POSITIVE DEFINITE FUNCTIONS ON SPHERES

by

Fatma Yılmaz

BS, Integrated BS and MS Program in Teaching Mathematics, Boğaziçi University,

2009

Submitted to the Institute for Graduate Studies in

Science and Engineering in partial fulfillment of

the requirements for the degree of

Master of Science

Graduate Program in Mathematics

Boğaziçi University

2011

ACKNOWLEDGEMENTS

I am grateful to the many people who helped me get to this stage and complete this thesis.

First and foremost, I would like to express my sincere gratitude to my supervisor Prof. Alp Eden, firstly for introducing me to this area of analysis, for his excellent guidance and valuable suggestions. His perpetual energy and enthusiasm in research had motivated me throughout this period.

Asst. Prof. Fatih Ecevit deserves special thanks for his endless support and encouragement on all sides of my master studies as well as participating in my thesis committee.

I would like to thank Prof. Ahmet Feyzioğlu for encouraging and supporting me to take the first step for becoming a mathematician.

My sincere thanks to Prof. Haluk Beker for participating in my thesis committee.

I am grateful to Çağan Özen for his technical support besides being a true friend. I also thank to all of my officemates for such a friendly atmosphere.

I am far beyond being thankful to Tevfik Terzioğlu who has always been with me with his endless love, support and understanding.

Finally and above all, I am indebted to my family for their self-sacrifice and patience through my education.

ABSTRACT

POSITIVE DEFINITE FUNCTIONS ON SPHERES

Positive definite functions play a central role in approximation theory as in many other areas of mathematical research. The methods for data interpolation on spheres can be effectively used for analysis of large data sets arising from geosciences. In this endeavor, studying positive definite functions on spheres is essential. The characterization of positive definite functions on spheres in $\mathbb{R}^m$ using ultraspherical polynomials is given by I. J. Schoenberg in his celebrated 1942 paper "Positive Definite Functions on Spheres" where he also characterizes positive definite functions in the unit sphere of a real Hilbert space utilizing the cosine function. In this thesis, our aim is to expand the underlying ideas in Schoenberg's characterization of positive definite functions and review the results on some of its extensions. For this purpose, we firstly present the fundamental results in the theory of positive definite functions. We also review basic concepts in ultraspherical polynomials in which we present a proof of the addition formula for ultraspherical polynomials by simplifying the one in Nielsen's book as much as possible. Then, we analyze the proofs of Schoenberg's characterization of positive definite functions on finite and infinite dimensional unit spheres. Finally, we introduce strictly and conditionally positive definite functions and review some partial results on their characterizations.

ÖZET

KÜRE ÜZERİNDE POZİTİF TANIMLI FONKSİYONLAR

Pozitif tanımlı fonksiyonlar matematiğin diğer birçok alanında olduğu gibi, yaklaşım teorisinde de merkezi bir rol oynar. Küre üzerinde veri interpolasyon metotları, yerbilimlerinde ortaya çıkan geniş veri kümelerinin analizi için verimli bir şekilde kullanılabilmektedir. Bu durum küre üzerinde pozitif tanımlı fonksiyonlari çalışmayı önemli kılar. Pozitif tanımlı fonksiyonların ultra-küresel fonksiyonları içeren karakterizasyonu, 1942 yılında I. J. Schoenberg tarafından yazılan "Küre Uzerinde Pozitif Tanımlı Fonksiyonlar" adlı makalede verilmiştir. Yine bu makalede Schoenberg, gerçek Hilbert uzayı üzerinde pozitif tanımlı fonksiyonları, kosinüs fonksiyonundan yararlanarak karakterize etmiştir. Bu tezin amacı, Schoenberg'in pozitif tanımlı fonksiyonlar için yaptığı karakterizasyonların temelindeki fikirleri açıklamak ve bu tür fonksiyonların bazı genellemeleri üzerindeki neticelerin bir özetini sunmaktır. Bu amaçla, öncelikle pozitif tanımlı fonksiyonlar teorisindeki temel neticeler sunulmuştur. Bunun yanısıra, ultra-küresel fonksiyonlar konusuna ilişkin temel kavramların üzerinden geçilmiştir. Ultra-küresel fonksiyonlar için toplama formülünün ispatı da Nielsen'in kitabındaki ispatı mümkün olduğunca sadeleştirilerek, yine bu bölümde verilmiştir. Daha sonra, Schoenberg'in sonlu ve sonsuz boyutlu küre üzerinde pozitif tanımlı fonksiyonların karakterizasyonlarını içeren teoremlerinin ispatları çözümlenmiştir. Son olarak da, tam ve koşullu pozitif tanımlı fonksiyonlar tanıtılıp, bu tür fonksiyonların küre üzerinde tanımlı olanları için verilen karakterizasyonları içeren kısmi neticelere değinilmiştir.

