ROCKY MOUNTAIN
JOURNAL OF MATHEMATICS
Volume 6, Number 3, Summer 1976

POSITIVE DEFINITE FUNCTIONS AND GENERALIZATIONS, AN HISTORICAL SURVEY

JAMES STEWART

1. **Introduction.** A complex-valued function f of a real variable is said to be *positive definite* (abbreviated as p.d.) if the inequality

$$\sum_{i,j=1}^{n} f(x_i - x_j)\xi_i\bar{\xi}_j \geqq 0 \tag{1.1}$$

holds for every choice of $x_1, \cdots, x_n \in R$(the real numbers) and $\xi_1, \cdots, \xi_n \in C$(the complex numbers). In other words,the matrix

$$[f(x_i - x_j)]_{i,j=1}^{n} \tag{1.2}$$

is positive definite (strictly speaking we should say positive semi-definite or non-negative definite) for all n, no matter how the x_i's are chosen. A synonym for positive definite function is *function of positive type.*

For example, the function $f(x) = \cos x$ is p.d. because

$$\begin{aligned}\sum_{i,j=1}^{n} \cos(x_i - x_j)\xi_i\bar{\xi}_j &= \sum_{i,j=1}^{n} (\cos x_i \cos x_j + \sin x_i \sin x_j)\xi_i\bar{\xi}_j \\ &= \left| \sum_{i=1}^{n} \xi_i \cos x_i \right|^2 + \left| \sum_{i=1}^{n} \xi_i \sin x_i \right|^2 \geqq 0.\end{aligned}$$

Likewise it is easily verified directly that $e^{i\lambda x}$ is p.d. for real λ, but it is not so straightforward to see that such functions as $e^{-|x|}$, e^{-x^2}, and $(1 + x^2)^{-1}$ are p.d. These and other examples are discussed in § 3.

Positive definite functions and their various analogues and generalizations have arisen in diverse parts of mathematics since the beginning of this century. They occur naturally in Fourier analysis, probability theory, operator theory, complex function-theory, moment problems, integral equations, boundary-value problems for partial differential equations, embedding problems, information theory, and other areas. Their history constitutes a good illustration of the words of Hobson [51, p. 290]: "Not only are special results, obtained independently of one another, frequently seen to be really included in

Received by the editors on March 6, 1975.

some generalization, but branches of the subject which have developed quite independently of one another are sometimes found to have connections which enable them to be synthesized in one single body of doctrine. The essential nature of mathematical thought manifests itself in the discernment of fundamental identity in the mathematical aspects of what are superficially very different domains." Fourier [**36**] put it more succinctly: "[Mathematics] compares the most diverse phenomena and discovers the secret analogies which unite them."

To cite a specific instance, Mathias and the other early workers with p.d. functions of a real variable were chiefly concerned with Fourier transforms and apparently did not realize that more than a decade previously Mercer and others had considered the more general concept of positive definite kernels $K(x, y)$ (satisfying (1.1) with $f(x_i - x_j)$ replaced by $K(x_i, x_j)$) in research on integral equations. I have likewise found that present-day mathematicians working with some of the manifestations of p.d. functions are unaware of other closely related ideas. Thus one of the purposes of this survey is to correlate some of the more important generalizations of p.d. functions with the hope of making them better known. For example, probabilists are acquainted on the one hand with p.d. functions (in the guise of characteristic functions) and on the other hand with the Kolmogorov or Lévy-Khintchine formula for the logarithm of the characteristic function of an infinitely divisible random variable, but probably very few of them realize that the latter functions are examples of a significant generalization of p.d. functions, namely functions with a finite number of negative squares, and that Krein's integral representation for such functions may be of use to them.

It is not possible to discuss all of the analogues and generalizations of p.d. functions in this article; there are simply too many of them. Those to which we devote an entire section are p.d. sequences (which arose first), p.d. functions on groups, integrally p.d. functions, p.d. distributions, p.d. kernels, functions with a finite number of negative squares, and Schoenberg's functions which are p.d. in metric spaces. Some other generalizations are mentioned in the final section.

2. **Positive definite sequences.** The concept of a p.d. sequence was inspired by a problem of Carathéodory in complex function-theory. Contained in his paper [**16**], which appeared in 1907, was the following problem: What are necessary and sufficient conditions on the coefficients of the power series representation

$$w = f(z) = 1 + \sum_{k=1}^{\infty} (a_k + ib_k) z^k$$

of an analytic function f in order that it map the unit disk $|z| < 1$ into the right half-plane $\mathrm{Re}(w) > 0$? Carathéodory's answer was that, for each $n = 1, 2, 3, \cdots$, the point $(a_1, b_1, a_2, b_2, \cdots, a_n, b_n)$ in R^{2n} should lie in the smallest convex set containing the curve with parametric representation

$$(2\cos\theta, -2\sin\theta, 2\cos 2\theta, -2\sin 2\theta, \cdots, 2\cos n\theta,$$
$$-2\sin n\theta), 0 \leqq \theta \leqq 2\pi.$$

In 1911 Toeplitz [115] noticed that Carathéodory's conditions could be reformulated algebraically in terms of the non-negativity of certain Hermitian forms:

$$\sum_{i,j=1}^{n} c_{i-j}\xi_i\bar{\xi}_j \geqq 0, \text{ for } n = 1, 2, \cdots, \tag{2.1}$$

where $c_0 = 2, c_k = a_k - ib_k, c_{-k} = a_k + ib_k$. Any sequence $\{c_n\}$ which satisfies (2.1) is called *positive definite.*

Within a year several of the ablest mathematicians of the day published papers offering alternative proofs of the Carathéodory-Toeplitz results and pointing out connections with other areas of mathematics. In particular F. Riesz [90] saw the application to systems of integral equations, Herglotz [46] established the connection with the trigonometric moment problem, Carathéodory himself [17] considered the series expansion of positive harmonic functions, and further related papers were written by Fischer [35], Schur [102], and Frobenius [37]. Of these, the paper of Herglotz has turned out to have the most far-reaching consequences. The trigonometric moment problem can be stated as follows: Given a sequence $\{c_n\}_{-\infty}^{\infty}$ of complex numbers, what are necessary and sufficient conditions for the existence of a bounded non-decreasing function σ on $[-\pi, \pi]$ such that $\{c_n\}$ is the sequence of Fourier-Stieltjes coefficients of σ, i.e., $c_n = \int_{-\pi}^{\pi} e^{in\theta}\, d\sigma(\theta)$ for every integer n? Herglotz solved this problem by proving that a necessary and sufficient condition is the non-negativity of the Toeplitz forms (2.1), i.e., the positive-definiteness of $\{c_n\}$. We shall see that this theorem of Herglotz has many important analogues and generalizations.

For proofs, further details, and applications of p.d. sequences to problems in analysis and probability theory, we refer the reader to the books of Akhiezer and Krein [2], Akhiezer [1], and Grenander and Szegö [44]. Fan [33] has established various properties of p.d. sequences (without using Herglotz's Theorem) by representing them in terms of stationary sequences of vectors in a Hilbert space.

3. **Positive definite functions of a real variable.** Mathias, in 1923, [**69**], was the first person to define and study the properties of p.d. functions of a real variable. Motivated by the results of Carathéodory and Toeplitz he defined a complex-valued function f on R to be positive definite if

$$f(-x) = \overline{f(x)}, \text{ for } x \in R, \tag{3.1}$$

and the Hermitian form

$$\sum_{i,j=1}^{n} f(x_i - x_j)\xi_i\bar{\xi}_j \geqq 0, \tag{3.2}$$

for every choice of $x_1, \cdots, x_n \in R$ and $\xi_1, \cdots, \xi_n \in C$.

Condition (3.1) is superfluous, as F. Riesz, [**91**], pointed out. To see this, set $n = 2, x_1 = 0, x_2 = x, \xi_1 = 1$, and $\xi_2 = \xi$ in (3.2). Then

$$(1 + |\xi|^2)f(0) + \xi f(x) + \bar{\xi}f(-x) \geqq 0 \tag{3.3}$$

for every $\xi \in C$, and so $\xi f(x) + \bar{\xi}f(-x)$ is real for every $\xi \in C$. Setting $\xi = 1$, we have that $f(x) + f(-x)$ is real; setting $\xi = i$, we see that $i(f(x) - f(-x))$ is real. Thus (3.1) holds.

Mathias [**69**] observed the following elementary properties of positive definite functions:

I. If f is p.d., then so is $\bar{f}$.
II. If $f_1, \cdots, f_n$ are p.d. and $c_i \geqq 0$ then $f(x) = \sum_{i=1}^{n} c_i f_i(x)$ is p.d.
III. If each f_n is p.d., then so is $f(x) = \lim_{n\to\infty} f_n(x)$.
IV. Any p.d. function f is bounded, and in fact, $|f(x)| \leqq f(0)$.
V. The product of p.d. functions is p.d.

The first three properties are immediate consequences of the definition. The fourth follows from (3.3) by choosing ξ so that $\xi f(x) = -|f(x)|$. The fifth is a consequence of Schur's theorem [**101**] that the product of p.d. matrices is p.d.

