Positive Definite Functions on Locally Compact Groups

Ikhan Choi

March 21, 2022

Contents

Abstract

Hi, I'm abstract.

1 Introduction

Definition 1.1. Let G be a group. A function $f : G \to \mathbb{C}$ is called *positive definite* if for each positive integer n a non-negativity condition

$$\sum_{k,l=1}^{n} f(x_l^{-1}x_k)\xi_k\overline{\xi}_l \geq 0$$

is satisfied for every n-tuple $(x_1, \cdots, x_n) \in G^n$ and every vector $(\xi_1, \cdots, \xi_n) \in \mathbb{C}^n$.

A function f is positive definite if and only if bilinear forms defined by matrices $(f(x_i x_j^{-1}))_{i,j=1}^n$ for each positive integer n are hermitian, and even more, positive *semi*-definite, regardless of any choices of $(x_1, \cdots, x_n) \in G^n$. Positive definite functions have several remarkable properties as follows:

Proposition 1.1 (Algebraic properties)**.** *Let G be a group. For the set of all positive definite functions, the following properties hold:*

(a) *It is closed under complex conjugation. Furthermore, $\overline{f(x)} = f(x^{-1})$.*

(b) *It is closed under positive scalar multiplication.*

(c) *It is closed under summation.*

(d) *It is closed under product.*

Proof. (a) Let $n = 1$ and $\xi_1 \neq 0$. Then, $0 \leq f(e)|\xi_1|^2$ implies $f(e) \in \mathbb{R}$.
Let $n = 2$ with $x_1 = e$ and $x_2 = x$, and let $\xi_2 = 1$. Then,

$$\begin{aligned} 0 &\leq f(e)|\xi_1|^2 + f(x^{-1})\xi_1\overline{\xi}_2 + f(x)\xi_2\overline{\xi}_1 + f(e)|\xi_2|^2 \\ &= f(e)(1 + |\xi_1|^2) + f(x^{-1})\xi_1 + f(x)\overline{\xi}_1, \end{aligned}$$

so

$$\begin{aligned} 0 &= \mathrm{Im}(f(x^{-1})\xi_1 + f(x)\overline{\xi}_1) \\ &= (\mathrm{Re}\, f(x^{-1}) - \mathrm{Re}\, f(x))\,\mathrm{Im}\,\xi_1 + (\mathrm{Im}\, f(x^{-1}) + \mathrm{Im}\, f(x))\,\mathrm{Re}\,\xi_1 \end{aligned}$$

for all $\xi_1 \in \mathbb{C}$. Therefore, $\overline{f}(x) = f(x^{-1})$.

(b) and (c) are clearly true by definition.

(d) It follows from the Schur product theorem, which states that the Hadamard product(componentwise product) of two positive semi-definite matrices is also positive semi-definite. $\square$

Proposition 1.2 (Analytic properties)**.** *Let G be a group.*

(a) *If f is positive definite, then* $\sup_{x\in G}|f(x)| \le f(e)$.

(b) *If f_n is a sequence of positive definite functions, then the pointwise limit* $\lim_{n\to\infty} f_n$ *is also positive definite.*

(c) *Let G be a locally compact group. If f_n is a sequence of positive definite functions that converges to f pointwisely and $f_n(e) = 1$, then f_n converges to f compactly.*

(d) *Let G be a locally compact group. If f is positive definite and continuous at the e, then it is both-sided uniformly continuous. (It holds for $G = \mathbb{R}$, but I have not checked for arbitrary G. I suspect it holds.)*

Proof. (a) Let $n = 2$ with $x_1 = e$ and $x_2 = x$, and let $|\xi_1| = |\xi_2| = 1$. Then,

$$0 \le f(e)|\xi_1|^2 + f(x^{-1})\xi_1\overline{\xi}_2 + f(x)\xi_2\overline{\xi}_1 + f(e)|\xi_2|^2 = 2f(e) + 2f(x)\xi_2\overline{\xi}_1.$$

Taking ξ_1 and ξ_2 such that $\xi_2\overline{\xi}_1$ has the same argument with $\overline{f}(x)$, we obtain $|f(x)| \le f(e)$.

(b) The defining property of positive definite functions is purely algebriac, so that it is preserved by pointwise limit.

(c) and (d) are too difficult to prove at this point, we will be proved later. □

This thesis follows the historical flows to extract mathematical ideas behind the positive definite functions. In particular, we are concerned with the results like the following *Bochner-type theorems*:

Theorem 1.3. *A function $c : \mathbb{Z} \to \mathbb{C}$ is positive definite if and only if there is a unique finite regular Borel measure μ on $\mathbb{T} = \mathbb{R}/2\pi\mathbb{Z}$ such that*

$$c(k) = \int_0^{2\pi} e^{-ik\theta}\, d\mu(\theta)$$

for all $k \in \mathbb{Z}$.