TABLE OF CONTENTS

LIST OF SYMBOLS

$\mathfrak{B}(M)$	The class of positive definite functions on a metric space M
$C(X)$	Continuous functions on a set X
$C^{\infty}(X)$	Infinitely differentiable functions on a set X
$D_x^n f$	n-th derivative of f wrt. x
$f\|_X$	Restriction of f on X
$\widehat{f}$	Fourier transform of f
$f^{\curlyvee}$	Inverse Fourier transform of f
$\mathcal{F}^*$	Dual space of a function space $\mathcal{F}$
$L_1(\mathbb{R}^n)$	Integrable functions on $\mathbb{R}^n$
$\mathcal{L}f$	Laplace transform of f
$\mathbb{N}$	The set of natural numbers
pp'	Spherical distance between p and p'
$P_n^{(\alpha,\beta)}(x)$	Jacobi polynomial of degree n with parameters α and β
$P_n^{(\lambda)}(x)$	Ultraspherical polynomial of degree n and order λ
$P_n(x)$	Legendre polynomial of degree n
$P_n^s(x)$	Associated Legendre function
$\mathbb{R}^n$	n-dimensional Euclidean space
$Re(z)$	Real part of z
$\mathcal{S}^n$	n-dimensional unit sphere
$\mathcal{S}^{\infty}$	Unit sphere in a real Hilbert sphere
$\mathcal{S}(\mathbb{R}^n)$	The Schwartz space of functions on $\mathbb{R}^n$
$T_n(x)$	Chebyshev Polynomial of degree n
$\mathbb{Z}_+$	The set of non-negative integers
δ_{nm}	Kronecker delta function
δ_y	Point evaluation functional at y
$\Gamma(z)$	Gamma function
$\pi_m(\mathbb{R}^n)$	Space of n-variate polynomials of total degree at most m

$\pi_m(\mathbb{R}^n)^\perp$	Space of point evaluation functionals on $\pi_m(\mathbb{R}^n)$
$\square$	End of proof
$:=$	Equality that includes a definition
$\|\cdot\|$	Euclidean norm
$\subset$	Subset

1. INTRODUCTION

In practical applications over wide fields of study one often faces the problem of reconstructing an unknown function f from a finite set of data. These data consist of data sites $X = \{x_1, ..., x_N\}$ and data values $f_j = f(x_j), 1 \leq j \leq N$, and the reconstruction has to approximate or recover the data values at the data sites usually from a class of functions. In other words, a function s is sought that either approximates the data, $s(x_j) \approx f_j$, or interpolates the data, i.e. that satisfies $s(x_j) = f_j, 1 \leq j \leq N,$. The former case is in particular important if the data have noise [1].

In many cases, the data sites are *scattered*, i.e. they bear no regular structure at all, and there is a very large number of them. In some applications, the data sites also exist in a space of very high dimensions [1]. Some of the methods used to solve this problem are spline methods, radial basis functions, least squares etc..

In the univariate setting, polynomial interpolation is a useful tool. However, it is a well established fact that a large data set is better dealt with by splines, i.e. piecewise polynomials, than by polynomials. In contrast to polynomials, the accuracy of interpolation process using splines is not based on the polynomial degree but on the spacing of the data sites [1].

Spline methods usually require a triangulation of the data set in order to define the space from which we approximate, unless the data sites are in a very special position, e.g. gridded or otherwise highly regularly distributed. The reason for this is that it has to be decided where the pieces of the piecewise polynomials lie and where they are joined together. Moreover, it then has to be decided with what smoothness they are joined together at common vertices, edges etc. and how it is done. This is not trivial in more than one dimension and it is highly relevant in connection with the dimension of the space. The quality of the spline approximation depends strongly on triangulation itself, but triangulations or similar structures (such as quadrangulations) can be very difficult to provide in more than two dimensions. This is one of the severe disadvantages of piecewise polynomial techniques [2]. For further details about the triangulation process, one may consult on [3, p. 420], for example.

In the multivariate setting, although spline interpolation is tough to handle, it gives a motivation for a framework in higher dimensions. In this alternative approach, called *Radial Basis Function Method*, instead of using splines, one can form the approximant by taking finite linear combinations of translates of a radially symmetric basis function, say $\phi(\|\cdot\|)$ where $\|\cdot\|$ is the Euclidean norm. Radial symmetry means that the value of a function depends only on the Euclidean distance of the argument from the origin, and any rotations thereof make no difference to the function value [2]. The simplest example is, [2], for a finite set of *centers* $X \subseteq \mathbb{R}^m$, we can form the space of approximants S by

$$S = \{\sum_{x_j \in X} a_j \| \cdot - x_j \| : a_j \in \mathbb{R}\}. \tag{1.1}$$

Here the radial basis function is simply $\phi(r) = r$, the radial symmetry stemming from the Euclidean norm $\|\cdot\|$, and we are shifting this norm in (1.1) by the centers x_j.

More generally, radial basis function spaces are spanned by translates

$$\phi(\| \cdot - x_j \|), \quad x_j \in X,$$

where $\phi : [0, \infty) \to \mathbb{R}$ is a given, continuous function, called *radial basis function.* Therefore approximants have the general form

$$s(x) = \sum_{x_j \in X} a_j \phi(\| x - x_j \|), \quad \text{for} \quad x \in \mathbb{R}^m, \tag{1.2}$$

with real coefficients a_j. In certain cases, low-degree polynomials have to be added but we will not discuss this until the last chapter. Other most common examples of radial basis functions $\phi(r)$, for $r = \| x - y \|$, $x, y \in \mathbb{R}^n$, are

- Thin plate splines: $\phi(r) = r^{2k} \log r$ $(k = 1, 2, ...,)$,
- Multiquadrics: $\phi(r) = (r^2 + c^2)^\beta$ $(c > 0,\ 0 < \beta < 1)$,
- Inverse multiquadrics: $\phi(r) = \frac{1}{(r^2+c^2)^\beta}$ $(c > 0,\ \beta > 0)$,
- The Gaussians: $\phi(r) = e^{-\alpha r^2}$ $(\alpha > 0)$,
- Logarithmic: $\phi(r) = log(r^2 + c^2)$.