The main theorem of Mathias, slightly rephrased, is that if f is p.d., then its Fourier transform $\hat{f}(t) = \int_{-\infty}^{\infty} f(x)e^{-itx}\,dx$ is non-negative (provided that it exists). Conversely if f satisfies the Fourier inversion formula and $f(t) \geqq 0$, then f is p.d. His proof makes use of the analogous result of Carathéodory and Toeplitz [**115**] for Fourier series.

We can make use of part of this theorem of Mathias to give some examples of p.d. functions. The functions $f_1(x) = e^{-|x|}, f_2(x) = e^{-x^2}$, $f_3(x) = (1 + x^2)^{-1}$, and $f_4(x) = 1 - |x|$ for $|x| \leqq 1, f_4(x) = 0$ for $|x| > 1$, are all p.d. because their Fourier transforms are positive and integrable. In fact it can be shown (see, e.g., Schoenberg [**97**]) that $f_\alpha(x)$

$= \exp[-|x|^\alpha]$ is p.d. if and only if $0 \leqq \alpha \leqq 2$. (Further examples are provided by Pólya's criterion [84]: Any real, even, continuous function f which is convex on $(0, \infty)$, i.e., $f(\frac{1}{2}(x_1 + x_2)) \leqq (\frac{1}{2})[f(x_1) + f(x_2)]$, and satisfies $\lim_{x\to\infty} f(x) = 0$, is p.d.)

Mathias did not prove the analogue of the theorem of Herglotz. That had to wait until 1932 when Bochner [11] proved the celebrated theorem which bears his name: *If f is a continuous p.d. function on R, then there exists a bounded non-decreasing function V on R such that f is the Fourier-Stieltjes transform of V, i.e.,*

$$f(x) = \int_{-\infty}^{\infty} e^{i\alpha x}\, dV(\alpha) \tag{3.4}$$

holds for all x.

The converse of this theorem is easy to prove, for if f has this form, then

$$\sum_{i,j=1}^{n} f(x_i - x_j)\xi_i\bar{\xi}_j = \sum_{i,j=1}^{n} \left\{ \int_{-\infty}^{\infty} \exp[i(x_i - x_j)\alpha]\, dV(\alpha) \right\} \xi_i\bar{\xi}_j$$

$$= \int_{-\infty}^{\infty} \left| \sum_{j=1}^{n} \xi_j \exp(ix_j\alpha) \right|^2 dV(\alpha) \geqq 0.$$

Bochner's Theorem itself is not so easy to establish, but in view of its great importance many, quite different, proofs have been given. (In particular, it can be deduced from the theorem of Herglotz. See, e.g., [54, p. 137].) The generalization to functions of several real variables was also given by Bochner [12].

In 1933, F. Riesz [91] extended Bochner's Theorem by proving that if we merely assume the measurability of a p.d. function f, then, for almost all x, $f(x)$ is equal to the Fourier-Stieltjes transform of a bounded non-decreasing function. An interesting refinement of this theorem of Riesz was given in 1956 by Crum [24] who showed that if f is a measurable p.d. function, then $f = p + r$, where p is a Fourier-Stieltjes integral, and the remainder function r is equal to zero almost everywhere (as Riesz had shown) and is also positive definite itself.

In order to indicate some of the many applications of Bochner's Theorem, we first mention that Bochner himself showed in 1933 [12] how it can be used to deduce and generalize much of the harmonic analysis of Wiener [120, 121].

About the same time both Bochner [13] and F. Riesz [91] showed how Bochner's Theorem can be applied to prove another important

theorem of the same period, namely Stone's Theorem of 1930 [**110**, **111**] on one-parameter groups of unitary operators. Let U_t, $-\infty < t < \infty$, be a group of unitary operators on a Hilbert space H, i.e., $U_0 = I$, the identity operator, and $U_{s+t} = U_s U_t$, and assume that $(U_t x, y)$ is a continuous function of t for every $x, y \in H$. Stone's Theorem asserts that there is a unique self-adjoint operator A on H with canonical resolution of the identity E such that $U_t = \int_{-\infty}^{\infty} e^{it\lambda}\, dE_\lambda$ $(\equiv e^{itA})$ in the sense that $(U_t x, y) = \int_{-\infty}^{\infty} e^{it\lambda}\, d(E_\lambda x, y)$, for $x, y \in H$. The proofs of Bochner and Riesz use the fact that for any $x \in H$ the function $f(t) = (U_t x, x)$ is p.d. Indeed

$$\sum_{i,j=1}^{n} (U_{t_i - t_j} x, x)\xi_i \bar{\xi}_j = \left(\sum_{i=1}^{n} \xi_i U_{t_i} x, \sum_{j=1}^{n} \xi_j U_{t_j} x \right) \geqq 0.$$

Stone's Theorem, in turn, has applications to quantum mechanics and ergodic theory. (See Riesz and Sz.-Nagy [**92**, §§138–139] for Bochner's proof of Stone's Theorem and applications.)

It is perhaps true to say that the area of mathematics in which the largest number of people are familiar with p.d. functions and Bochner's Theorem is that of probability theory. The Fourier-Stieltjes transform of the distribution function of a random variable is called a *characteristic function*, and so, by virtue of Bochner's Theorem, f is a characteristic function if and only if f is continuous, p.d., and $f(0) = 1$. Although characteristic functions can be traced back as far as Laplace and Cauchy, it was Paul Lévy [**63**, **64**] who first exploited systematically the fact that characteristic functions are in general easier to work with than distribution functions. This is especially true in connection with sums of independent random variables (which correspond to products of characteristic functions) and convergence of sequences of random variables (in part because for sequences $\{f_n\}$ of p.d. functions, $f_n \to f$ a.e. if and only if $f_n \to f$ uniformly on every finite interval; see § 4). Thus it is not surprising that the Central Limit Problem (the problem of convergence of laws of sequences of sums of random variables) was solved with the aid of p.d. functions.

Another occurrence of p.d. functions in probability theory is in the theory of stationary stochastic processes. Khintchine [**55**] used Bochner's Theorem to show that $R(t)$ is the correlation (covariance) function of a continuous stationary stochastic process if and only if it is of the form $R(t) = \int_{-\infty}^{\infty} \cos t\alpha\, dV(\alpha)$ where V is bounded and non-decreasing. See Fan [**34**] for this and other connections between p.d. functions and probability theory. Further connections can be found in Yaglom [**122**] and Schreiber, Sun and Barucha-Reid [**100**].

4. **Positive definite functions on groups.** With the advent of harmonic analysis on groups, and especially the Banach algebra approach to the subject in the 1940's, the central role of positive definite functions in Fourier analysis became apparent. The classical treatises on Fourier series and integrals of Zygmund [125] and Titchmarsh [114] had certainly managed to thrive without ever mentioning such functions. However in their present-day counterparts which deal with Fourier series and integrals, e.g., Edwards [27] and Katznelson [54], p.d. functions do have a role to play. Furthermore, virtually every book which treats harmonic analysis on groups gives prominence to p.d. functions. In some of these books, e.g., Loomis [67] and Rudin [93], p.d. functions play a very fundamental role indeed. The inversion formulas for the Fourier transform are based on the analogue of Bochner's Theorem, and then the duality theorem and Plancherel's Theorem are deduced. Consequently, in such treatments everything depends on p.d. functions.

The definition given in § 1 generalizes easily. A complex-valued function f defined on an arbitrary group G is *positive definite* if the inequality

$$\sum_{i,j=1}^{n} f(x_j^{-1} x_i)\xi_i\bar{\xi}_j \geqq 0 \tag{4.1}$$

holds for every choice of $x_i \in G$ and $\xi_i \in C$. Let P denote the set of all continuous p.d. functions on G. For the case where G is a locally compact abelian group (LCAG), Bochner's Theorem was generalized in 1940 (almost simultaneously) by Weil [118], Povzner [85], and Raikov [86] as follows. *A character* $\hat{x}$ of G is a homomorphism of G into the circle group, i.e., it satisfies $\hat{x}(xy) = \hat{x}(x)\hat{x}(y)$ and $|\hat{x}(x)| = 1$ for $x, y \in G$. The dual group $\hat{G}$ is the group of all continuous characters of G under pointwise multiplication with the topology of uniform convergence on compact subsets of G. The Weil-Povzner-Raikov Theorem says that if $f \in P$, then there is a positive bounded measure μ on $\hat{G}$ such that f is the Fourier-Stieltjes transform of μ, i.e.,

$$f(x) = \int_{\hat{G}} \hat{x}(x)\, d\mu(\hat{x}). \tag{4.2}$$

Let $B(G)$ be the algebra of all finite linear combinations of functions in P. Then (4.2) in conjunction with the Jordan decomposition theorem shows that $B(G)$ is precisely the set of all Fourier-Stieltjes transforms of bounded measures on $\hat{G}$.