Theorem 1.4. *A continuous function $\varphi : \mathbb{R} \to \mathbb{C}$ is positive definite if and only if there is a unique finite regular Borel measure μ on $\mathbb{R}$ such that*

$$\varphi(t) = \int e^{itx}\, d\mu(x)$$

for all $t \in \mathbb{R}$.

They have similar forms in that they describe the necessary and sufficient conditions for a function to have a Fourier-Stieltjes integral representation of a finite regular Borel measure. One of our primary goals is to investigate the nature of positive definite functions and their harmonic-analytic relation to Borel measures within more familiar cases of $G = \mathbb{Z}$ or $\mathbb{R}$. Now then, we finally extend the Bochner-type results in the more general setting, where G is a locally compact group, and assign a new perspective of measures in terms of the representation theory of groups.

Each theorem above has its own taste in different subfields of mathematics. Theorem 1.1, which is a corollary of the celebrated Herglotz-Riesz representation theorem, is related to a classical problem in complex analysis that asks to give a characterization of a special class of analytic functions on the open unit disk $\mathbb{D}$ called the Carathéodory class. The positive definiteness arises as a property of coefficients of functions in the Caracthéodory class, and their connection to Fourier coefficients leads the complex analysis problem into harmonic analysis. In Section 2, with the methods of elementary complex variable function theory, our first Bochner-type theorem will be proved, giving a geometric description of the space of positive definite functions in addition.

In Section 3, we review the well-known results of the positive definite functions on the real line and their "weak convergence". They have been studied by probabilists, to attack the weak convergence of probaility measures. Recall that a probaility distribution of a real-valued random variable is defined by a probability measure on $\mathbb{R}$. The extended Fourier transform, but reversing the sign convention on the phase term, with respect to not only integrable functions but also finte measures, called Fourier-Stieltjes transform, of a probability measure μ is called a characteristic function of the distribution μ. In terms of probability theory, it is nothing but the function defined by the expectation $\varphi(t) := Ee^{itX}$, where X is a random variable of law μ. The Bochner theorem states that the necessary and sufficient condition for being a characteristic function is the positive definiteness and continuity.

Summaries for Section 4 and Section 5... will be here.

2 On the group $\mathbb{Z}$: complex analysis

Before the discussion of big theorems including the Carathéodory-Toeplitz theorem and the Herglotz-Riesz representation theorem, we develop a lemma as a preparation for the interplay between complex analysis and Fourier analysis.

Lemma 2.1 (Fourier series of analytic functions). *Let f be an analytic function on the open unit disk $\mathbb{D}$ with $f(0) \in \mathbb{R}$ with*

$$f(z) = c_0 + \sum_{k=1}^{\infty} 2c_k z^k,$$

the power series expansion of f at $z = 0$. Use the notation $c_{-k} := \overline{c}_k$.

(a) *For $0 \le r < 1$ and $0 \le \theta < 2\pi$, we have*

$$\operatorname{Re} f(re^{i\theta}) = \sum_{k=-\infty}^{\infty} c_k r^{|k|} e^{ik\theta}.$$

(b) *For $0 \le r < 1$, we have*

$$c_k r^{|k|} = \frac{1}{2\pi} \int_0^{2\pi} \operatorname{Re} f(re^{i\theta}) e^{-ik\theta} \, d\theta.$$

Proof. (a) Easy computation shows the identity

$$\begin{aligned}
\operatorname{Re} f(re^{i\theta}) &= \frac{1}{2}[f(re^{i\theta}) + \overline{f(re^{i\theta})}] \\
&= \frac{1}{2}\left[\left(1 + \sum_{k=1}^{\infty} 2c_k (re^{i\theta})^k\right) + \overline{\left(1 + \sum_{k=1}^{\infty} 2c_k (re^{i\theta})^k\right)}\right] \\
&= \frac{1}{2}\left[\left(1 + \sum_{k=1}^{\infty} 2c_k r^k e^{ik\theta}\right) + \left(1 + \sum_{k=1}^{\infty} 2\overline{c_k} r^k e^{-ik\theta}\right)\right] \\
&= \sum_{k=-\infty}^{\infty} c_k r^{|k|} e^{ik\theta}.
\end{aligned}$$

(b) It is clear from the uniform convergence of the series in the part (a) and the orthogonality

$$\frac{1}{2\pi i} \int_0^{2\pi} e^{-ik\theta} e^{il\theta} \, d\theta = \begin{cases} 1 & \text{if } k = l \\ 0 & \text{if } k \neq l \end{cases}.$$

□

2.1 The Carathéodory coefficient problem

The positive definiteness of functions were originally inspired by "Carathéodory coefficient problem" in complex analysis. The problem asks the condition on the power series coefficients for an analytic function defined on the open unit disk to have values of positive real part. The original paper deals with the functions analytic on a neighborhood of the closed unit disk, but the idea is extended well to the functions that has harsh behavior on the boundary.

Definition 2.1. The *Carathéodory class* is the set of all analytic functions f that map the open unit disk into the region of positive real part, with normalization condition $f(0) = 1$.

