PUBLICATIONS MATHÉMATIQUES ET INFORMATIQUES DE RENNES

M. S. BINGHAM

K. R. PARTHASARATHY

A probabilistic Proof of Bochner's Theorem on positive definite Functions

Publications des séminaires de mathématiques et informatique de Rennes, 1966-1967
« Séminaires de probabilités et statistiques », , exp. n° 1, p. 1-13

<http://www.numdam.org/item?id=PSMIR_1966-1967____A1_0>

© Département de mathématiques et informatique, université de Rennes, 1966-1967, tous droits réservés.

L'accès aux archives de la série « Publications mathématiques et informatiques de Rennes » implique l'accord avec les conditions générales d'utilisation (http://www.numdam.org/conditions). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques
http://www.numdam.org/

A PROBABILISTIC PROOF OF BOCHNER'S THEOREM ON POSITIVE DEFINITE FUNCTIONS

by

M. S. Bingham and K. R. Parthasarathy
University of Sheffield

1. Introduction

The purpose of this paper is to give a proof of Bochner's theorem on positive definite functions in a locally compact second countable abelian group by a purely probabilistic method without using the theory of Banach algebras or general theory of representations. We reduce the problem to the Daniel-Kolmogorov construction of a stochastic process on the basis of finite dimensional distributions. The [illegible]-Cramer continuity theorem is deduced as a corollary.

2. Preliminaries

Let X be a locally compact second countable abelian group and Y its character group. Y consists of all continuous homomorphisms from X into the multiplicative circle group K of complex numbers with modulus unity. We shall denote the Haar measure on Y by H and use dy to denote integration with respect to H. For any $x \in X$, $y \in Y$, we shall denote by $\langle x,y \rangle$ the value of the character y at x. We assume the fact that X and Y are character groups of each other (cf. Rudin [2]).

A complex valued function φ defined on Y is said to be positive definite if for all $y_1, y_2, \ldots, y_k \in Y$, complex numbers $c_1, c_2, \ldots, c_k$ and positive integers k, the inequality

$$\sum_{1 \leq i,j \leq k} c_i \bar{c}_j \varphi(y_i - y_j) \geq 0 \tag{2.1}$$

is satisfied. With these notations we have the following theorem due to Bochner [2].

Theorem 2.1 Let φ be a complex-valued positive definite function defined on Y such that φ is continuous at e and $\varphi(e) = 1$, e being the identity element of Y. Then there exists a unique probability measure μ defined on the Borel σ-field of X such that

$$\varphi(y) = \int_X \langle x, y \rangle \, d\mu(x), \quad y \in Y$$

3. The case when X is a finite-dimensional torus

In this section we shall consider the case when $X = K^r$ where K is the multiplicative circle group consisting of all complex numbers of modulus unity and K^r is the r-fold cartesian product of K. Then $Y = I^r$, where I is the additive group of all integers. Any point $x \in X$ can be represented by $\underline{\theta} = (\theta_1, \theta_2, \ldots, \theta_r)$ where $0 \leq \theta_j < 2\pi$, $j = 1, 2, \ldots, r$. With this representation the group operation becomes coordinatewise addition modulo 2π. Any point $y \in Y$ can be represented by $\underline{n} = (n_1, n_2, \ldots, n_r)$ where all the n_j's are integers. Then $\langle \underline{\theta}, \underline{n} \rangle = \exp i \sum_{j=1}^{r} n_j \theta_j$. Let φ be a positive

definite function on I^r and $\varphi(0) = 1$. φ is automatically continuous because I^r is discrete. Even though an elementary proof of Theorem 2.1 is well known in this case, we reproduce it for the sake of completeness.

Consider the measure μ_N which is absolutely continuous with respect to the uniform distribution on K^r and which has the density function

$$f_N(\theta) = \frac{1}{N^r} \sum_{\substack{1 \le m_j \le N \\ 1 \le n_j \le N \\ j = 1, \dots, r}} \varphi(\underset{\sim}{m} - \underset{\sim}{n}) \langle \underset{\sim}{\theta}, \underset{\sim}{n} - \underset{\sim}{m} \rangle$$