Since every continuous character on R is of the form $\hat{x}(x) = e^{i\alpha x}$ for some $\alpha \in R$, we can consider $\hat{R} = R$, and (4.2) reduces to Bochner's

Theorem for $G = R$. Similarly if $G = Z$, the group of integers, then every continuous character is of the form $\hat{x}(n) = e^{in\theta}$ for some $\theta \in [-\pi, \pi]$, and (4.2) becomes Herglotz's Theorem.

Just as for $G = R$, many proofs of (4.2) have since appeared. In addition to the proofs in current texts, most of which, like Raikov's, use Banach algebra techniques, we cite in particular the proof of Cartan and Godement [19] which uses the Krein-Milman Theorem, the proof of Bingham and Parthasarathy [9] which uses probabilistic methods, and the proof of Bucy and Maltese [15] which uses the Choquet Representation Theorem. See also Phelps [82] and Choquet [20].

The proofs of Weil and Raikov referred to above also generalized Riesz's extension of Bochner's Theorem, i.e., any measurable p.d. function on a LCAG can be written as $f = p + r$ where p is the Fourier-Stieltjes transform of a positive measure and $r = 0$ locally almost everywhere. We have seen that for $G = R$, the residual function r is actually p.d. itself (Crum's Theorem, 1956). For an arbitrary LCAG this result is in fact a consequence of an earlier theorem (1950) of Segal and von Neumann [104, Thm. 2] on unitary representations, though they did not explicitly point this out. In 1960, Devinatz [26] made it explicit and gave a direct proof. A far-reaching generalization of this fact was given by de Leeuw and Glicksberg [62]. They showed that an arbitrary p.d. function (not necessarily measurable) on an arbitrary topological group G can be expressed as $f = p + r$, where $p \in P$ and r is a p.d. function which averages uniformly to 0 at e in the sense that $0 \in \bigcap_{V \in} \mathcal{C}(R_V r)$, where $\mathcal{V}$ is the set of all neighborhoods of the identity e of G, $\mathcal{C}$ denotes the closed convex hull in the set of all bounded functions on G, and the partial orbit $R_V r = \{R_g r : g \in V\}$, where $R_g r(h) = r(hg)$.

The importance of p.d. functions does not diminish when we turn our attention to non-abelian groups. The importance stems in part from the intimate connection among p.d. functions, unitary representations of the group, and positive functionals on the group algebra.

Let U be a unitary representation of a locally compact group G, i.e., each $U(x)$ is a unitary operator on a Hilbert space H, $U(e) = I$, and $U(xy) = U(x)U(y)$, for $x, y \in G$. Then, as in § 3, the function

$$f(x) = (U(x)\xi, \xi) \tag{4.3}$$

is p.d. for any $\xi \in H$. Conversely if f is p.d. Gelfand and Raikov [38] showed how to construct a unitary representation U which satisfies (4.3). Let H_0 be the set of functions ϕ on G such that $\phi(x) = 0$ except for finitely many x. If $(\phi, \psi) = \sum_{s,t} f(t^{-1}s)\phi(s)\overline{\psi(t)}$ then $(\phi, \phi) \geqq 0$.

Let H be the completion of H_0/N, where $N = \{\phi : (\phi, \phi) = 0\}$, and define U via $(U(x)\phi)(y) = \phi(xy^{-1})$. Then U is a unitary representation of G, and if ξ corresponds to the function which is equal to 1 at e and is zero elsewhere, we have $(U(x)\xi, \xi) = \sum_{s,t} f(t^{-1}s)\xi(xs)\overline{\xi(t)} = f(x)$. Both halves of this correspondence between unitary representations and p.d. functions are useful. For example, Ambrose [4] and Godement [41] used the first half, together with the Weil-Povzner-Raikov Theorem, to generalize Stone's Theorem to the situation where U is a continuous unitary representation of a LCAG. Conversely, the other half can be used to deduce (4.2) from the theorem of Ambrose and Godement. (See Nakano [79] for $G = R$ and Nakamura and Umegaki [78] in general.)

In order to describe another application, we introduce an ordering on P as follows: $\phi_1 > \phi_2 \Leftrightarrow \phi_1 - \phi_2 \in P$. A function $\phi_1 \in P$ is said to be *elementary* if the only functions $\phi_2 \in P$ with $\phi_1 > \phi_2$ are of the form $\phi_2 = \lambda\phi_1$, where $\lambda \in C$. It was proved by Gelfand and Raikov in 1943 [38], and independently by Godement [42], that ϕ is elementary if and only if the corresponding unitary representation U is irreducible (i.e., if S is a closed subspace of H with $U(x)S \subset S$ for all $x \in G$, then $S = H$ or $S = \{0\}$). Gelfand and Raikov expoited this correspondence to prove their famous theorem that every locally compact group admits "sufficiently many" irreducible unitary representations. Indeed it was Gelfand and Raikov who pointed out the full significance of the central role that p.d. functions play in analysis on locally compact groups.

Let $P_0 = \{\phi \in P : \phi(e) = 1\}$, the set of normalized functions in P. (If G is abelian, then $\hat{G}$ is the set of all elementary functions in P_0.) Then P_0 is a compact convex subset of $L^\infty(G)$ in the weak topology, and, by identifying the extreme points of P_0 as precisely the elementary functions in P_0, Gelfand and Raikov [38] used the Krein-Milman Theorem to show that any $f \in P$ is the weak limit in L^∞ of functions of the form $\sum \lambda_i\phi_i$, where $\lambda_i \geqq 0$, $\sum \lambda_i \leqq f(e)$, and the ϕ_i are elementary functions in P_0.

There have been various extensions of Bochner's Theorem to non-abelian locally compact groups. For the details we refer the reader to Godement [42, p. 52] and, for a more explicit version in the case where G is compact, to Krein [60, § 7]. (See also Hewitt and Ross [49, p. 334] for Krein's version.)

Convergence in P holds some surprises. Raikov [88] (and independently Yoshizawa [123]) proved that the mere assumption of pointwise convergence of a sequence of functions $f_n \in P$ to a function $f \in P$ implies that $f_n \to f$ uniformly on compact subsets of a locally compact

group. For an historical discussion of the relationship between such theorems and the various Cramér-Lévy convergence theorems of the 1920's and 1930's see McKennon [**70**, p. 62].

Several authors have recently shown interest in p.d. functions and analogues of Bochner's Theorem on groups which are abelian but not locally compact. For instance, the underlying additive group of an infinite-dimensional Banach space is such a group. We cite in particular the papers of Minlos [**74**] who deals with nuclear spaces (see also Gelfand and Vilenkin [**39**, p. 350]), Sazonov [**96**] and Gross [**45**] who deal with Hilbert spaces, and Waldenfels [**117**] who also deals with vector spaces. Shah [**106**] obtains a Bochner-type theorem for abelian groups on which the only restriction is the existence of a "quasi-invariant" measure.

5. **The extension problem.** In 1940, M. Krein [**57**] posed the following problem. Suppose that f is continuous and p.d. on the finite interval $[-A, A]$, i.e., the inequality (1.1) holds whenever $\xi_i \in C$ and $0 \leqq x_1 < x_2 < \cdots < x_n \leqq A$. Can f be extended to a continuous p.d. function on R, i.e., does there exist $g \in P$ such that $g(x) = f(x)$ for $-A \leqq x \leqq A$? Krein answered this question in the affirmative, and so by Bochner's Theorem, any such f is a Fourier-Stieltjes transform. In the same year, Raikov [**87**] proved this fact directly.

In the same paper, Krein also showed that the extension need not be unique, and, by employing methods reminiscent of those used in the classical moment problems, he gave several criteria for uniqueness of the extension. An example is the following result. Let B_A be the set of entire functions g which are bounded on the real axis and satisfy

$$\limsup_{r\to\infty} \frac{\log M(r)}{r} \leqq A, \text{ where } M(r) = \max\{|g(z)| : |z| \leqq r\}.$$

If f is p.d. on $[-A, A]$, it has a representation $f(x) = \int e^{itx}\, dF(t)$, for $|x| \leqq A$, by the theorems of Krein and Bochner. The functional $\Phi_f(g) = \int_{-\infty}^{\infty} g(t)\, dF(t)$ on B_A turns out to be independent of F. Krein proved that the p.d. extension of f is unique if there exists $g \in B_A$, $g \geqq 0$, $g \not\equiv 0$, satisfying $\Phi_f(g) = 0$. He also gave the example $f(x) = 1 - |x|$, $|x| \leqq A$, which is p.d. on $[-A, A]$ if and only if $0 < A \leqq 2$; the extension is unique for $A = 2$ but not unique for $0 < A < 2$.

Many authors have since given other criteria for uniqueness and tried to classify all possible extensions in the case of non-uniqueness. For example, both Akutowicz [3] and Devinatz [**25**] gave criteria in terms of the self-adjointness of certain operators on Hilbert spaces.

Part of the interest in such questions stems from probability theory. Several authors in the 1930's gave examples of distinct characteristic

functions which coincide on a finite interval, and this, of course, implies the non-uniqueness of the extension problem for p.d. functions. Esseen [31] and Lévy [65] treated the extension problem from the point of view of probability theory. See also Loève ([66, p. 212].