Typical examples of functions in the Carathéodory class are given by the family of functions

$$f_\theta(z) = \frac{e^{i\theta} + z}{e^{i\theta} - z} = 1 + \sum_{k=1}^{\infty} 2e^{-ik\theta} z^k$$

parametrized by $\theta \in [0, 2\pi)$. We can check they are eactly the Möbius transformations that map the unit circle to the imaginary axis satisfying $f(0) = 1$. Note the Carathéodory class is convex; if f_0 and f_1 are in the Carathéodory class, then the real part of the image of the function $f_t(z) = (1-t)f_0(z) + tf_1(z)$ is also positive for $0 < t < 1$ and $f_t(0) = (1-t) + t = 1$, so f_t also belongs to the Carathéodory class. It clearly implies that the convex combination of f_θ is also in the Carathéodory class. Carathéodory's result tells us that the converse statement also holds in a modified sense, so that f_θ can be viewed as "extreme points" in the Carathéodory class. We discuss about the extreme points later. Precisely, it is stated as follows:

Theorem 2.2 (Carathéodory)**.** *Let f be an analytic function on the open unit disk with the power series expansion*

$$f(z) = 1 + \sum_{k=1}^{\infty} 2c_k z^k.$$

Then, f belongs to the Carathéodory class if and only if the point $(c_1, \cdots, c_n) \in \mathbb{C}^n$ belongs to the convex hull of the curve $(e^{-i\theta}, \cdots, e^{-in\theta}) \in \mathbb{C}^n$ parametrized by $\theta \in [0, 2\pi)$ for each n.

Proof. ($\Leftarrow$) Denote by K_n the convex hull of the curve $\theta \mapsto (e^{-i\theta}, \cdots, e^{-in\theta}) \in \mathbb{C}^n$. Suppose first that $(c_1, \cdots, c_n) \in K_n$. For each n, there exists a finite sequence of pairs

$(\lambda_{n,j}, \theta_{n,j})_j$ having the following convex combination

$$(c_1, \cdots, c_n) = \sum_j \lambda_{n,j}(e^{-i\theta_{n,j}}, \cdots, e^{-in\theta_{n,j}})$$

with coefficients $\lambda_{n,j} \geq 0$ such that $\sum_j \lambda_{n,j} = 1$. Define

$$f_n(z) := \sum_j \lambda_{n,j} \frac{e^{i\theta_{n,j}} + z}{e^{i\theta_{n,j}} - z},$$

which has positive real part on $|z| < 1$ because $\operatorname{Re}(e^{i\theta_{n,j}}+z)/(e^{i\theta_{n,j}}-z) > 0$ for $|z| < 1$. Then,

$$\begin{aligned} f_n(z) &= \sum_j \lambda_{n,j}(1 + \sum_{k=1}^{\infty} 2e^{-ik\theta_{n,j}} z^k) \\ &= 1 + \sum_{k=1}^{n} 2c_k z^k + \sum_{k=n+1}^{\infty} \left(\sum_j 2\lambda_{n,j} e^{-ik\theta_{n,j}} \right) z^k \end{aligned}$$

implies

$$\begin{aligned} |f_n(z) - f(z)| &= \left| \sum_{k=n+1}^{\infty} \left(\sum_j 2\lambda_{n,j} e^{-ik\theta_{n,j}} \right) z^k - \sum_{k=n+1}^{\infty} 2c_k z^k \right| \\ &\leq \sum_{k=n+1}^{\infty} \left| \left(\sum_j 2\lambda_{n,j} e^{-ik\theta_{n,j}} \right) - 2c_k \right| |z|^k \\ &\leq \sum_{k=n+1}^{\infty} 4|z|^k \end{aligned}$$

converges to zero for $|z| < 1$. Therefore, f has non-negative real part on the open unit disk. The non-negativity is strengthen to the positivity by the open mapping theorem so that f belongs to the Carathéodory class.

($\Rightarrow$) Conversely, suppose that f is in the Carathéodory class. Let $(\gamma_1, \cdots, \gamma_n)$ be any point on the surface ∂K_n of K_n and S any supporting hyperplane of K_n tangent at $(\gamma_1, \cdots, \gamma_n)$. Let $(u_1, \cdots, u_n)$ be the outward unit normal vector of the supporting hyperplane S. Note that this unit normal vector is uniquely determined with respect to the induced real inner product structure on $2n$-dimensional space $\mathbb{C}^n$ described by

$$\langle (z_1, \cdots, z_n), (w_1, \cdots, w_n) \rangle = \sum_{k=1}^{n} (\operatorname{Re} z_k \operatorname{Re} w_k + \operatorname{Im} z_k \operatorname{Im} w_k) = \operatorname{Re} \sum_{k=1}^{n} z_k \overline{w}_k.$$

Then, $\sum_{k=1}^{n} |u_k|^2 = 1$ and further that the maximum

$$M := \max_{(x_1,\cdots,x_n)\in K_n} \operatorname{Re}\sum_{k=1}^{n} x_k \overline{u}_k > 0$$

is attained at $(\gamma_1, \cdots, \gamma_n)$. Our goal is to verify the bound

$$\operatorname{Re}\sum_{k=1}^{n} c_k \overline{u}_k \leq M,$$

which implies that $(c_1, \cdots, c_n)$ is contained in every half space tangent to K_n so that we finally obtain $(c_1, \cdots, c_n) \in K_n$.