A good choice of the radial basis function is important for the quality of the approximation and for the existence of the interpolants. The constants in the above examples are usually adjusted using experimental techniques. More examples and approximation properties of the above functions can be found in [1], [2].

Now it is apparent that radial basis functions allows to work for large dimensional spaces because the function reduces the multivariate setting to the univariate setting. Further remarkable properties of radial basis functions that render them highly efficient in practice are their easily adjustable smoothness and their powerful convergence properties [2].

In fact, several results hold if we replace $\phi(\| \cdot - x_j \|)$ in (1.2) with a more general function $\Phi : \Omega \times \Omega \to \mathbb{R}$. Of course the latter case can only work if $X \subseteq \Omega$. We will call such a Φ a *kernel* rather than a function. Now, we can restate our interpolation problem: Given the data values $f_1, ..., f_N$ at given data sites $X = \{x_1, ..., x_N\} \subseteq \mathbb{R}^m$, choose a fixed function $\Phi : \Omega \times \Omega \to \mathbb{R}$ and form the interpolant as

$$s(x) = \sum_{k=1}^{N} \alpha_k \Phi(x, x_k) \tag{1.3}$$

where the coefficients α_k are determined, if possible, by the interpolation conditions

$$s(x_j) = f_j, \quad 1 \le j \le N. \tag{1.4}$$

This is equivalent to asking for a non-singular interpolation matrix

$$A_{\Phi,X} := (\Phi(x_j, x_k))_{1\leq j,k\leq N}. \tag{1.5}$$

There are certain cases that ensure the invertibility of the interpolation matrix. One of them is that if the interpolation matrix is *positive definite*, then it is invertible; thus the interpolation problem is well-posed. Also, positive definiteness of the matrix enables to use efficient algorithms like conjugate gradient method [3]. Hence, when forming the interpolant, choosing a kernel Φ that generate a positive definite matrix will greatly facilitate the analysis. This leads to the definition of positive definite kernels.

Definition 1.1. *A continuous kernel $\Phi : \Omega \times \Omega \to \mathbb{C}$ is called positive definite on $\Omega \subseteq \mathbb{R}^m$ if, for all $N \in \mathbb{N}$, for all sets of pairwise distinct centers $X = \{x_1, ..., x_N\} \subseteq \mathbb{R}^m$ and all $\alpha \in \mathbb{C}^N$, the quadratic form*

$$\sum_{j=1}^{N}\sum_{k=1}^{N} \alpha_j \overline{\alpha_k} \Phi(x_j, x_k) \tag{1.6}$$

is nonnegative. The kernel Φ is called strictly positive definite on $\Omega \subseteq \mathbb{R}^m$ if the quadratic form is positive for all $\alpha \in \mathbb{C}^N \setminus \{0\}$.

The theory of positive definite functions is well studied. In fact, positive definite functions and their various analogues and generalizations have arisen in diverse parts of mathematics since the beginning of this century [4]. They occur naturally in Fourier analysis, probability theory [5], operator theory [6], moment problems [7], integral equations [8], boundary-value problems for partial differential equations, embedding problems [9], machine learning [10], and other areas. Mathias [11], was the first person to define and study the properties of positive definite functions of a real variable. But according to [4], he apparently did not realize that more than a decade previously Mercer [8] and others had considered the more general concept of positive definite kernels in research on integral equations.

There is a close connection between positive definite kernels and the reproducing-kernel Hilbert spaces which is well covered in Aronszajn's classical paper, "Theory of Reproducing Kernels" [12]. Before showing this connection, we recall the basic concepts related to the theory of reproducing kernels following mostly [13] and [1]. Here, we state only the results. For the proofs one may consult on [12, 13, 1, 14].

Definition 1.2. *Let $\mathcal{F}$ be a real Hilbert space of functions $f : \Omega \to \mathbb{R}$. A function $\Phi : \Omega \times \Omega \to \mathbb{R}$ is called a reproducing kernel for $\mathcal{F}$ if*

i. $\Phi(\cdot, y) \in \mathcal{F}$ *for all* $y \in \Omega$.

ii. $f(y) = (f, \Phi(\cdot, y))_{\mathcal{F}}$ *for all* $f \in \mathcal{F}$ *and all* $y \in \Omega$.

It follows from the definition that the reproducing kernel of a Hilbert space is uniquely determined.

Theorem 1.3. *Suppose that $\mathcal{F}$ is a Hilbert space of functions $f : \Omega \to \mathbb{R}$, and $\mathcal{F}^*$ is the dual space of $\mathcal{F}$. Then the following statements are equivalent:*

i. the point evaluation functionals are continuous, i.e. $\delta_y \in \mathcal{F}^*$ *for all* $y \in \Omega$.

ii. $\mathcal{F}$ *has a reproducing kernel.*

A reproducing-kernel Hilbert space has several properties:

Theorem 1.4. *Suppose that $\mathcal{F}$ is a Hilbert space of functions $f : \Omega \to \mathbb{R}$ with reproducing kernel Φ. Then we have*

i. $\Phi(x, y) = (\Phi(\cdot, x), \Phi(\cdot, y))_{\mathcal{F}} = (\delta_x, \delta_y)_{\mathcal{F}^*}$ *for* $x, y \in \Omega$,

ii. $\Phi(x, y) = \Phi(y, x)$ *for* $x, y \in \Omega$,

iii. if $f, f_n \in \mathcal{F}$, $n \in \mathbb{N}$, *are given such that* f_n *converges to* f *in the Hilbert space norm then* f_n *also converges pointwise to* f.