Chover [21] has dealt with an application of the extension problem to information theory. He showed that under certain conditions there is a unique extension of a p.d. function f which maximizes the entropy of the extension measure, i.e., among all the measures μ for which $f(x) = \int e^{ixt}\, d\mu(t)$, $|x| \leqq A$, there is one which carries the minimum amount of additional information.

The extension problem can be formulated in a more general context. If S is any subset of a group G, we say that f is positive definite on $S^{-1}S = \{y^{-1}x : x, y \in S\}$ if the inequality (4.1) holds whenever $\xi_i \in C$ and $x_i \in S$.

Rudin [94] has shown that the analogue of Krein's Theorem (on the existence of an extension if S is an interval of the real line) is false in higher-dimensional Euclidean spaces. Specifically, if S is an n-dimensional cube in R^n, where $n \geqq 2$, then there is a function which is p.d. on $S^{-1}S$ but has no p.d. extension to all of R^n. However, Rudin [95] has also shown that such extensions *do* exist in R^n if S is a ball instead of a cube and if the functions are radial. As far as the other classical groups are concerned, see Devinatz [25, § 7] for the non-existence of extensions when G is the circle group and Rudin [94] for the existence of extensions when $G = Z$ and non-existence when $G = Z^n$ and $n \geqq 2$.

For more general groups the interest lies in the case where S is a subgroup of G (and so, $S^{-1}S = S$). Actually it is not hard to see that any function which is p.d. on a subgroup G_0 can be extended so as to be p.d. on G simply by defining it to be zero outside of G_0. However the real interest lies in continuous p.d. extensions. Hewitt and Ross [49, p. 364] have shown that if G_0 is closed and G is either compact or LCA, then the problem of extending continuous p.d. functions from G_0 to G always has a solution. McMullen [71] has proved the same result for the case where G is locally compact and G_0 is compact. McMullen's monograph may be consulted for further results in this direction.

6. **Integrally positive definite functions.** The early workers in the field realized that for many purposes it is convenient to replace the inequality (1.1) by its integral analogue

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x - y)\phi(x)\overline{\phi(y)}\, dxdy \geqq 0, \tag{6.1}$$

where the function ϕ ranges over L^1 or C_c (the continuous functions

with compact support). Indeed if f is continuous, then (1.1) is equivalent to (6.1). (See Bochner [11] and also § 8.)

In view of this situation, Cooper [23] proposed the following definition in 1960. Given a set F of complex-valued functions on R, he called a function f *positive definite for* F if the integral in (6.1) exists as a Lebesgue integral and is non-negative for every $\phi \in F$. Let us denote by $P(F)$ the class of all functions which are p.d. for F. Clearly $F_1 \subset F_2$ implies that $P(F_1) \supset P(F_2)$. It turns out that $P(L^1)$ is identical, up to sets of measure zero, with the class of ordinary continuous p.d. functions. However, $P(C_c)$ is a much more extensive class of functions; ordinary p.d. functions are necessarily bounded, whereas the functions in $P(C_c)$ may be unbounded.

Cooper showed that $P(C_c) = P(L_c^p)$ for every $p \geqq 2$, where L_c^p is the set of functions in L^p with compact support. Furthermore, if $1 \leqq p \leqq 2$ and $q = p/2(p-1)$, then any function in $P(L_c^2)$ which is locally in L^q is in $P(L_c^p)$. (The converse is false; see Stewart [108].) These results indicate that for all practical purposes the gamut of p.d. functions is spanned by the classes $P(L_c^p)$, where $1 \leqq p \leqq 2$; as p increases from 1 to 2, $P(L_c^p)$ increases from the smallest class of p.d. functions to the largest such class. (However the examples at the end of this section and in § 9 show that $P(F)$ can be larger when F is *substantially* smaller than the usual function classes C_c, C_c^∞, L_c^p, etc.)

Cooper's principal result is a generalization of Bochner's Theorem for these integrally p.d. functions: if $f \in P(C_c)$, then there is a non-decreasing function V, not necessarily bounded, such that the equation $f(x) = \int e^{i\alpha x}\, dV(\alpha)$ holds in the sense of Cesàro summability almost everywhere. Furthermore the function V must satisfy $V(\alpha) = o(\alpha)$ as $\alpha \to \pm\infty$. For example, if we take $V(\alpha) = \sqrt{\alpha}$ for $\alpha \geqq 0$, and $V(\alpha) = 0$ for $\alpha < 0$, then

$$f(x) = \int_{-\infty}^{\infty} e^{i\alpha x}\, dV(\alpha) = \begin{cases} \dfrac{1+i}{2}\sqrt{\dfrac{\pi}{2x}}, & \text{for } x > 0 \\[2ex] \dfrac{1-i}{2}\sqrt{\dfrac{\pi}{-2x}}, & \text{for } x < 0 \end{cases}$$

is p.d. for C_c, but it is not an ordinary p.d. function, since it is unbounded.

Functions of this wider class $P(C_c)$ had been studied earlier under certain restrictions. Weil [118] had considered those functions in $P(C_c)$ which belong to L^p for some $p \geqq 1$, whereas Cooper [22] had

investigated the functions in $P(C_c)$ which satisfy the condition $\int_0^h |f(x)|\,dx = 0(|h|^\alpha)$ as $h \to 0$, where $0 \leqq \alpha \leqq 1$.

The notion of integrally p.d. functions makes sense on any locally compact group G. If integration with respect to left-invariant Haar measure on G is denoted by dx, then $P(F)$ consists of those f for which $\int\int f(y^{-1}x)\phi(x)\overline{\phi(y)}\,dxdy$ exists and is non-negative for every $\phi \in F$. In recent papers by Hewitt and Ross [48], Edwards [28], and Rickert [89], constructions have been given on non-discrete locally compact groups for functions in $P(C_c)$ which are not in L^∞ and, therefore, not almost everywhere equal to the ordinary continuous p.d. functions. A Bochner-type theorem, which generalizes Cooper's Theorem on the one hand and the Weil-Povzner-Raikov Theorem on the other hand, was proved by Stewart [107, Thm. 4.2]. Any $f \in P(C_c)$ on a LCAG is the Fourier-Stieltjes transform (in a suitable summability sense) of a positive measure μ, possibly unbounded, on $\hat{G}$. Furthermore, μ must satisfy $\mu(\hat{x} + K) \to 0$ as $\hat{x} \to \infty$, where K is any compact subset of G.

It would be interesting to find integral representation theorems for functions in $P(F)$ which would reflect any kind of symmetry that the class F might possess. For instance, if $G = R$ and E denotes the even functions in C_c, then any even continuous function $f \in P(E)$ is of the form

$$f(x) = \int_0^\infty \cos \lambda x\, d\mu_1(\lambda) + \int_0^\infty \cosh \lambda x\, d\mu_2(\lambda), \tag{6.2}$$

where μ_1 and μ_2 are positive measures, μ_1 is finite, and μ_2 is such that the second integral converges. (See Gelfand and Vilenkin [39, p. 197] where the result is attributed to Krein.) More generally, if $G = R^n$ and F is symmetric with respect to a group of rotations (or more general linear transformations), is there an analogue of the formula (6.2) for functions in $P(F)$ which conveys the symmetry of F? Partial answers have been given by Nussbaum [81] (for the orthogonal group) and Tang [113].

7. **Distributions.** Schwartz [103] has extended the theory of p.d. functions to distributions. Let C_c^∞ be the space of infinitely differentiable functions with compact support on R, and give C_c^∞ the topology usual for the theory of distributions, i.e., $\phi_n \to 0$ in $C_c^\infty \Leftrightarrow$ the supports of all the ϕ_n's lie in a common compact set, and ϕ_n and all its derivatives converge uniformly to 0. It T is a distribution, i.e., a continuous linear functional on C_c^∞, then T is called positive definite if $T(\phi * \phi^*) \geqq 0$ for all $\phi \in C_c$, where $\phi^*(x) = \overline{\phi(-x)}$ and $\phi * \psi$ denotes convolution: $\phi * \psi(x) = \int_{-\infty}^{\infty} \phi(x - y)\psi(y)\,dy$. In order to see why this definition can be considered as a generalization of p.d. functions, let us

consider the distribution T_f which is associated with any locally integrable function f:

$$T_f(\phi) = \int_{-\infty}^{\infty} f(x)\phi(x)\,dx.$$

If f is p.d., or more generally is in $P(C_c^\infty)$, then the associated distribution T_f is p.d. according to Schwartz's definition, because

$$\begin{aligned} T_f(\phi * \phi^*) &= \int_{-\infty}^{\infty} f(x)\phi * \phi^*(x)\,dx \\ &= \int_{-\infty}^{\infty} f(x)\,dx \int_{-\infty}^{\infty} \phi(x+y)\overline{\phi(y)}\,dy \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x-y)\phi(x)\overline{\phi(y)}\,dxdy \geqq 0. \end{aligned}$$

The analogue of Bochner's Theorem is Schwartz's representation for a p.d. distribution as the Fourier transform of a positive tempered measure μ, i.e., $T(\phi) = \int_{-\infty}^{\infty}\hat{\phi}(x)\,d\mu(x)$ where $\hat{\phi}$ is the Fourier transform of ϕ, and, for some $p \geqq 0$, $\int_{-\infty}^{\infty} d\mu(x)/(1+|x|^2)^p < \infty$. This Bochner-Schwartz Theorem has been extended to distributions on LCA groups by Maurin and Wawrzyńczyk [**116**].