Since $\langle \underset{\sim}{\theta}, \underset{\sim}{n} - \underset{\sim}{m} \rangle = \langle \underset{\sim}{\theta}, \underset{\sim}{n} \rangle \overline{\langle \underset{\sim}{\theta}, \underset{\sim}{m} \rangle}$, it follows from (2.1) that $f_N(\underset{\sim}{\theta}) \ge 0$ for all $\underset{\sim}{\theta}$. Trivially

$$\frac{1}{(2\pi)^r} \int_0^{2\pi} \cdots \int_0^{2\pi} f_N(\underset{\sim}{\theta}) \, d\theta_1 \, d\theta_2 \cdots d\theta_r = 1$$

An easy computation shows that, for $\underset{\sim}{j} = (j_1, j_2, \dots, j_r)$, we have

$$\frac{1}{(2\pi)^r} \int f_N(\theta) \langle \underset{\sim}{\theta}, \underset{\sim}{j} \rangle \, d\underset{\sim}{\theta} = \varphi(\underset{\sim}{j}) \prod_{k=1}^{r} \left(1 - \frac{|j_k|}{N}\right), \quad \text{if } |j_k| < N, \; k = 1, 2, \dots, r.$$

$$= 0 \quad \text{otherwise}$$

Since any set of probability measures on K^r is weakly conditionally compact, $\{\mu_N\}$ has a convergent subsequence. Let μ be any limit of $\{\mu_N\}$. Then

$$\int \langle \underset{\sim}{\theta}, \underset{\sim}{j} \rangle d\mu = \lim_N \int \langle \underset{\sim}{\theta}, \underset{\sim}{j} \rangle d\mu_N = \varphi(\underset{\sim}{j})$$

for all $\underset{\sim}{j}$. This establishes the existence of a μ with the required property. Uniqueness follows from the well known fact that any continuous function on K^r is a uniform limit of linear combinations of functions of the form $\langle \underset{\sim}{\theta}, \underset{\sim}{j} \rangle$. This completes the proof.

Before proceeding to the proof of Theorem 2.1 in the general case we shall mention a simple corollary. To this end we need some notations. For any two abstract sets A and B, we shall denote by A^B the set of all functions defined on B and taking values in A. Then K^Y is the space of all functions defined on Y and taking values in K. Any element $\underset{\sim}{\theta} \in K^Y$ can be represented as a function $\underset{\sim}{\theta}(y)$ where $0 \leq \underset{\sim}{\theta}(y) < 2\pi$ for all $y \in Y$. Let π_y denote the projection map $\underset{\sim}{\theta} \longrightarrow \underset{\sim}{\theta}(y)$ from K^Y onto K. Let $\underset{\sim}{\mathcal{B}}$ denote the smallest σ-field of subsets of K^Y with respect to which all the π_y are measurable. We shall denote by $\mathcal{N}^Y$ the character group of K^Y. $\mathcal{N}^Y$ is the set of all integer valued functions $\underset{\sim}{n}(y)$ defined on Y and vanishing outside a finite subset. We shall denote by $\underset{\sim}{0}$ the function which is identically 0. With these notations we have the following corollary.

Corollary 3.1 Let $\psi(\underset{\sim}{n})$, $\underset{\sim}{n} \in \mathcal{N}^Y$ be any positive definite function defined on $\mathcal{N}^Y$ and $\psi(\underset{\sim}{0}) = 1$. Then there exists a unique probability measure

P on $\mathcal{S}$ such that

$$\int \left[\exp i \sum_{y \in Y} \underset{\sim}{n}(y)\, \underset{\sim}{\theta}(y) \right] dP(\underset{\sim}{\theta}) = \psi(\underset{\sim}{n})$$

for all $\underset{\sim}{n} \in \mathcal{N}$.

Proof For any finite subset $F \subseteq Y$, let

$$\mathcal{N}_F = \left\{ \underset{\sim}{n} : \underset{\sim}{n} \in \mathcal{N},\ \underset{\sim}{n}(y) = 0 \text{ if } y \notin F \right\}$$

When ψ is restricted to $\mathcal{N}_F$ we obtain a positive definite function on I^F. Hence by the validity of Theorem 2.1 for K^F, we deduce the existence of a measure P_F on the Borel σ-field of K^F, satisfying the property

$$\int_{K^F} \left[\exp\left(i \sum_{y \in F} \underset{\sim}{n}(y)\, \underset{\sim}{\theta}(y) \right) \right] dP_F(\underset{\sim}{\theta}) = \psi(\underset{\sim}{n}),\ \underset{\sim}{n} \in \mathcal{N}_F.$$

It follows from the uniqueness of P_F for every F that if $F_1 \supseteq F_2$ are two finite subsets of Y, the measure P_{F_1} induces P_{F_2} through the natural projection from K^{F_1} onto K^{F_2} obtained by dropping the yth coordinate whenever $y \notin F_2$. In other words the family of measures $\left\{ P_F : F \subseteq Y,\ F \text{ finite} \right\}$ is consistent. Hence, by the Daniel-Kolmogorov theorem [1] there exists a unique probability measure P on $\mathcal{S}$ with the required properties.