Our next result discloses the connection between reproducing-kernel Hilbert spaces and positive definite kernels.

Theorem 1.5. *Suppose that $\mathcal{F}$ is a reproducing-kernel Hilbert function space with reproducing kernel $\Phi : \Omega \times \Omega \to \mathbb{R}$. Then Φ is positive definite. Moreover, Φ is* *strictly positive definite if and only if the point evaluation functionals are linearly independent in* $\mathcal{F}^*$.

Hence, the reproducing kernel of a function space $\mathcal{F}$ leads to a real valued positive definite kernel. If the function space $\mathcal{F}$ is a complex vector space containing complex valued functions everything said so far remains true with mild modifications. In particular, the reproducing kernel is now a complex valued positive definite function [1].

So far, we have seen that a positive definite kernel appears naturally as the reproducing kernel of a Hilbert function space. But since we normally do not start with a function space but with a positive definite kernel we are confronted by the problem of finding the associated function space that has this kernel as the reproducing kernel. Although this is not trivial, we can construct the corresponding Hilbert function space for a strictly positive definite symmetric kernel $\Phi : \Omega \times \Omega \to \mathbb{R}$ [13], [1].

Definition 1.6. *If a (strictly) positive definite kernel* $\Phi : \Omega \times \Omega \to \mathbb{R}$ *is the reproducing kernel of a real Hilbert function space* $\mathcal{F}$ *of real valued functions on* Ω*, then* $\mathcal{F}$ *is called the* native space *for* Φ.

Theorem 1.7. *Any strictly positive definite kernel* Φ *on some domain* Ω *has a unique native space. It is the closure of the space*

$$\mathcal{F}_\Phi(\Omega) = \Big\{ \sum_{j=1}^{N} \alpha_j \Phi(\cdot, x_j) \quad | \quad \alpha_j \in \mathbb{R}, \quad N \in \mathbb{N}, \quad x_j \in \Omega \Big\} \tag{1.7}$$

under the inner product

$$(\Phi(\cdot, x), \Phi(\cdot, y))_\Phi = \Phi(x, y) \tag{1.8}$$

for all $x, y \in \Omega$. The elements of the native space can be interpreted as functions via

$$\delta_x(f) = (f, \Phi(\cdot, x))_\Phi \quad \textit{for all} \quad x \in \Omega, f \in \mathcal{F}_\Phi(\Omega). \tag{1.9}$$

Note that (1.9) makes sense since point evaluation functionals δ_x extend continuously to the completion.

In many cases, the domain Ω of functions allows a group $\mathbb{T}$ of geometric transformations, and the Hilbert space $\mathcal{F}$ is invariant under this group. This means that

$$f \circ T \in \mathcal{F}$$

and

$$(f \circ T, g \circ T)_\mathcal{F} = (f, g)_\mathcal{F} \tag{1.10}$$

for all $f, g \in \mathcal{F}$, and $T \in \mathbb{T}$.
The invariance of the function space is inherited by the kernel.

Theorem 1.8. *Suppose that the reproducing-kernel Hilbert function space $\mathcal{F}$ is invariant under the transformations of $\mathbb{T}$; then the reproducing kernel Φ satisfies*

$$\Phi(Tx, Ty) = \Phi(x, y)$$

for all $x, y \in \Omega$ and all $T \in \mathbb{T}$.

By some easy additional arguments one can read off the following invariance properties inherited by reproducing kernels Φ from their Hilbert spaces $\mathcal{F}$ on Ω [13]:

- Invariance on $\Omega = \mathbb{R}^m$ under translations from $\mathbb{R}^m$ leads to *translation invari-*

ant functions $\Phi(x, y) = \phi(x - y)$ with $\phi(x) = \phi(-x) : \mathbb{R}^m \to \mathbb{R}$.

- In case of additional invariance under all orthogonal transformations we get *radial* functions $\Phi(x, y) = \phi(\parallel x - y \parallel)$ with $\phi : [0, \infty) \to \mathbb{R}$. Thus, radial basis functions arise naturally in all Hilbert spaces on $\mathbb{R}^m$ which are invariant under Euclidean rigid-body motions.
- Invariance on the sphere $\mathcal{S}^m$ under all orthogonal transformations leads to *zonal* functions $\Phi(x, y) = \phi(x \cdot y)$ for $\phi : [-1, 1] \to \mathbb{R}$ where $x \cdot y$ denotes the usual dot product of x and y.
- Spaces of periodic functions induce periodic reproducing kernels.

The paper [15] introduces the theory of basis functions on general manifolds, and corresponding error bounds can be found in [16].

There has been an increasing awareness of the importance of approximation on the sphere with applications to meteorology, oceanography, jeodesy etc.. See [17] for other applications.

Over the last several years, there has been much fundamental work done by Freeden and colleagues [18] as well as by Wahba [19], [20] concerning approximation on $\mathcal{S}^m$ [21]. Nevertheless, even though positive definite functions on spheres were introduced and characterized long ago by Schoenberg [9], the approximation power of such functions on $\mathcal{S}^m$ has not yet been nearly as well understood as on $\mathbb{R}^m$ [21].