Positive definite distributions have found applications in the theory of generalized random processes. See Gelfand and Vilenkin [**39**, Ch. 3].

8. **Kernels.** We can generalize the notion of a p.d. function by replacing $f(x_i - x_j)$ in (1.1) by $K(x_i, x_j)$. If $K(x, y)$ is any complex-valued function on R^2, we call K a *positive definite kernel* if

$$\sum_{i,j=1}^{n} K(x_i, x_j)\xi_i\bar{\xi}_j \geqq 0 \tag{8.1}$$

holds whenever $x_i \in R$ and $\xi_i \in C$. Although this concept is, of course, more general than that of p.d. functions, it appeared earlier. In fact p.d. kernels, as defined by (8.1), seem to have arisen first in 1909 in a paper by Mercer [**72**] on integral equations, and, although several other authors made use of this concept in the following two decades, none of them explicitly considered kernels of the form $K(x, y) = f(x-y)$, i.e., p.d. functions. Indeed Mathias and Bochner seem not to have been aware of the study of p.d. kernels.

Mercer's work arose from Hilbert's paper of 1904 [**50**] on Fredholm integral equations of the second kind:

$$f(s) = \phi(s) - \lambda \int_a^b K(s, t)\phi(t)\, dt. \tag{8.2}$$

In particular, Hilbert had shown that

$$\int_a^b \int_a^b K(s, t)x(s)x(t)\, dsdt = \sum \frac{1}{\lambda_n} \left[\int_a^b \psi_n(s)x(s)\, ds \right]^2, \tag{8.3}$$

where K is a continuous real symmetric [$K(s, t) = K(t, s]$ kernel, x is continuous, $\{\psi_n\}$ is a complete system of orthonormal eigenfunctions, and the λ_n's are the corresponding eigenvalues of (8.2), i.e., $\psi_n(s) = \lambda_n \int_a^b K(s, t)\psi_n(t)\, dt$. Hilbert defined a "definite" kernel as one for which the double integral $J(x) = \int_a^b \int_a^b K(s, t)x(s)x(t)\, dsdt$ satisfies $J(x) > 0$ except for $x(s) \equiv 0$.

The original object of Mercer's paper [72] was to characterize the kernels which are definite in the sense of Hilbert, but Mercer soon found that the class of such functions was too restrictive to characterize in terms of determinants. He therefore defined a continuous real symmetric kernel $K(s, t)$ to be of *positive type* if $J(x) \geqq 0$ for all real continuous functions x on $[a, b]$, and he proved that (8.1) is a necessary and sufficient condition for a kernel to be of positive type. (In view of (8.3) a necessary and sufficient condition for a kernel to be p.d. is that all its eigenvalues be positive.) Mercer then proved that for any continuous p.d. kernel the expansion

$$K(s, t) = \sum \frac{\psi_n(s)\psi_n(t)}{\lambda_n}$$

holds absolutely and uniformly.

At about the same time, W. H. Young [124], motivated by a different question in the theory of integral equations, showed that for continuous kernels (8.1) is equivalent to $J(x) \geqq 0$ for all x in $L^1[a, b]$. In a later paper, Mercer [73] considered unbounded kernels of positive type, thereby anticipating the integrally p.d. functions considered in § 6.

E. H. Moore [75, 76] initiated the study of a very general kind of p.d. kernel. If E is an abstract set, he called functions $K(x, y)$ defined on $E \times E$ "positive Hermitian matrices" if they satisfy (8.1) for all $x_i \in E$. Moore was interested in a generalization of integral equations and showed that to each such K there is a Hilbert space H of functions such that, for each $f \in H$, $f(y) = (f(x), K(x, y))$. This property is called the *reproducing property* of the kernel and turns out to have importance in the solution of boundary-value problems for elliptic partial differential equations. For an account of reproducing kernels,

with further indications of their applications, see Aronszajn [6], especially the historical introduction.

Another line of development in which p.d. kernels played a large role was the theory of harmonics on homogeneous spaces as begun by E. Cartan [18] in 1929 and continued by Weyl [119] and Ito [53]. The most comprehensive theory of p.d. kernels on homogeneous spaces is that of Krein [60] which includes as special cases not only the work of the above three authors but also the work of Gelfand and Raikov (described in § 4) on p.d. functions and irreducible unitary representations of locally compact groups. Krein was able to represent a broad class of kernels on homogeneous spaces in the form $K(x, y) = \int_T Z(x, y; t)\, d\sigma(t)$, where σ is bounded and positive on a certain space T, and each $Z(x, y; t)$ is a "zonal" kernel (analogous to the elementary p.d. functions). As special cases of this very comprehensive representation we can mention the Weil-Povzner-Raikov Theorem and Schoenberg's results described in § 10, specifically equations (10.1) and (10.2). See Hewitt [47; pp. 145–149] for an exposition of this work of Krein.

In probability theory p.d. kernels arise as covariance kernels of stochastic processes. (See Loève [66, p. 466].)

9. **Functions with a finite number of negative squares.** In this section we discuss a generalization of p.d. functions which is due to M. Krein. Although certain special cases had been dealt with earlier, it was in 1959 [61] that Krein formulated the concept of functions with k negative squares. These are the complex-valued functions which are Hermitian, in the sense that $\overline{f(-x)} = f(x)$, and such that the form (1.1) has at most k negative squares (when reduced to diagonal form) for every choice of n and $x_1, \cdots, x_n \in R$, and at least one of these forms has exactly k negative squares. In other words the matrix (1.2) has at most k negative eigenvalues no matter how the x_i's are chosen, and has exactly k negative eigenvalues for some choice of the x_i's. For example, the function $f(x) = \cosh x$ has one negative square because

$$\sum \cosh(x_i - x_j)\xi_i\bar{\xi}_j = |\sum \xi_i \cosh x_i|^2 - |\sum \xi_i \sinh x_i|^2.$$

Krein proved that if f is a continuous function with k negative squares, then there is a positive measure μ and a polynomial Q of degree k such that

$$f(x) = h(x) + \int_{-\infty}^{\infty} \frac{e^{i\lambda x} - S(x, \lambda)}{|Q(\lambda)|^2}\, d\mu(\lambda), \tag{9.1}$$

where h is a solution of the differential equation

$$\bar{Q}\left(-i\frac{d}{dx}\right)Q\left(-i\frac{d}{dx}\right)h(x) = 0, (\bar{Q}(\lambda) = Q(\bar{\lambda})),$$

and $S(x, \lambda)$ is a regularizing correction (compensating for the real zeros of Q). Notice that in the case $k = 0$ the definition reduces to that of a p.d. function, and Krein's integral representation (9.1) becomes Bochner's Theorem. The proof makes use of Pontryagin's Theorem on invariant subspaces associated with self-adjoint operators in spaces with indefinite scalar product.

Krein had earlier given an integral representation for continuous functions with one negative square in connection with the problem of the continuation of screw lines in infinite-dimensional Lobachevski space. See [**59**] or [**52**, Theorem 6.2] for a simplified version of (9.1) in the case $k = 1$. The generalization of (9.1) to functions of several variables appears in Gorbachuk [**43**].

Iohvidov and Krein [**52**, Theorem 5.2] proved an integral representation analogous to (9.1) for sequences with k negative squares which reduces to the theorem of Herglotz when $k = 0$.

Functions with a finite number of negative squares have applications to probability theory, in particular to infinitely divisible distribution laws and, therefore, to stochastic processes with stationary increments. A random variable is said to be infinitely divisible if for every positive integer n it can be expressed as the sum of n independent and identically distributed random variables. We shall show that the logarithm $f(x)$ of the characteristic function $g(x)$ of such a random variable has at most one negative square. For any n, $g(x) = [g_n(x)]^n$, where g_n is a characteristic function, and hence is p.d., i.e., $\exp[f(x)/n]$ is p.d. But the product of p.d. functions is p.d., and so $\exp[(m/n)f(x)]$ is p.d. Since the limit of p.d. functions is p.d., it follows that $\exp[tf(x)]$ is p.d. whenever $t \geqq 0$. Thus for any $x_1, \cdots, x_n \in R$ and $t \geqq 0$, we have

$$\begin{aligned} 0 &\leqq \sum_{i,j=1} \exp[tf(x_i - x_j)]\,\xi_i\bar{\xi}_j \\ &= |\sum \xi_i|^2 + t\sum f(x_i - x_j)\xi_i\bar{\xi}_j \\ &\quad + \frac{t^2}{2}\sum \exp[\theta_{ij}tf(x_i - x_j)]\,f^2(x_i - x_j)\xi_i\bar{\xi}_j, \end{aligned}$$

where $0 < \theta_{ij} < 1$. By letting $t \to 0$, we see from this that $\sum f(x_i - x_j)\,\xi_i\bar{\xi}_j \geqq 0$ whenever $\sum \xi_i = 0$, and so f has at most one negative square. If we apply Krein's formula (9.1) with $k = 1$, we obtain Kolmogorov's formula

$$\log g(x) = i\gamma x + \int_{-\infty}^{\infty} \frac{e^{i\lambda x} - 1 - i\lambda x}{\lambda^2}\, d\mu(\lambda) \tag{9.2}$$

for the logarithm of the characteristic function of an infinitely divisible distribution with finite variance. Formula (9.2) was first proved by Kolmogorov [**56**] in 1932, but, of course, not by the above method.