Since for any $\theta \in [0, 2\pi)$ the point $(e^{-i\theta}, \cdots, e^{-in\theta})$ is in K_n so that

$$\operatorname{Re}\sum_{k=1}^{n} e^{-ik\theta} \overline{u}_k \leq M,$$

we have for arbitrarily small $\varepsilon > 0$ that

$$\operatorname{Re}\sum_{k=1}^{n} \frac{1}{r^k} e^{-ik\theta} \overline{u}_k \leq M + \varepsilon$$

for any $0 < r < 1$ sufficiently close to 1, thus we can write

$$\begin{aligned}
\operatorname{Re}\sum_{k=1}^{n} c_k \overline{u}_k &= \operatorname{Re}\sum_{k=1}^{n} \frac{1}{2\pi r^k} \int_0^{2\pi} \operatorname{Re} f(re^{i\theta}) e^{-ik\theta} \overline{u}_k \, d\theta \\
&= \frac{1}{2\pi} \int_0^{2\pi} \operatorname{Re} f(re^{i\theta}) \operatorname{Re}\sum_{k=1}^{n} \frac{1}{r^k} e^{-ik\theta} \overline{u}_k \, d\theta \\
&\leq \frac{1}{2\pi} \int_0^{2\pi} \operatorname{Re} f(re^{i\theta}) \, d\theta \cdot (M + \varepsilon) \\
&= M + \varepsilon
\end{aligned}$$

thanks to the positivity of $\operatorname{Re} f$, and by limiting $r \to 1$ from left we get the bound

$$\operatorname{Re}\sum_{k=1}^{n} c_k \overline{u}_k \leq M. \qquad \square$$

Here we introduce an infinite-dimentional description of this theorem, combining with the Herglotz representation theorem, suggests a rough but intuitional description of the space of probability measures on the unit circle $\mathbb{T}$.

Proposition 2.3. *Consider a sequence space $\mathbb{C}^{\mathbb{N}}$, endowed with the standard product topology. Then, the condition addressed in Caracthéodory's theorem is equivalent to the following: the point $(c_1, c_2, \cdots) \in \mathbb{C}^{\mathbb{N}}$ belongs to the closed convex hull of the curve $(e^{-i\theta}, e^{-i2\theta}, \cdots) \in \mathbb{C}^{\mathbb{N}}$ parametrized by $\theta \in [0, 2\pi)$.*

Furthermore, the curve $(e^{-i\theta}, e^{-i2\theta}, \cdots) \in \mathbb{C}^{\mathbb{N}}$ is the set of extreme points of its closed convex hull.

Proof. Denote by K_n the convex hull of the curve $\theta \mapsto (e^{-i\theta}, \cdots, e^{-in\theta}) \in \mathbb{C}^n$, and by K the closed convex hull of the curve $\theta \mapsto (e^{-i\theta}, e^{-i2\theta}, \cdots) \in \mathbb{C}^{\mathbb{N}}$. If we assume the Carathéodory coefficient condition is true, then since for each n we have a convex combination

$$(c_1, \cdots, c_n) = \sum_j \lambda_{n,j} (e^{-i\theta_{n,j}}, \cdots, e^{-in\theta_{n,j}})$$

with coefficients such that $\lambda_{n,j} \geq 0$ and $\sum_j \lambda_{n,j} = 1$, the sequence

$$\begin{aligned}(c_1, \cdots, c_n, &\sum_j \lambda_{n,j} e^{-i(n+1)\theta_{n,j}}, \sum_j \lambda_{n,j} e^{-i(n+2)\theta_{n,j}} \cdots) \\ &= \sum_j \lambda_{n,j} (e^{-i\theta_{n,j}}, \cdots, e^{-in\theta_{n,j}}, e^{-i(n+1)\theta_{n,j}}, e^{-i(n+2)\theta_{n,j}}, \cdots)\end{aligned}$$

is contained in and converges to the point $(c_1, c_2, \cdots)$ in the product topology as $n \to \infty$, so we are done with the desired result. For the opposite direction, let $(c_1, c_2, \cdots) \in K$. By definition of K we have an expression

$$c_k = \lim_{m \to \infty} \sum_{j=1}^{m} \lambda_{m,j} e^{-ik\theta_{m,j}}$$

with $\lambda_{m,j} \geq 0$ and $\sum_{j=1}^{m} \lambda_{m,j} = 1$, for each k. Then,

$$(c_1, \cdots, c_n) = \lim_{m \to \infty} \sum_{j=1}^{m} \lambda_{m,j} (e^{-i\theta_{m,j}}, \cdots, e^{-in\theta_{m,j}})$$

belongs to K_n because K_n is closed.