Schoenberg, [9], showed that a continuous function f was positive definite (see Definition 2.1) on $\mathcal{S}^m$ if its expansion in *ultraspherical (or Gegenbauer) polynomials*,

$$f(\cos\theta) = \sum_{n=0}^{\infty} a_n P_n^{(\lambda)}(\cos\theta)$$

had all $a_n \geq 0$ where $\lambda = \frac{1}{2}(m-1)$.

In his article [9], Schoenberg also investigated what happens if the space dimension tends to infinity.

The fundamental aim of this thesis is to analyze the ideas given in [9] and present them in a self contained way. The rest of the thesis is organized as follows. In *Chapter 2*, we give some notations, definitions and preliminary results in the theory of positive definite functions.

For his celebrated result mentioned above, Schoenberg uses *the addition formula for ultraspherical polynomials*. To make the proof of the formula more accessible, in *Chapter 3* we give the proof in [22] by simplifying it as much as possible. Since the ideas given in the proof of the addition theorem and also other notions in the article require considerable amount of knowledge on ultraspherical polynomials, we give these results in *Chapter 3*.

Chapter 4 is devoted to positive definite functions on spheres. First, we analyze the proof of the theorem which characterizes positive definite functions in $\mathcal{S}^m$ in terms of ultraspherical polynomials. Then, using the results in $\mathcal{S}^m$, we prove a theorem which characterizes positive definite functions in $\mathcal{S}^\infty$ utilizing the cosine function following [9].

In *Chapter 5*, we mention several extensions of positive definite functions which are useful in scattered data interpolation. Firstly, we give the definitions of strictly positive definite functions and conditionally positive definite ones. Then, we review the results on the characterization of such functions on $\mathcal{S}^m$.

2. POSITIVE DEFINITE FUNCTIONS

In this chapter, we present the definition and basic properties of positive definite functions. After some examples of positive definite functions, we give two characterizations for them, one of which is due to Bochner [23], and the other is established by Schoenberg [24]. This chapter is mostly based on Wendland's book [1].

2.1. Definition and Basic Properties of Positive Definite Functions

In Chapter 1, we have already defined positive definiteness for arbitrary kernels. Now, we define it for complex valued functions on arbitrary metric spaces.

Definition 2.1. *Let M be a metric space with the distance function pq. A complex-valued continuous function $g(t)$ $(0 \le t \le$ diameter of $M)$ is said to be positive definite in M if we have*

$$\sum_{j,k=1}^{n} \alpha_j \overline{\alpha_k} g(p_j p_k) \ge 0, \tag{2.1}$$

for any n points $p_1, ..., p_n$ of M, arbitrary $\alpha_j \in \mathbb{C}$, and for all $n \in \mathbb{N}$. The function $g(t)$ is called strictly positive definite on M if the quadratic form (2.1) is positive for all $\alpha \in \mathbb{C}^n \setminus \{0\}$, for all $n \in \mathbb{N}$.

We denote the class of positive definite functions by the symbol $\mathfrak{B}(M)$.

Theorem 2.2. *$\mathfrak{B}(M)$ enjoys the following closure properties:*

i. If $g_1(t) \in \mathfrak{B}(M)$, $g_2(t) \in \mathfrak{B}(M)$, then also $c_1 g_1(t) + c_2 g_2(t) \in \mathfrak{B}(M)$, provided $c_1 \ge 0$, $c_2 \ge 0$.

ii. The same assumptions imply also that $g_1(t) g_2(t) \in \mathfrak{B}(M)$.

iii. If $g_m(t) \in \mathfrak{B}(M)$, $g_m(t) \to g(t)$ as $m \to \infty$, and $g(t)$ is continuous, then also $g(t) \in \mathfrak{B}(M)$.

Proof. (i) Assume $g_1(t)$ and $g_2(t)$ are positive definite functions in M. Let $p_1, ..., p_n$ be arbitrary n points of M, and $\alpha_j \in \mathbb{C}$ be arbitrary. Then, for all $n \in \mathbb{N}$, the quadratic form

$$\sum_{j,k=1}^{n} \alpha_j \overline{\alpha_k}[c_1 g_1(p_j p_k) + c_2 g_2(p_j p_k)] = c_1 \sum_{j,k=1}^{n} \alpha_j \overline{\alpha_k} g_1(p_j p_k) + c_2 \sum_{j,k=1}^{n} \alpha_j \overline{\alpha_k} g_2(p_j p_k)$$

is nonnegative for $c_1 \geq 0$, $c_2 \geq 0$, since $g_1(t)$ and $g_1(t)$ are positive definite on M. Hence, $c_1 g_1(t) + c_2 g_2(t)$ is also positive definite on M for $c_1 \geq 0$, $c_2 \geq 0$.

(ii) Since $g_2(t)$ is positive definite, the matrix $A_{g_2,X} := [g_2(p_l p_j)]_{l,j=1,\ldots,n}$ is positive semi-definite and hermitian. Then, there exist a unitary matrix $U \in \mathbb{C}^{n\times n}$, $U = [u_{lk}]$, $(U\overline{U}^{\mathrm{T}} = I)$, such that $A_{g_2,X} = UDU^{\mathrm{T}}$, where $D = diag\{\lambda_1, ..., \lambda_n\}$ is the diagonal matrix with eigenvalues $0 \leq \lambda_1 \leq ... \leq \lambda_n$ of $A_{g_2,X}$ as diagonal entries. This means that