One can attempt to generalize the concept of functions with a finite number of negative squares in the same manner as described in §§ 4, 6, 7, and 8 for p.d. functions. There is no problem in formulating the definition and elementary properties of such functions on general groups (see [**109**]), but no one seems to have found an analogue of Krein's integral representation for groups. However the results of §§ 6, 7, 8 do have analogues. Shah Tao-Shing [**105**] has characterized distributions with k negative squares by a formula which generalizes both Krein's formula (9.1) and the Bochner-Schwartz Theorem. (Such distributions are useful in generalized random processes. See [**39**, chapters 2, 3].) Gorbachuk [**83**, **43**] has generalized Krein's work by considering kernels $K(x, y)$ with k negative squares. Stewart [**109**] has enlarged Krein's class of functions with k negative squares and generalized (9.1) in a sense similar to that in which Cooper's results (§ 6) on integrally p.d. functions extend those of Bochner.

In this connection we mention that although functions with k negative squares are clearly a *generalization* of p.d. functions, in another sense they can be regarded as a *special case* of p.d. functions. The explanation of this paradox lies in the fact that if f has k negative squares, then it satisfies the inequality

$$\int\int f(x-y)\bar{Q}\left(i\frac{d}{dx}\right)\phi(x)\overline{\bar{Q}\left(i\frac{d}{dx}\right)\phi(y)}\, dxdy \geqq 0$$

for every $\phi \in C_c^{\infty}$, the infinitely differentiable functions with compact support, where Q is the polynomial in (9.1), and so, in the notation of § 6, $f \in P(F)$, where $F = \bar{Q}(id/dx)C_c^{\infty}$.

10. **Metric spaces.** In 1938 Schoenberg published a certain generalization of real p.d. functions which arose from the problem of isometrically embedding metric spaces in Hilbert space [**97**, **98**, **99**] and the related problem of determining the screw lines and screw functions of Hilbert space [**80**].

Schoenberg calls a set S a *quasi-metric space* if it has a distance function d with the following two properties: (i) $d(P, P') = d(P', P) \geqq 0$, (ii) $d(P, P) = 0$, where $P, P' \in S$. A real, continuous, even function f, defined in the range of values of $\pm\, d(P, P')$, is called *positive definite in* S if for any $P_i \in S$ and $\xi_i \in R$, we have $\sum_{i,j=1}^{n} f(d(P_i, P_j))\xi_i\xi_j \geqq 0$.

(Notice that, for $S = R$, this agrees with the definition of the ordinary continuous real p.d. functions.) For example, the equation

$$\exp\Big(-\sum_1^m x_j^2\Big) = \int\cdots\int \exp\Big(-\tfrac{1}{4}\sum_1^m u_j^2\Big)\exp\Big(i\sum_1^m x_j u_j\Big)\,du_1\cdots du_m$$

shows that the function $f(t) = e^{-t^2}$ is p.d. in R^m, and, hence, in a real Hilbert space. One connection between these p.d. functions and the embedding problem is Schoenberg's result [97] that a separable quasi-metric space is isometrically embeddable in a real Hilbert space if and only if the function $f(t) = e^{-\lambda t^2}$ is p.d. in S for every $\lambda > 0$.

Schoenberg has proved a number of interesting integral representation theorems for such functions. For example, in [98] any function p.d. in R^m was shown to be of the form

$$f(t) = \int_0^\infty \Omega_m(tu)\,d\alpha(u), \tag{10.1}$$

where α is bounded and non-decreasing, and Ω_m is essentially a Bessel function:

$$\Omega_m(t) = \Gamma\left(\frac{m}{2}\right)\left(\frac{2}{t}\right)^{(\frac{1}{2})(m-2)} J_{(\frac{1}{2})(m-2)}(t).$$

In particular $\Omega_1(t) = \cos t$, $\Omega_2(t) = J_0(t)$, $\Omega_3(t) = (\sin t)/t$, and so (10.1) reduces to Bochner's Theorem for real functions when $m = 1$. By letting $m \to \infty$ in (10.1), the representation

$$f(t) = \int_0^\infty e^{-t^2u^2}\,d\alpha(u) \tag{10.2}$$

is deduced for functions p.d. in Hilbert space.

Let S_m denote the unit sphere in R^{m+1} and $d(P, P')$ denote spherical distance. Schoenberg's representation for functions p.d. in S_2 [99] is $f(t) = \sum_{n=0}^\infty a_n P_n(\cos t)$, where $a_n \geqq 0$, $\sum a_n < \infty$, and P_n is a Legendre polynomial. A similar formula, in terms of ultraspherical polynomials, holds for functions p.d. in S_m and was extended by Bochner [14] from spheres to compact spaces with transitive groups of transformations using the generalized spherical harmonics of Cartan and Weyl. Again, letting $m \to \infty$ gives the general form of functions p.d. in the unit sphere in Hilbert space: $f(t) = \sum_{n=0}^\infty a_n \cos^n t$, where

$a_n \geqq 0$, $\sum a_n < \infty$. For recent extensions of Schoenberg's work, see Bingham [10] and the references therein.

Finally we note Einhorn's discovery [29, 30] of the extreme paucity of functions p.d. in $C[0, 1]$, the space of continuous functions on $[0, 1]$ with the supremum metric. The only such functions are the positive constants.

11. **Other generalizations.** A complete list of the current generalizations of p.d. functions would be very long and exhibit great variety. In this final section we give brief mention to some of the more interesting ideas in this list which have not already been covered.

(i) Krein [58] and Berezanski [8] have constructed a theory which generalizes Bochner's in that $e^{i\alpha x}$ is replaced by eigenfunctions of differential (and more general) operators. Their integral representation theorem includes as special cases not only Bochner's Theorem but also Bernstein's Theorem on the representation of completely monotone functions as Laplace-Stieltjes integrals.

(ii) Positive functionals on Banach algebras with involution (see, e.g., Loomis [67, p. 96]) can be considered as a generalization of p.d. functions in view of the 1-1 correspondence established by Gelfand and Raikov [38] between P and the positive functionals on the group algebra $L^1(G)$. Any functional of the form $L(\phi) = \int f(x)\phi(x)\,dx$, where $f \in P$, and dx denotes integration with respect to Haar measure, is a positive functional (cf. § 7), and, conversely, any positive functional on $L^1(G)$ is of this form. Furthermore there is a version of the Weil-Povzner-Raikov Theorem which is valid for positive functionals on certain algebras. See [67, Thm. 26I] and Lumer [68].

(iii) There is a natural generalization from p.d. functions to p.d. measures. The inequality (6.1) can be rewritten as $\int \phi * \phi^*(x) f(x)\,dx \geqq 0$, for all $\phi \in C_c$ (cf. § 7). In view of this, a measure is said to be positive definite if $\int \phi * \phi^*(x)\,d\mu(x) \geqq 0$ for all $\phi \in C_c$. Such measures have been studied by Godement [42] and Argabright and Gil de Lamadrid [5].

(iv) So far all the generalizations that we have discussed have been numerical-valued functions, but several authors have considered functions with more general range. Falb and Haussmann [32] have given a representation like Bochner's for p.d. functions with values in a Banach space. Of particular interest are p.d. functions on a group whose values are operators in a Hilbert space, and such functions have been studied by Naimark [77] and Sz.-Nagy [112]. There are two possible ways of defining such functions. For the connection between the definitions and further references see the exposition in § 2 of

Berberian [7]. In the appendix to [92], Sz.-Nagy considers p.d. operator-valued functions on semigroups with involution and gives applications to contraction operators and extensions of operators.

References

1. N. Akhiezer, *The Classical Moment Problem and some Related Questions in Analysis*, Hafner, New York, 1965.

2. N. Akhiezer and M. Krein, *Some Questions in the Theory of Moments*, Amer. Math. Soc. Translations of Math. Monographs, Vol. 2, Providence, R.I., 1962.

3. E. J. Akutowicz, *On extrapolating a positive definite function from a finite interval*, Math. Scand. 7 (1959), 157–169.

4. W. Ambrose, *Spectral resolution of groups of unitary operators*, Duke Math J. 11(1944), 589–595.

5. L. Argabright and J. Gil de Lamadrid, *Fourier analysis of unbounded measures on locally compact abelian groups*, Memoirs Amer. Math. Soc. 145, Providence, R.I., 1974.