4. The general case

Before proceeding to the proof of Theorem 2.1 we need a simple lemma.

Lemma 4.1 Let $f(y)$ be a complex-valued Haar-measurable function satisfying the following conditions:

(1) $|f(y)| = 1$ for all $y \in Y$

(2) $f(y) \cdot f(y') = f(y+y')$ a.e. $y, y' \ (H \times H)$.

Then there exists a unique $x \in X$ such that

$$f(y) = \langle x, y \rangle \quad \text{a.e. } y \ (H).$$

Proof Consider the set function

$$\mu(C) = \int_C f(y)\, dy$$

for all Haar-measurable sets contained in a compact set. If $C + y'$ denotes the translate of C by y', we have from the invariance of Haar measure under translations,

$$\begin{aligned} \mu(C+y') &= \int_{C+y'} f(y)\, dy \\ &= \int_C f(y+y')\, dy \\ &= f(y')\mu(C) \quad \text{a.e. } y' \ (H). \end{aligned}$$

Choose and fix a compact set C such that $\mu(C) \neq 0$. Then

$$f(y) = \frac{\mu(C+y)}{\mu(C)} \quad \text{a.e. } y \ (H).$$

We shall now prove that $\mu(C + y)$ is continuous in y. Indeed, we have from condition (1) of the lemma

$$\begin{aligned} &|\mu(C+y_1) - \mu(C+y_2)| \\ &= \left| \int \chi_{C+y_1}(y) f(y)\,dy - \int \chi_{C+y_2}(y) f(y)\,dy \right| \\ &\leq \int |\chi_{C+y_1}(y) - \chi_{C+y_2}(y)|\,dy \\ &= \int |\chi_{C+y_1-y_2}(y) - \chi_C(y)|\,dy \end{aligned}$$

where χ_C denotes the indicator function of C. The last term on the right hand side of the above inequalities is simply the Haar measure of $(C + y_1 - y_2) \Delta C$ which tends to zero as $y_1 - y_2$ approaches the identity.

Thus $f(y)$ is almost everywhere equal to a continuous function $g(y)$. It follows from the conditions of the lemma that $|g(y)| = 1$ for all $y \in Y$ and $g(y).g(y') = g(y + y')$ for all $y, y' \in Y$. In other words g is a character on Y. Hence, by the duality theorem, there is a unique point $x \in X$ such that $g(y) = \langle x, y \rangle$ for all y. This completes the proof of the lemma.

Proof of Theorem 2.1 Adopting the notations described before the statement of corollary 3.1, we introduce the function ψ defined on $\mathcal{N}^p$ by

$$\psi(\underset{\sim}{n}) = \varphi\left(\sum_{y\in Y} \underset{\sim}{n}(y)\, y\right)$$

The positive definiteness of φ in Y implies the positive definiteness of ψ in $\mathcal{N}$. Further $\psi(\underset{\sim}{0}) = \varphi(e) = 1$. Thus, by corollary 3.1, there exists a probability measure P on $\mathcal{S}$ such that

$$\int_{K^Y} \exp\left[\sqrt{-1} \sum_{y\in Y} \underset{\sim}{n}(y)\, \underset{\sim}{\theta}(y)\right] dP(\underset{\sim}{\theta}) = \psi(\underset{\sim}{n}) \,, \quad \underset{\sim}{n} \in \mathcal{N}$$