$$g_2(p_l p_j) = \sum_{k=1}^{n} u_{lk} \overline{u_{jk}} \lambda_k.$$

As $g_1(t)$ is positive definite, we have

$$\begin{aligned}
\alpha A_{g_1 g_2,X} \overline{\alpha} &= \sum_{l,j=1}^{n} \alpha_l \overline{\alpha_j} g_1(p_l p_j) g_2(p_l p_j) \\
&= \sum_{l,j=1}^{n} \alpha_l \overline{\alpha_j} g_1(p_l p_j) \sum_{k=1}^{n} u_{lk} \overline{u_{jk}} \lambda_k \\
&= \sum_{k=1}^{n} \lambda_k \sum_{l,j=1}^{n} \alpha_l u_{lk} \overline{\alpha_j u_{jk}} g_1(p_l p_j) \\
&\geq \lambda_1 \sum_{l,j=1}^{n} \alpha_l \overline{\alpha_j} g_1(p_l p_j) \sum_{k=1}^{n} u_{lk} \overline{u_{jk}} \\
&= \lambda_1 \sum_{l=1}^{n} |\alpha_l|^2 g_1(0).
\end{aligned}$$

The last expression is nonnegative for all $\alpha \in \mathbb{C}^n$ since taking $n = 1$ and $\alpha_1 = 1$ in

(2.1), we have

$$\sum_{j,k=1}^{n} g_1(p_j p_k)\alpha_j \overline{\alpha_k} = g_1(0) \geq 0.$$

(iii) Assume $g_m(t)$ are positive definite functions in M for all $m \in \mathbb{N}$. Let $p_1, ..., p_n$ be arbitrary n points of M, and $\alpha_j \in \mathbb{C}$ for $1 \leq j \leq n$ given. Then, for all $n \in \mathbb{N}$, it follows that

$$\sum_{j,k=1}^{n} \alpha_j \overline{\alpha_k} g(p_j p_k) = \sum_{j,k=1}^{n} \alpha_j \overline{\alpha_k} \lim_{m\to\infty} g_m(p_j p_k) = \lim_{m\to\infty} \sum_{j,k=1}^{n} \alpha_j \overline{\alpha_k} g_m(p_j p_k) \geq 0.$$

The last inequality follows from positive definiteness of g_m, for all $m \in \mathbb{N}$. $\square$

Now, if we take $M = \mathbb{R}^m$, and the distance function as the Euclidean distance, then for a positive definite function $\Phi : \mathbb{R}^m \to \mathbb{C}$, we have

$$\sum_{j,k=1}^{n} \alpha_j \overline{\alpha_k} \Phi(x_j - x_k) \geq 0 \tag{2.2}$$

for all pairwise distinct $x_1, ..., x_n$ of $\mathbb{R}^m$, for all $n \in \mathbb{N}$, and for all $\alpha_j \in \mathbb{C}$. Moreover, the following additional properties are satisfied.

Theorem 2.3. *Suppose $\Phi : \mathbb{R}^m \to \mathbb{C}$ is a positive definite function. Then,*

i. $\Phi(0) \geq 0$

ii. $\Phi(-x) = \overline{\Phi(x)}$ *for all* $x \in \mathbb{R}^m$.

iii. Φ *is bounded. More precisely,* $|\Phi(x)| \leq \Phi(0)$ *for all* $x \in \mathbb{R}^m$.

Proof. (i) This follows by choosing $n = 1$ and $\alpha_1 = 1$ in (2.2).

(ii) Setting $n = 2, \quad \alpha_1 = 1, \quad \alpha_2 = c, \quad x_1 = 0, \quad x_2 = x$ in (2.2) gives

$$\Phi(0)(1 + |c|^2) + c\Phi(x) + \overline{c}\Phi(-x) \geq 0$$

for every $c \in \mathbb{C}$, and so $c\Phi(x) + \overline{c}\Phi(-x)$ is real for every $c \in \mathbb{C}$. Setting $c = 1$, we have that $\Phi(x) + \Phi(-x)$ is real; setting $c = i$, we see that $i(\Phi(x) - \Phi(-x))$ is real. This can only be satisfied if $\Phi(x) = \overline{\Phi(-x)}$. This proves the second property.

(iii) Taking $n = 2$, $\alpha_1 = |\Phi(x)|$, $\alpha_2 = -\overline{\Phi(x)}$, $p_1 = 0$, $p_2 = x$ in (2.2) and using $\Phi(-x) = \overline{\Phi(x)}$ gives

$$2|\Phi(x)|^2\Phi(0) - 2|\Phi(x)|^3 \geq 0.$$

If $\Phi(x) = 0$ then we have seen from the first property that $\Phi(0) \geq 0$ and (2) follows. If not, then $|\Phi(x)| \leq \Phi(0)$ for all $x \in \mathbb{R}^m$. □

Example 2.4. Let $y \in \mathbb{R}^m$ be fixed. Then, the function

$$\Phi_y(x) = e^{ix\cdot y}$$

is positive definite on $\mathbb{R}^m$ since

$$\begin{aligned}
\sum_{j=1}^{n}\sum_{k=1}^{n} \alpha_j\overline{\alpha_k}\Phi_y(x_j - x_k) &= \sum_{j=1}^{n}\sum_{k=1}^{n} \alpha_j\overline{\alpha_k}e^{i(x_j - x_k)\cdot y} \\
&= \sum_{j=1}^{n} \alpha_j e^{ix_j\cdot y} \sum_{k=1}^{n} \overline{\alpha_k}e^{-ix_k\cdot y} \\
&= \left| \sum_{j=1}^{n} \alpha_j e^{ix_j\cdot y} \right|^2 \geq 0,
\end{aligned}$$

for all $\alpha_j \in \mathbb{C}$, $n \in \mathbb{N}$, and all pairwise distinct points $x_1, ..., x_n$ of $\mathbb{R}^m$.