6. N. Aronszajn, *Theory of reproducing kernels*, Tran. Amer. Math. Soc. 68 (1950), 337–404.

7. S. K. Berberian, *Naimark's moment theorem*, Michigan Math. J. 13(1966), 171–184.

8. Yu. M. Berezanski, *A generalization of a multidimensional theorem of Bochner*, Soviet Math. Dokl. 2(1961), 143–147.

9. M. S. Bingham and K. R. Parthasarathy, *A probabilistic proof of Bochner's Theorem on positive definite functions*, J. London Math. Soc. 43(1968), 626–632.

10. N. H. Bingham, *Positive definite functions on spheres*, Proc. Camb. Phil. Soc. 73(1973), 145–156.

11. S. Bochner, *Vorlesungen über Fouriersche Integrale*, Akademische Verlagsgesellschaft, Leipzig, 1932.

12. ——, *Monotone Funktionen*, Stieltjessche Integrale und harmonische Analyse, Math. Ann. 108(1933), 378–410. [References 11 and 12 were translated by M. Tenenbaum and H. Pollard as: *Lectures on Fourier Integrals* (with the author's supplement on Monotonic Functions, Stieltjes Integrals and Harmonic Analysis), Annals of Mathematics Studies No. 42, Princeton Univ. Press, 1959.]

13. ——, *Spektralzerlegung linearer Scharen unitärer Operatoren*, Sitzungsberichte Preuss. Akad. Wiss. phys.-math. Kl., (1933), 371–376.

14. ——, *Hilbert distances and positive definite functions*, Ann. of Math. 42(1941), 647–656.

15. R. S. Bucy and G. Maltese, *Extreme positive definite functions and Choquet's representation theorem*, J. Math. Anal. Appl. 12(1965), 371–377.

16. C. Carathéodory, *Über den Variabilitätsbereich der Koeffizienten von Potenzreihen*, die gegebene Werte nicht annehmen, Math. Ann. 64(1907), 95–115.

17. ——, *Über den Variabilitätsbereich der Fourierschen Konstanten von positiven harmonischen Funktionen*, Rend. Circ. Mat. Palermo 32(1911), 193–217.

18. E. Cartan, *Sur la détermination d'un système orthogonal complet dans un espace de Riemann symétrique clos*, Rend. Circ. Mat. Palermo 53(1929), 217–252.

19. H. Cartan and R. Godement, *Théorie de la dualité et analyse harmonique dans les groupes abéliens localement compacts*, Ann. Sci. Ecole Norm. Sup. (3), 64(1947), 79–99.

20. G. Choquet, *Deux exemples classiques de représentation intégrale*, Enseignement Math. (2), 15(1969), 63–75.

21. J. Chover, *On normalized entropy and the extensions of a positive-definite function*, J. Math. Mech. 10(1961), 927–945.

22. J. L. B. Cooper, *Fourier-Stieltjes integrals*, Proc. London Math. Soc. (2), 51(1950), 265–284.

23. ——, *Positive definite functions of a real variable*, Proc. London Math. Soc. (3), 10(1960), 53–66.

24. M. Crum, *On positive-definite functions*, Proc. London Math. Soc. (3), 6(1956), 548–560.

25. A. Devinatz, *On the extensions of positive definite functions*, Acta Math. 102(1959), 109–134.

26. ——, *On measurable positive definite operator functions*, J. London Math. Soc. 35(1960), 417–424.

27. R. E. Edwards, *Fourier Series: a Modern Introduction*, 2 vols., Holt, Rinehart and Winston, New York, 1967.

28. ——, *Unbounded integrally postitive definite functions*, Studia Math. 33(1969), 185–191.

29. S. J. Einhorn, *Functions positive definite in the space C*, Amer. Math. Monthly 75(1968), 393.

30. ——, *Functions positive definite in C*[0, 1], Proc. Amer. Math. Soc. 22(1969), 702–703.

31. C. G. Esseen, *Fourier analysis of distribution functions*, *Acta Math.* 77 (1945), 1–125.

32. P. L. Falb and U. Haussmann, *Bochner's Theorem in infinite dimensions*, Pacific J. Math. 43(1972), 601–618.

33. K. Fan, *On positive definite sequences*, Ann. of Math. 47(1946), 593–607.

34. ——, *Les Fonctions Définies-Positives et les Fonctions Complètement Monotones*, Gauthier-Villars, Paris, 1950.

35. E. Fisher, *Über das Carathéodory'sche Problem*, Potenzreihen mit positivem reelen Teil betreffend, Rend. Circ. Mat. Palermo 32(1911), 240–256.

36. J. Fourier, *Théorie Analytique de la Chaleur*, Discours Préliminaire, Didot, Paris, 1822, p. xxiii.

37. G. Frobenius, *Ableitung eines Satzes von Carathéodory aus einer Formel von Kronecker*, Sitzungsberichte Preuss. Akad. Wiss., (1912), 4–15.

38. I. M. Gelfand and D. A. Raikov, *Irreducible unitary representations of locally bicompact groups* (in Russian), Mat. Sbornik N.S., 13(1943), 301–316. English translation: Amer. Math. Soc. Translations Ser. 2, 36(1964), 1–15.

39. I. M. Gelfand and N. Vilenkin, *Generalized Functions*, Vol. 4, Applications of Harmonic Analysis, Academic Press, New York, 1964.

40. B. V. Gnedenko, *Sur une propriété caractéristique des lois infiniment divisibles*, Bull. Univ. d'Etat Moscou, Ser. Internat. Sec. A, 1, no. 5, (1937), 9–15.

41. R. Godement, *Sur une généralisation d'un théorème de Stone*, C. R. Acad. Sci. Paris, 218(1944), 901–903.

42. ——, *Les fonctions de type positif et la théorie des groupes*, Trans. Amer. Math. Soc. 83(1948), 1–84.

43. V. Gorbachuk, *On integral representations of Hermitian-indefinite forms (the case of several variables)* (in Russian), Ukrain, Mat. Ž. 16(1964), 232–236.

44. U. Grenander and G. Szegö, *Toeplitz Forms and their Applications*, Univ. of Calif. Press, Berkeley, 1958.

45. L. Gross, *Harmonic Analysis on Hilbert Space*, Memoirs Amer. Math. Soc. **46**, Providence, R.I., 1963.

46. G. Herglotz, *Über Potenzreihen mit positivem, reelen Teil im Einheitskreis*, Leipziger Berichte, math.-phys. Kl. **63**(1911), 501–511.

47. E. Hewitt, *A survey of abstract harmonic analysis, Some Aspects of Analysis and Probability*, pp. 101–168. Surveys in Applied Mathematics, Vol. 4, Wiley, New York, 1958.

48. E. Hewitt and K. A. Ross, *Integrally positive-definite functions on groups*, Studia Math. **31**(1968), 145–151.

49. ——, *Abstract Harmonic Analysis* Vol. II, Springer-Verlag, New York, 1970.

50. D. Hilbert, *Grundzüge einer allgemeinen Theorie der linearen Integralgleichungen* I, Gött. Nachrichten, math.-phys. Kl. (1904), 49–91.

51. E. W. Hobson, *Presidential address to Brit. Ass. for the Advancement of Science*, Sec. A, Nature **84**(1910), 284–291.

52. I. Iohvidov and M. Krein, *Spectral theory of operators in spaces with an indefinite metric* II (in Russian), Trudy Moskov. Mat. Obšč. **8**(1959), 413–496. English translation: Amer. Math. Soc. Translations Ser. 2, **34**(1963), 283–373.

53. S. Ito, *Positive definite functions on homogeneous spaces*, Proc. Japan Acad. **26**(1950), 17–28.

54. Y. Katznelson, *An Introduction to Harmonic Analysis*, Wiley, New York, 1968.

55. A. Khintchine, *Korrelationstheorie der stationaren stochastischen Prozesse*, Math. Ann. **109**(1934), 604–615.

56. A. Kolmogorov, *Sulla forma generale di un processo stocastico omogeneo*, Atti Acad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Nat. (6), **15**(1932), 805–808, 866–869.

57. M. Krein, *Sur le problème du prolongement des fonctions hermitiennes positives et continues*, Dokl. Akad. Nauk SSSR **26**(1940), 17–22.

58. ——, *On a general method for decomposing positive-definite kernels into elementary products* (in Russian), Dokl. Akad. Nauk SSSR **53**(1946), 3–6.

59. ——, *Screw lines in infinite-dimensional Lobachevski space and the Lorentz transformation* (in Russian), Ups. Mat. Nauk **3**(1948), 158–160.

60. ——, *Hermitian-positive kernels on homogeneous spaces* I *and* II (in Russian), Ukrain. Mat. Ž. **1**(1949), 64–98 and **2**(1950), 10–59. English translation: Amer. Math. Soc. Translations Ser. 2, **34**(1963), 69–164.

61. ——, *Integral representation of a continuous Hermitian-indefinite function with a finite number of negative squares* (in Russian), Dokl. Akad. Nauk SSSR **125**(1959), 31–34.

62. K. de Leeuw and I. Glicksberg, *The decomposition of certain group representations*, J. Analyse Math. **15**(1965), 135–192.