Fix $\theta \in [0, 2\pi)$ and suppose two complex sequences $(c_1, c_2, \cdots)$ and $(d_1, d_2, \cdots)$ in $\mathbb{C}^{\mathbb{N}}$ are contained in K and satisfy

$$\frac{c_k + d_k}{2} = e^{-ik\theta}$$

for all $k \in \mathbb{N}$. For each k, since all components of K are bounded by one so that $|c_k| \leq 1$ and $|d_k| \leq 1$, and since $e^{-ik\theta}$ is an extreme point of the closed unit disk $\overline{\mathbb{D}} \subset \mathbb{C}$, we have $c_k = d_k = e^{-ik\theta}$, we deduce that $(e^{-i\theta}, e^{-i2\theta}, \cdots)$ is an extreme point of K. Conversely, every extreme point of K is contained in the curve $(e^{-i\theta}, e^{-i2\theta}, \cdots)$ by Milman's "converse" theorem of the Krein-Milman theorem[citation: Phelps]. $\square$

2.2 Toeplitz's algebraic condition

Toeplitz discovered the coefficient condition addressed in the Carathéodory's paper which regards convex bodies enveloped by a curve can be equivalently described in terms of an algebraic condition that the hermitian matrices

$$H_n := (c_{k-l})_{k,l=1}^n = \begin{pmatrix} c_0 & c_1 & c_2 & \cdots & c_{n-1} \\ c_{-1} & c_0 & c_1 & \cdots & c_{n-2} \\ c_{-2} & c_{-1} & c_0 & \cdots & c_{n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c_{-n+1} & c_{-n+2} & c_{-n+3} & \cdots & c_0 \end{pmatrix}$$

of size $n \times n$ always have non-negative determinant for any n. This algebraic condition is equivalent to that H_n are all positive semi-definite matrices. Since the principal minors of a positive semi-definite matrix is positive semi-definite, and since a hermitian matrix such that every leading principal minor has non-negative determinant is positive semi-definite, the bilateral sequence $(c_k)_{k=-\infty}^{\infty}$ is positive definite function when we consider it as a complex-valued function on $\mathbb{Z}$ that maps an integer k to c_k if and only if it is a positive definite *sequence* in the following sense:

Definition 2.2. A bilateral complex sequence $(c_k)_{k=-\infty}^{\infty}$ is said to be *positive definite* if

$$\sum_{k,l=1}^{n} c_{k-l}\xi_k\overline{\xi_l} \geq 0$$

for each n and $(\xi_1, \cdots, \xi_n) \in \mathbb{C}^n$.

Theorem 2.4 (Carathéodory-Toeplitz). *Let f be an analytic function on the open unit disk with the power series expansion*

$$f(z) = 1 + \sum_{k=1}^{\infty} 2c_k z^k.$$

Then, f belongs to the Carathéodory class if and only if the sequence $(c_k)_{k=-\infty}^{\infty}$ is positive definite, where we use the notations $c_0 = 1$ and $c_{-k} = \overline{c_k}$.

Proof. ($\Rightarrow$) If f is in the Carathéodory class, then because

$$c_{k-l} r^{|k-l|} = \frac{1}{2\pi} \int_0^{2\pi} \operatorname{Re} f(re^{i\theta}) e^{-i(k-l)\theta} \, d\theta,$$

we have

$$\sum_{k,l=1}^{n} c_{k-l}\xi_k\overline{\xi}_l = \lim_{r\uparrow 1}\frac{1}{2\pi}\int_0^{2\pi} \operatorname{Re} f(re^{i\theta})\left|\sum_{k=1}^{n} e^{-ik\theta}\xi_k\right|^2 d\theta \geq 0$$

for each n.

($\Leftarrow$) Conversely, assume that the coefficient sequence $(c_k)_{k=-\infty}^{\infty}$ is positive definite. Put $\xi_k = z^{k-1}$ and $z = re^{i\theta}$ to write

$$\begin{aligned}
0 &\leq \sum_{k,l=1}^{n+1} c_{k-l} z^{k-1}(\overline{z})^{l-1} \\
&= \sum_{k,l=0}^{n} c_{k-l} r^{k+l} e^{i(k-l)\theta} \\
&= \sum_{k,l=0}^{n} c_{k-l} r^{|k-l|} r^{2\min\{k,l\}} e^{i(k-l)\theta} \\
&= \sum_{k=-n}^{n} c_k r^{|k|} e^{ik\theta} \sum_{l=0}^{n-|k|} r^{2l} \\
&= \sum_{k=-n}^{n} c_k r^{|k|} e^{ik\theta} \frac{1-r^{2(n-|k|+1)}}{1-r^2} \\
&= \frac{1}{1-r^2}\sum_{k=-n}^{n} c_k r^{|k|} e^{ik\theta} - \frac{r^{n+2}}{1-r^2}\sum_{k=-n}^{n} c_k r^{n-|k|} e^{ik\theta}.
\end{aligned}$$

