math. J. **26** (1959), 617–622.
6. I. J. Schoenberg, *Positive definite functions on spheres*, Duke Math. J. **9** (1942), 96–108.
7. I. Schur, *Bemerkungen zur Theorie der beschränkten Bilinearformen mit unendlich vielen Veränderlichen*, J. Reine Angew. Math. **140** (1911), 1–29.
8. D. Widder, *The Laplace transform*, Princeton Univ. Press, Princeton, N.J., 1946.

KOBENHAVNS UNIVERSITETS MATEMATISKE INSTITUT, UNIVERSITETSPARKEN 5, 2100 KOBENHAVN, DENMARK (Current address of J. P. R. Christensen)

Current address (Paul Ressel): University of Freiburg, Freiburg, Federal Republic of Germany