Consider the probability space $(K^Y, \overline{\mathcal{S}}, P)$ where $\overline{\mathcal{S}}$ denotes the P-completion of $\mathcal{S}$. Treating the elements of Y as a time variable, define the stochastic process

$$z(y, \underset{\sim}{\theta}) = \exp \sqrt{-1}\, \underset{\sim}{\theta}(y) \,, \quad y \in Y \,, \ \underset{\sim}{\theta} \in K^Y \,.$$

for all $y_1, y_2, \ldots, y_k \in Y$, integers $n_1, n_2, \ldots, n_k$ and positive integers k, we have

$$(4.1) \qquad E \prod_{j=1}^{k} z(y_j, \underset{\sim}{\theta})^{n_j} = \varphi\left(\sum_{j=1}^{k} n_j y_j\right)$$

where E denotes expectation with respect to P. In particular,

$$(4.2) \qquad E\, z(y, \underset{\sim}{\theta})\, z(y', \underset{\sim}{\theta})\, z(y+y', \underset{\sim}{\theta})^{-1} = 1 \,, \quad y, y' \in Y \,.$$

Since the random variable within the expectation sign above is of modulus unity, it follows that

$$z(y+y',\underset{\sim}{\theta}) = z(y,\underset{\sim}{\theta})\,z(y',\underset{\sim}{\theta}) \quad \text{a.e. } \underset{\sim}{\theta}\ (P) \tag{4.3}$$

for every pair $y, y' \in Y$. We also have

$$E\, z(y,\underset{\sim}{\theta}) = \varphi(y) \quad , \ y \in Y \ . \tag{4.4}$$

From (4.1) an easy computation gives

$$E\left| z(y,\underset{\sim}{\theta}) - z(y',\underset{\sim}{\theta}) \right|^2 = 2 - \varphi(y-y') - \varphi(y'-y) \tag{4.5}$$

Since $\varphi(e) = 1$ and φ is continuous at the identity, the above equation shows that $z(y,\underset{\sim}{\theta})$ is stochastically continuous in y. Therefore, by a theorem similar to Theorem 2.6, page 61 of Doob [1], we may assume that $z(y,\underset{\sim}{\theta})$ is $H \times \mathcal{B}$-measurable. By applying Fubini's theorem to the function $\left| z(y,\underset{\sim}{\theta})\, z(y',\underset{\sim}{\theta}) - z(y+y',\underset{\sim}{\theta}) \right|$ in the space $C \times C \times K^Y$ for every compact $C \subseteq Y$, we conclude that there exists a set $A \subseteq K^Y$ such that $P(A) = 1$, and for every $\underset{\sim}{\theta} \in A$,

$$z(y,\underset{\sim}{\theta})\, z(y',\underset{\sim}{\theta}) = z(y+y',\underset{\sim}{\theta}) \quad \text{a.e. } y,y' \ (H \times H)$$

It now follows from lemma 4.1 that there exists an element $x(\underset{\sim}{\theta}) \in X$ such that

$$\langle x(\underset{\sim}{\theta}), y \rangle = z(y,\underset{\sim}{\theta}) \quad \text{a.e. } y\ (H) \tag{4.6}$$

for every $\underset{\sim}{\theta} \in \Lambda$. Define $x(\underset{\sim}{\theta}) = e$, the identity of X, for every $\theta \notin \Lambda$. Since $\langle x(\underset{\sim}{\theta}), y \rangle$ is continuous in y for every $\underset{\sim}{\theta}$, it follows that $\langle x(\underset{\sim}{\theta}), y \rangle$ is P-measurable in $\underset{\sim}{\theta}$ for every $y \in Y$. Since the smallest σ-field with respect to which all the characters are measurable, is the Borel σ-field in X, it follows that $x(\underset{\sim}{\theta})$ is an X-valued random variable defined on $(K^Y, \overline{\mathcal{E}}, P)$. Let μ be the distribution of $x(\underset{\sim}{\theta})$ in X.

Using Fubini's theorem, we deduce from equation (4.6) that

$$\int \langle x(\underset{\sim}{\theta}), y \rangle \, dP(\theta) = \int z(y, \underset{\sim}{\theta}) \, dP(\underset{\sim}{\theta}) \quad \text{a.e. } y \ (H)$$

Since the left hand integral in the above equation is simply $\int \langle x, y \rangle \, d\mu(x)$, equation (4.4) gives

$$\int \langle x, y \rangle \, d\mu(x) = \varphi(y) \quad \text{a.e. } y \ (H) . \tag{4.7}$$

Further, (4.4) and (4.5) imply the continuity of φ at all points of Y. Hence we have

$$\int \langle x, y \rangle \, d\mu(x) = \varphi(y) , \quad y \in Y . \tag{4.8}$$