One can observe by Theorem 2.3(i) that if a positive definite function is real valued then it must be even. Conversely, for an even real valued function it suffices to show that the quadratic form is nonnegative for all $\alpha \in \mathbb{R}^n$. This leads to the following theorem:

Theorem 2.5. *Suppose that $\Phi : \mathbb{R}^m \to \mathbb{R}$ is a continuous function. Then Φ is strictly positive definite if and only if Φ is even and we have, for all $n \in \mathbb{N}$, for all*

$\alpha \in \mathbb{R}^n \setminus \{0\}$, and for all pairwise distinct $x_1, ..., x_n \in \mathbb{R}^m$,

$$\sum_{j=1}^{n}\sum_{k=1}^{n} \alpha_j \alpha_k \Phi(x_j - x_k) > 0. \tag{2.3}$$

Proof. If Φ is strictly positive definite then it is even by the previous theorem; by definition of positive definiteness (2.3) is satisfied. Conversely, if Φ is even and satisfies (2.3), then we have for $\alpha_j = a_j + ib_j$

$$\begin{aligned}\sum_{j=1}^{n}\sum_{k=1}^{n} \alpha_j \overline{\alpha_k} \Phi(x_j - x_k) &= \sum_{j=1}^{n}\sum_{k=1}^{n} (a_j a_k + b_j b_k)\Phi(x_j - x_k) \\ &+ i\sum_{j=1}^{n}\sum_{k=1}^{n} a_k b_j [\Phi(x_j - x_k) - \Phi(x_k - x_j)].\end{aligned}$$

As Φ is even, the second sum on the right hand side is zero. The first sum is nonnegative because of the assumption and vanishes only if all the a_j's and b_j's vanish. $\square$

Definition 2.6. *Let $\mathcal{S}^m$ be the m-dimensional unit sphere on the Euclidean space $\mathbb{R}^{m+1}$, with norm $\| \cdot \|$, which is defined by*

$$\mathcal{S}^m = \{p \in \mathbb{R}^{m+1} : \quad \| p \| = 1\}.$$

The spherical distance (or geodesic distance) between two points $p, p' \in \mathcal{S}^m$ is defined as the length of the shorter part of the great circle joining p and p' or, in other words,

$$pp' := dist(p, p') = \arccos(p \cdot p')$$

where $p_j \cdot p_k$ denotes the usual dot product of p_j and p_k.

Example 2.7. $g(t) = \cos t$ is positive definite for $t \in [0, \pi]$ on every $\mathcal{S}^m$.

Proof. Let $p_1, ..., p_n$ be n arbitrary points and o be the center of $\mathcal{S}^m$. For n real variables $\alpha_1, ..., \alpha_n$, we have

$$\begin{aligned}
\sum_{j,k=1}^{n} \alpha_j \alpha_k \cos(p_j p_k) &= \sum_{j,k=1}^{n} \alpha_j \alpha_k p_j \cdot p_k \\
&= \Big(\sum_{j=1}^{n} \alpha_j p_j\Big) \cdot \Big(\sum_{k=1}^{n} \alpha_k p_k\Big) \\
&= \Big\| \sum_{j=1}^{n} \alpha_j p_j \Big\|^2 \geq 0.
\end{aligned}$$

$\square$

Example 2.8. The Gaussian $\phi(t) = e^{-c\|t\|^2}$, $c > 0$, is positive definite on every $\mathbb{R}^m$.

Proof. For arbitrary n points $p_1, ..., p_n$ of $\mathbb{R}^m$, we have by the Cauchy-Schwarz inequality

$$\|p_j - p_k\|^2 \leq 2(\|p_j\|^2 + \|p_k\|^2),$$

and since e^{-cx} is a monotone decreasing function on $[0, \infty)$ for $c > 0$,

$$e^{-c\|p_j - p_k\|^2} \geq e^{-2c(\|p_j\|^2 + \|p_k\|^2)}.$$

Therefore, for n real variables $\alpha_1, ..., \alpha_n$, we have

$$\begin{aligned}
\sum_{j,k=1}^{n} \alpha_j \alpha_k \phi(p_j - p_k) &= \sum_{j,k=1}^{n} \alpha_j \alpha_k e^{-c\|p_j - p_k\|^2} \\
&\geq \sum_{j,k=1}^{n} \alpha_j \alpha_k e^{-2c(\|p_j\|^2 + \|p_k\|^2)} \\
&= \sum_{j,k=1}^{n} \alpha_j \alpha_k e^{-2c\|p_j\|^2} e^{-2c\|p_k\|^2} \\
&= \Big(\sum_{j=1}^{n} \alpha_j e^{-2c\|p_j\|^2}\Big)^2 \geq 0.
\end{aligned}$$

$\square$

2.2. Integrally Positive Definite Functions

There is another characterization of positive definite functions using an integral analogue of (2.2).

Theorem 2.9. *A continuous function $\Phi : \mathbb{R}^m \to \mathbb{C}$ is positive definite if and only if Φ is bounded and satisfies*

$$\int_{\mathbb{R}^m}\int_{\mathbb{R}^m} \Phi(x-y)\gamma(x)\overline{\gamma(y)}dxdy \geq 0 \tag{2.4}$$

for all test functions γ from the Schwartz space $\mathcal{S}(\mathbb{R}^m)$.