For $r = |z| < 1$ the first term tends to

$$\lim_{n\to\infty}\frac{1}{1-r^2}\sum_{k=-n}^{n} c_k r^{|k|} e^{ik\theta} = \frac{1}{1-|z|^2}\operatorname{Re} f(z),$$

and $|c_k| \leq c_0 = 1$ implies the second term vanishes as

$$\left|\frac{r^{n+2}}{1-r^2}\sum_{k=-n}^{n} c_k r^{n-|k|} e^{ik\theta}\right| \leq \frac{r^{n+2}}{1-r^2}(2n+1) \to 0$$

as $n \to \infty$. It proves $\operatorname{Re} f(z) \geq 0$ for $|z| < 1$, and we obtain $\operatorname{Re} f(z) > 0$ by the open mapping theorem. $\square$

2.3 The Herglotz-Riesz representation theorem

Herglotz proved another equivalent condition in 1911, considered as the first Bochner-type theorem, which states the correspondence between the Carathéodory class and probability Borel measure on the unit circle. The essential difficulty comes from the construction of a measure, and here we resolve this in the aid of either Helly's selection theorem or the Riesz-Markov-Kakutani representation theorem. Suppose a function f in the Carathéodory class is analytic on a neighborhood of the closed unit disk $\overline{\mathbb{D}}$. In this case, by appropriately manipulate the identities for $r = 1$ in Lemma 2.1, or by using the Cauchy integral formula along the unit circle, we can get

$$f(z) = \frac{1}{2\pi} \int_0^{2\pi} \frac{e^{i\theta} + z}{e^{i\theta} - z} \operatorname{Re} f(e^{i\theta})\, d\theta.$$

Based on this representation of f, we will approximate the measure $d\mu$ with the absolutely continuous measures $(2\pi)^{-1} \operatorname{Re} f(re^{i\theta})\, d\theta$ by limiting $r \uparrow 1$. More precisely, we will use the following equation, which is justified by the uniform convergence of power series: for each $|z| < 1$,

$$\begin{aligned}
\lim_{r\uparrow 1} \frac{1}{2\pi} \int_0^{2\pi} \frac{e^{i\theta} + z}{e^{i\theta} - z} \operatorname{Re} f(re^{i\theta})\, d\theta &= \lim_{r\uparrow 1} \frac{1}{2\pi} \int_0^{2\pi} \left(1 + \sum_{k=1}^{\infty} 2e^{-ik\theta} z^k\right) \operatorname{Re} f(re^{i\theta})\, d\theta \\
&= 1 + \lim_{r\uparrow 1} \sum_{k=1}^{\infty} 2 \left(\frac{1}{2\pi} \int_0^{2\pi} e^{-ik\theta} \operatorname{Re} f(re^{-i\theta})\, d\theta \right) z^k \\
&= 1 + \lim_{r\uparrow 1} \sum_{k=1}^{\infty} 2c_k r^k z^k \\
&= f(z).
\end{aligned}$$

Theorem 2.5 (The Herglotz-Riesz representation theorem). *Let f be a complex-valued function defined on the open unit disk. Then, f belongs to the Carathéodory class if and only if f is represented as the following Stieltjes integral*

$$f(z) = \int_0^{2\pi} \frac{e^{i\theta} + z}{e^{i\theta} - z}\, d\mu(\theta),$$

where μ is a probability Borel measure on $\mathbb{T} = \mathbb{R}/2\pi\mathbb{Z}$.

First proof: using Helly's selection theorem. ($\Leftarrow$) Take a probability Borel measure μ on $\mathbb{T}$. Then, we can check the function defined by

$$f(z) := \int_0^{2\pi} \frac{e^{i\theta} + z}{e^{i\theta} - z}\, d\mu(\theta)$$

is analytic on the open unit disk easily by using Morera's theorem and Fubini's theorem. Recall that $z \mapsto (e^{i\theta} + z)/(e^{i\theta} - z)$ has positive real part since it is a conformal mapping that maps open unit disk onto the right half plane. The function f belongs to the Carathéodory class by the open mapping theorem since

$$\operatorname{Re} f(z) = \int_0^{2\pi} \operatorname{Re} \frac{e^{i\theta}+z}{e^{i\theta}-z}\, d\mu(\theta) \geq 0.$$