In order to prove the uniqueness of μ, choose and fix a dense subset $\{ y_k \}$, $k = 1, 2, \ldots$ of Y. Consider the map

$$T : x \to \left(\langle x, y_1 \rangle, \langle x, y_2 \rangle, \ldots \ldots \right)$$

of X into the space K^∞ which is a countable product of K. Since characters

are continuous and separate points, it follows that τ is a one-one and continuous map of X into K^{∞}. Since X is a countable union of compact sets, it follows by considering the restriction of τ to these compact sets that τ^{-1} is measurable. In other words τ is a Borel isomorphism between X and the σ-compact set $\tau(X)$. If μ and ν are two measures satisfying (4.8), then $\mu\tau^{-1}$ and $\nu\tau^{-1}$ are measures in K^{∞} with the same finite dimensional distributions. Further $\mu\tau^{-1}$ and $\nu\tau^{-1}$ are concentrated in $\tau(X)$. Therefore $\mu = \nu$. This completes the proof of Theorem 2.1.

5. The continuity theorem

For any probability measure μ defined on the Borel σ-field of X, the function φ defined on Y by the equation

$$\varphi(y) = \int \langle x, y \rangle \, d\mu(x)$$

is called the characteristic function of μ. It is easy to show that φ is continuous and positive definite. Bochner's theorem asserts that every continuous positive definite function arises in this way. Further the correspondence $\mu \longleftrightarrow \varphi$ is one-one. The uniform continuity of the characters implies the uniform continuity of φ. Further if a sequence of probability measures μ_n converges weakly to μ, then the corresponding characteristic functions φ_n converge uniformly over compact sets to the characteristic function φ of μ. The following theorem asserts a much stronger version of the converse.

Theorem 5.1 Let $\{\mu_n\}$, $n = 1, 2, \ldots$ be a sequence of probability measures defined on the Borel σ-field of X and $\{\varphi_n\}$, $n = 1, 2, \ldots$ be the corresponding characteristic functions. If φ_n converges pointwise to a function φ which is continuous at the identity of Σ, then φ is the characteristic function of a probability measure μ and μ_n converges weakly to μ.

Proof First of all it is obvious that φ is a positive definite function on Σ and $\varphi(e) = 1$. Hence by Theorem 2.1 there exists a measure μ with φ as its characteristic function.

In order to prove the weak convergence consider the map τ described in the last paragraph of section 4. τ maps X into the space K^∞. It is clear that the characteristic function of $\mu_n \tau^{-1}$ converges to the characteristic function of $\mu \tau^{-1}$. Since the set of all probability measures in K^∞ is weakly compact, it follows that $\mu_n \tau^{-1}$ converges weakly to $\mu \tau^{-1}$ in K^∞.

Choose and fix a sequence $\{C_j\}$ of pairwise disjoint Borel subsets of X whose union is X and such that the boundary of each C_j has μ-measure 0 and $\overline{C}_j$ is compact. Then for any closed set $F \subseteq X$, we have

$$\overline{\lim_{n\to\infty}}\ \mu_n(F) = \overline{\lim_{n\to\infty}} \sum_{j=1}^{\infty} \mu_n(F \cap C_j) \tag{5.1}$$

$$\leq \sum_{j=1}^{\infty} \overline{\lim_{n\to\infty}}\ \mu_n(F \cap \overline{C}_j) .$$

Since τ is a Borel isomorphism which maps compact subsets of X onto closed subsets of X^∞ and $\mu_n \tau^{-1}$ converges weakly to $\mu \tau^{-1}$, we have

$$\begin{aligned}(5.2)\qquad \overline{\lim_{n\to\infty}}\ \mu_n(F\cap C_j) &= \overline{\lim_{n\to\infty}}\ \mu_n \tau^{-1}\left[\tau(F\cap C_j)\right] \\ &\le \mu\tau^{-1}\left[\tau(F\cap C_j)\right] \\ &= \mu(F\cap \bar{C}_j) = \mu(F\cap C_j)\end{aligned}$$

Combining (5.1) and (5.2) we have

$$\overline{\lim_{n\to\infty}}\ \mu_n(F) \le \mu(F).$$

Since F is an arbitrary closed set, it follows that μ_n converges weakly to μ. This completes the proof.

REFERENCES

[1] J. L. Doob, Stochastic Processes. John Wiley & Sons, New York, 1953.

[2] W. Rudin, Fourier Analysis on Groups. Interscience Pub. Co., New York, 1962.