63. P. Lévy, *Calcul des Probabilités*, Gauthier-Villars, Paris, 1925.

64. ——, *Théorie de l'Addition des Variables Aléatoires*, Gauthier-Villars, Paris, 1937.

65. ——, *Quelques problèmes non résolus de la théorie des fonctions caractéristiques*, Annali Mat. Pura e Appl. (4), **53**(1961), 315–332.

66. M. Loève, *Probability Theory*, 2nd ed., Van Nostrand, Princeton, N.J., 1960.

67. L. Loomis, *An Introduction to Abstract Harmonic Analysis*, Van Nostrand, Princeton, N.J., 1953.

68. G. Lumer, *Bochner's Theorem, states, and the Fourier transform of measures,* Studia Math. **46**(1973), 135–140.

69. M. Mathias, *Über positive Fourier-Integrale,* Math. Zeit, **16**(1923), 103–125.

70. K. McKennon, *Multipliers, Positive Functionals, Positive-Definite Functions, and Fourier-Stieltjes Transforms,* Memoirs Amer. Math. Soc. **111**, Providence, R.I., 1971.

71. J. McMullen, *Extensions of Positive-Definite Functions,* Memoirs Amer. Math. Soc. **117**, Providence, R.I., 1972.

72. J. Mercer, *Functions of positive and negative type and their connection with the theory of integral equations,* Philos. Trans. Roy. Soc. London, Ser. A, **209**(1909), 415–446.

73. ——, *Linear transformations and functions of positive type,* Proc. Roy. Soc. London, Ser. A, **99**(1921), 19–38.

74. R. A. Minlos, *Generalized random processes and their extension to a measure* (in Russian), Trudy Moskov. Mat. Obšč. **8**(1959), 497–518. English translation: Selected Translations in Math. Stat. and Prob. **3**(1963), 291–313.

75. E. H. Moore, *On properly positive Hermitian matrices,* Bull. Amer. Math. Soc. **23**(1916), 59, 66–67.

76. ——, *General Analysis, Part* I, Memoirs Amer. Philos. Soc. **1**, Philadelphia, 1935.

77. M. A. Naimark, *Positive definite operator functions on a commutative group* (in Russian, English summary), Izv. Akad. Nauk SSSR, Ser. Mat. **7**(1943), 237–244.

78. M. Nakamura and H. Umegaki, *A remark on theorems of Stone and Bochner,* Proc. Japan Acad. **27**(1951), 506–507.

79. H. Nakano, *Reduction of Bochner's Theorem to Stone's Theorem,* Ann. of Math. (2), **49**(1948), 279–280.

80. J. Von Neumann and I. J. Schoenberg, *Fourier integrals and metric geometry,* Trans Amer. Math. Soc. **50**(1941), 226–251.

81. A. E. Nussbaum, *Integral representation of functions and distributions* positive definite relative to the orthogonal group, Trans. Amer. Math. Soc. **175** (1973), 355–387.

82. R. Phelps, *Lectures on Choquet's Theorem,* Van Nostrand, Princeton, N.J., 1966.

83. V. Plyuschcheva (Gorbachuk), *On integral representations of continuous Hermitian-indefinite forms* (in Russian), Dokl. Akad. Nauk SSSR, **145**(1962), 534–537.

84. G. Pólya, *Remarks on characteristic functions,* Proc. First Berkeley Conf. on Math. Stat. and Prob., pp. 115–123, Univ. of Calif. Press, Berkeley, 1949.

85. A. Povzner, *Über positive Funktionen auf einer Abelschen Gruppe,* Dokl. Akad. Nauk SSSR **28**(1940), 294–295.

86. D. A. Raikov, *Positive definite functions on commutative groups with an invariant measure,* Dokl. Adad. Nauk SSSR **28**(1940), 296–300.

87. ——, *Sur les fonctions positivement définies,* Dokl. Akad. Nauk SSSR **26**(1940), 860–865.

88. ——, *On various types of convergence of positive definite functions* (in Russian), Dokl. Akad. Nauk SSSR **58**(1947), 1279–1282.

89. N. W. Rickert, *A note on positive definite functions,* Studia Math. **36** (1970), 223–225.

90. F. Riesz, *Sur certains systèmes singuliers d'équations intégrales*, Ann. Sci. Ecole Norm. Sup. **28**(1911), 33–62.

91. ——, *Über Sätze von Stone und Bochner*, Acta Sci. Math. Szeged **6**(1933), 184–198.

92. F. Riesz and B. Sz.-Nagy, *Functional Analysis*, Ungar, New York, 1955.

93. W. Rudin, *Fourier Analysis on Groups*, Interscience New York, 1962.

94. ——, *The extension problem for positive definite functions*, Ill. J. Math. **7**(1963), 532–539.

95. ——, *An extension theorem for positive-definite functions*, Duke Math. J. **37**(1970), 49–53.

96. V. Sazonov, *A note on characteristic functionals*, Theory Prob. and Appl. **3**(1958), 188–192.

97. I. J. Schoenberg, *Metric spaces and positive definite functions*, Trans. Amer. Math. Soc. **44**(1938), 522–536.

98. ——, *Metric spaces and completely monotone functions*, Ann. of Math. **39**(1938), 811–841.

99. ——, *Positive definite functions on spheres*, Duke Math. J. **9**(1942), 96–108.

100. B. M. Schreiber, T. C. Sun and A. T. Barucha-Reid, *Algebraic models for probability measures associated with stochastic processes*, Trans. Amer. Math. Soc. **158**(1971), 93–105.

101. I. Schur, *Bemerkungen zur Theorie der beschränkten Bilinearformen mit unendlich vielen Veränderlichen*, J. Reine Angew. Math. **140**(1911), 1–28.

102. ——, *Über einen Satz von C. Carathéodory*, Sitzungsberichte Preuss. Akad. Wiss. (1912), 4–15.

103. L. Schwartz, *Théorie des Distributions*, Vol. 2, Actualités Sci. Indust., no. 1122, Hermann, Paris, 1951.

104. I. Segal and J. von Neumann, *A theorem on the representation of semi-simple Lie groups*, Ann. of Math. (2), **52**(1950), 509–517.

105. T.-S. Shah, *On conditionally positive-definite generalized functions*, Scientia Sinica **11**(1962), 1147–1168.

106. ——, *Positive definite functions on commutative topological groups*, Scientia Sinica **14**(1965), 653–665.

107. J. Stewart, *Unbounded positive definite functions*, Canad. J. Math. **21** (1969), 1309–1318.

108. ——, *Counterexample to a conjecture on positive definite functions*, Canad. J. Math. **24**(1972), 926–929.

109. ——, *Functions with a finite number of negative squares*, Canad. Math. Bull. **15**(1972), 399–410.

110. M. Stone, *Linear transformations in Hilbert space* III, Proc. Nat. Acad. Sci. U.S.A. **16**(1930), 172–175.

111. ——, *On one-parameter unitary groups in Hilbert space*, Ann. of Math. (2), **33**(1932), 643–648.

112. B. Sz.-Nagy, *Trnasformations de l'espace de Hilbert, fonctions de type positif sur un groupe*, Acta Sci. Math. Szeged **15** (1954), 104–114.

113. L. M. Tang, *Symmetrically positive definite functions*, M.Sc. Thesis, McMaster University, Hamilton, Ontario, 1974.

114. E. C. Titchmarsh, *Introduction to the Theory of Fourier Integrals*, Oxford University Press, 1937.

115. O. Toeplitz, *Über die Fourier'sche Entwickelung positiver Funktionen*, Rend. Circ. Mat. Palermo **32**(1911), 191–192.

116. A. Wawrzyńczyk, *On tempered distributions and Bochner-Schwartz Theorem on arbitrary locally compact abelian groups*, Colloq. Math. **19**(1968), 305–318.

117. W. von Waldenfels, *Positiv definite Funktionen auf einem unendlich dimensionalen Vektorraum*, Studia Math., **30**(1968), 153–162.

118. A. Weil, *L'Intégration dans les Groupes Topologiques et ses Applications*, Actualités Sci. Indust., no. 869, Hermann, Paris, 1940.

119. H. Weyl, *Harmonics on homogeneous manifolds*, Ann. of Math. (2), **35** (1934), 486–499.

120. N. Wiener, *Generalized harmonic analysis*, Acta Math. **55**(1930), 117–258.

121. ——, *The Fourier Integral and Certain of its Applications*, Cambridge Univ. Press, 1933.

122. A. M. Yaglom, *Positive-definite functions and homogeneous random fields on groups and homogeneous spaces*, Societ Math. Dokl. 1(1960), 1402–1405.

123. H. Yoshizawa, *On some types of convergence of positive definite functions*, Osaka Math. J. 1(1949), 90–94.

124. W. H. Young, *A note on a class of symmetric functions and on a theorem required in the theory of integral equations*, Philos. Trans. Roy. Soc. London, Ser. A, **209**(1909), 415–446.

125. A. Zygmund, *Trigonometrical Series*, Monografje Matematyczne, Warsaw, 1935.

Department of Mathematics, McMaster University, Hamilton, Ontario.