($\Rightarrow$) Fix z in the open unit disk $\mathbb{D}$. Define $f_n(\theta) := (2\pi)^{-1} \operatorname{Re} f((1-n^{-1})e^{i\theta})$ and

$$F_n(\theta) := \int_0^{\theta} \operatorname{Re} f_n(\psi)\, d\psi$$

for $\theta \in [0, 2\pi]$. Note $F_n(0) = 0$ and $F_n(2\pi) = 1$ for all n. Since $\operatorname{Re} f \geq 0$, F_n is also monotonically increasing. Therefore, the sequence $(F_n)_n$ has a pointwise convergent subsequence $(F_j)_j$ on $[0, 2\pi]$ by the Helly's selection theorem. Let

$$F(\theta) := \lim_{\psi \downarrow \theta} \lim_{j \to \infty} F_j(\psi).$$

Then, we have $F(0) = 0$ and $F(2\pi) = 1$, and F_j converges to F at every continuity point θ of F. It means F_j converges to F weakly, so by the Portemanteau theorem, we get

$$\int_0^{2\pi} \frac{e^{i\theta}+z}{e^{i\theta}-z} dF_j(\theta) \to \int_0^{2\pi} \frac{e^{i\theta}+z}{e^{i\theta}-z} dF(\theta)$$

as $j \to \infty$ since $\theta \mapsto (e^{i\theta}+z)/(e^{i\theta}-z)$ is continuous and bounded on $\mathbb{T}$. On the other hand,

$$\int_0^{2\pi} \frac{e^{i\theta}+z}{e^{i\theta}-z} dF_j(\theta) = \frac{1}{2\pi} \int_0^{2\pi} \frac{e^{i\theta}+z}{e^{i\theta}-z} \operatorname{Re} f((1-j^{-1})e^{i\theta})\, d\theta \to f(z)$$

as $j \to \infty$. Therefore, by the uniqueness of limit, we have

$$f(z) = \int_0^{2\pi} \frac{e^{i\theta}+z}{e^{i\theta}-z} dF(\theta) = \int_0^{2\pi} \frac{e^{i\theta}+z}{e^{i\theta}-z} d\mu(\theta),$$

where μ is the probability measure on $\mathbb{T}$ defined by the distribution function F as $\mu([0, \theta]) = F(\theta)$. □

Second proof: using the Riesz representation theorem. As we have seen in the first proof using Helly's selection theorem, one direction is trivial. Suppose f is in the Carathéodory class. Let $g \in C(\mathbb{T})$ be a complex-valued test function. Define a complex linear functional l on $C(\mathbb{T})$ as

$$l[g] := \lim_{r \uparrow 1} \frac{1}{2\pi} \int_0^{2\pi} g(\theta) \operatorname{Re} f(re^{i\theta})\, d\theta.$$

It is positive and bounded since $\operatorname{Re} f \geq 0$ and $l[1] = 1$. By the Riesz-Markov-Kakutani representation theorem, there is a measure μ on $\mathbb{T}$ such that

$$l[g] = \frac{1}{2\pi} \int_0^{2\pi} g(\theta)\, d\mu(\theta)$$

for all $g \in C(\mathbb{T})$. If we expand f as

$$f(z) = 1 + \sum_{k=1}^{\infty} 2c_k z^k,$$

then for each fixed z in the open unit disk we have

$$\frac{1}{2\pi} \int_0^{2\pi} \frac{e^{i\theta} + z}{e^{i\theta} - z} d\mu(\theta) = \lim_{r \uparrow 1} \frac{1}{2\pi} \int_0^{2\pi} \frac{e^{i\theta} + z}{e^{i\theta} - z} \operatorname{Re} f(re^{i\theta})\, d\theta = f(z). \qquad \square$$

As a corollary of the Herglotz' theorem, we finally arrive at:

Corollary 2.6 (The Bochner theorem on $\mathbb{Z}$). *A function $c : \mathbb{Z} \to \mathbb{C}$ is positive-definite if and only if there is a finite regular Borel measure μ on $\mathbb{T} = \mathbb{R}/2\pi\mathbb{Z}$ such that $2\pi\widehat{\mu} = c$.*

Proof. Let μ be a finite Borel measure on $\mathbb{T}$. Then, $\widehat{\mu}$ is positive definite because

$$\begin{aligned}
\sum_{k,l=1}^{n} \widehat{\mu}(k-l)\xi_k\overline{\xi}_l &= \sum_{k,l=1}^{n} \frac{1}{2\pi} \int_0^{2\pi} e^{-i(k-l)\theta}\, d\mu(\theta)\; \xi_k\overline{\xi}_l \\
&= \frac{1}{2\pi} \int_0^{2\pi} \left| \sum_{k=1}^{n} e^{-ik\theta} \xi_k \right|^2 d\mu(\theta) \geq 0
\end{aligned}$$

for any $(\xi_1, \cdots, \xi_n) \in \mathbb{C}^n$.

On the other hand, if the sequence $(c_k)_{k=-\infty}^{\infty}$ is positive definite, then, with the assumption $c_0 = 1$, the function $z \mapsto 1 + \sum_{k=1}^{\infty} 2c_k z^k$ is in the Carathéodory class. By the Herglotz-Riesz representation theorem, there is a probability regular Borel measure μ on $\mathbb{T}$ such that

$$\begin{aligned}
1 + \sum_{k=1}^{\infty} 2c_k z^k &= \int_0^{2\pi} \frac{e^{it} + z}{e^{it} - z}\, d\mu(t) \\
&= \int_0^{2\pi} \left(1 + \sum_{k=1}^{\infty} 2e^{-ik\theta} z^k \right) d\mu(t) \\
&= 1 + \sum_{k=1}^{\infty} 2 \left(\int_0^{2\pi} e^{-ik\theta}\, d\mu(\theta) \right) z^k
\end{aligned}$$

in $z \in \mathbb{D}$, hence that

$$c_k = \int_0^{2\pi} e^{-ik\theta}\, d\mu(\theta) = 2\pi\widehat{\mu}(k). \qquad \square$$

So far, we have proved the following one-to-one correspondences:

$$
\begin{array}{ccccc}
 & & \begin{array}{c}\text{Closed convex hull of}\\ \text{the curve } (e^{-i\theta}, e^{-i2\theta}, \cdots)\end{array} & & \\
 & & \Big\updownarrow & & \\
\begin{array}{c}\text{Positive definite}\\ \text{sequences } (c_k)_k\\ \text{with } c_0 = 1\end{array} & \longleftrightarrow & \text{Carathéodory class} & \longleftrightarrow & \begin{array}{c}\text{Probability Borel}\\ \text{measures on } \mathbb{T}\end{array}
\end{array}.
$$

3 On the group $\mathbb{R}$: probability theory

3.1 Weak convergence of probability measures

We have seen the relation of positive definite sequences and measures on the unit circle $\mathbb{T}$. On the real line $\mathbb{R}$, predictably, we can also prove that there exists a correspondence between measures and positive definite functions. Herglotz' representation theorem used measures to characterize certain complex functions, but on $\mathbb{R}$, and more generally on a locally compact group G, we will see how the positive functions play an important role in studying measures. The use of positive definite functions to study measures virtually starts in probability theory; a probability distribution is defined as a measure on a state space, which is $\mathbb{R}$ for usual random variables. Especially to see several probability distributions and their limit behaviors, the theory and tools to deal with weak convergence of probability distributions have occured, and positive definite functions were the most successful. The direct connection between convergences in two different realms, measures and positive definite functions, is encoded in the Lévy continuity theorem. For example, several limit theorems in probability theory such as Lindeberg's central limit theorem are described and proved in terms of positive definite functions.

Definition 3.1 (Weak convergence)**.** Let $(\mu_n)_n$ and μ be probability Borel measures on a metric space. We say μ_n *weakly converges to* μ if

$$\int g\, d\mu_n \to \int g\, d\mu$$

as $n \to \infty$ for any g continuous bounded functions.

A characteristic function is defined as the Fourier transform, but reversed the sign convention on the phase term, of a probability measure, and is the place where the positive definiteness comes.

Definition 3.2 (Characteristic functions)**.** Let μ be a probability measure on $\mathbb{R}$ and X a random variable of distributioin μ. Note that such random variable always exists. The *characteristic function* of X is a function $\varphi : \mathbb{R} \to \mathbb{C}$ defined by $\varphi(t) := Ee^{itX}$. Equilvalently, φ is given by

$$\varphi(t) := \int e^{itx}\, d\mu(x).$$

Proposition 3.1. *Let φ be a characteristic function of a probability Borel measure μ on $\mathbb{R}$.*

(a) φ *is positive definite.*

(b) φ *is uniformly continuous.*

Proof. □

Characteristic functions take an advantage that we can learn the information about probability measures by investigating the continuous functions instead of studying measures directly.

Theorem 3.2 (Lévy's continuity theorem)**.** *Let* $(\mu_n)_n$ *be a sequence of probability Borel measures and* φ_n *their characteristic functions. Then,* μ_n *converges weakly to a probability Borel measure* μ *if and only if* φ_n *converges pointwise to a function* φ *that is continuous at zero.*

Proof. □

3.2 Vague convergence of probability measures

We now have a question: why is the weak convergence defined like that? In subsequent sections 4 and 5, we will extend the domain $\mathbb{R}$ to some general spaces like locally compact groups G. The topology of weak convergence is expected to be generalized on G. In this regard, we suggests a possible answer to the above question in view of functional analysis. We are also now going to see here a small hint of why we extend the theory only to *locally compact groups*, but not to the more general topological groups.

Vauge convergence of probability measures, which is equivalent to the weak convergence in nice domains such as $\mathbb{R}$, has an entirely suitable functional-analytic interpretation: the weak* topology. When probability measures are recognized as positive linear functionals on a function space via integration, they are embeded into the dual of the function space, and the weak* topology can be inherited onto the space of probability measures. Furthermore, the Riesz-Markov-Kakutani representation theorem allows to describe the exact geometry of the space of probability measures as a *compact* convex subset of the locally convex space $C_0(\mathbb{R})^*$.

We will see that why

3.3 Examples of positive definite functions

Mathias' examples Polya's criterion

3.4 A proof of Bochner's theorem

3.5 Application: Stone-von Neumann theorem

4 On locally compact abelian groups

4.1 Fourier transform and character group

4.2 Proofs of Bochner's theorem

4.3 Application: Pontryagin duality

5 On locally compact non-abelian groups

References