For detailed information about $\mathcal{S}(\mathbb{R}^m)$, see [25] or [26] for example.

Proof. Suppose that Φ is positive definite Then, Φ is bounded by Theorem 2.3, and since $\gamma \in \mathcal{S}$ decays rapidly, the integral (2.4) is well-defined. Moreover, for every $\epsilon > 0$, there exist a closed cube $W \subseteq \mathbb{R}^m$ such that

$$\Big|\int_{\mathbb{R}^m}\int_{\mathbb{R}^m} \Phi(x-y)\gamma(x)\overline{\gamma(y)}dxdy - \int_W\int_W \Phi(x-y)\gamma(x)\overline{\gamma(y)}dxdy\Big| < \frac{\epsilon}{2}.$$

But the double integral over the cubes is the limit of Riemannian sums. Hence we can find $x_1, ..., x_n \in \mathbb{R}^m$ and weights $\omega_1, ..., \omega_n$ such that

$$\Big|\int_W\int_W \Phi(x-y)\gamma(x)\overline{\gamma(y)}dxdy - \sum_{i,k=1}^{n} \Phi(x_i - x_k)\gamma(x_i)\omega_i\overline{\gamma(x_k)}\overline{\omega_k}\Big| < \frac{\epsilon}{2}.$$

This means that

$$\Big|\int_{\mathbb{R}^m}\int_{\mathbb{R}^m} \Phi(x-y)\gamma(x)\overline{\gamma(y)}dxdy - \sum_{i,k=1}^{n} \Phi(x_i - x_k)\gamma(x_i)\omega_i\overline{\gamma(x_k)}\overline{\omega_k}\Big| < \epsilon.$$

Letting ϵ tend to zero and using that Φ is positive definite the above inequality shows that (2.4) is true for all $\gamma \in \mathcal{S}(\mathbb{R}^m)$.

Conversely, assume that Φ is bounded and satisfies (2.4). Our aim is to find a sequence of γ_l's, $\gamma_l \in \mathcal{S}$, satisfying

$$\sum_{j,k=1}^{n} \Phi(x_j - x_k)\alpha_j \overline{\alpha_k} = \lim_{l\to\infty} \int_{\mathbb{R}^m} \int_{\mathbb{R}^m} \Phi(x-y)\gamma_l(x)\overline{\gamma_l(y)}dxdy$$

allowing us to bound the left hand side through (2.4). To this end, we first show that

$$\int_{\mathbb{R}^m} \int_{\mathbb{R}^m} \Phi(x-y)\gamma(x)\overline{\gamma(y)}dxdy = \int_{\mathbb{R}^m} \Phi(x)(\gamma(x) * \widetilde{\gamma}(x))dx$$

where $\widetilde{\gamma}(x) = \overline{\gamma(-x)}$. Indeed, letting $s = x - y$ and $\tau = -y$, we have by change of area formula and Fubini's theorem

$$\begin{aligned}
\int_{\mathbb{R}^m} \int_{\mathbb{R}^m} \Phi(x-y)\gamma(x)\overline{\gamma(y)}dxdy &= \int_{\mathbb{R}^m} \int_{\mathbb{R}^m} \Phi(s)\gamma(s-\tau)\overline{\gamma(-\tau)}d\tau ds \\
&= \int_{\mathbb{R}^m} \Phi(s)\Big(\int_{\mathbb{R}^m} \gamma(s-\tau)\widetilde{\gamma}(\tau)d\tau\Big)ds \\
&= \int_{\mathbb{R}^m} \Phi(s)\big(\gamma * \widetilde{\gamma}(s)\big)ds \\
&= \int_{\mathbb{R}^m} \Phi(x)\big(\gamma * \widetilde{\gamma}(x)\big)dx.
\end{aligned}$$

Next, let $x_1, ..., x_n \in \mathbb{R}^m$, $\alpha_1, ..., \alpha_n \in \mathbb{C}$, and

$$\gamma = \gamma_l = \sum_{j=1}^{n} \alpha_j g_{2l}(\cdot - x_j)$$

where $g_l(x) = (\frac{l}{\pi})^{\frac{m}{2}} e^{-l\|x\|^2}$ (see Theorem (A.4)). Then,

$$(\gamma_l * \widetilde{\gamma}_l)(x) = \sum_{j,k=1}^{n} \alpha_j \overline{\alpha_k} g_l(x - (x_j - x_k)). \tag{2.5}$$

To see this, we use to compute the Fourier transform of γ_l as A.4(ii),

$$
\begin{aligned}
\widehat{\gamma}_l(\omega) &= \sum_{j=1}^{n} \alpha_j \widehat{\tau_{x_j} g_{2l}}(\omega) \\
&= (2\pi)^{-m/2} \sum_{j=1}^{n} \alpha_j e^{-i\omega^{\mathrm{T}} x_j} e^{-\|\omega\|_2^2/8l}.
\end{aligned}
$$

Then, we utilize Theorem A.3(iv), A.4(ii), and A.3(iii) to conclude

$$
\begin{aligned}
(\gamma_l * \widetilde{\gamma}_l)^{\wedge}(\omega) &= (2\pi)^{m/2} |\widehat{\gamma}_l|^2(\omega) \\
&= (2\pi)^{-m/2} \Big| \sum_{j=1}^{n} \alpha_j e^{-i\omega^{\mathrm{T}} x_j} \Big|^2 e^{-\|\omega\|^2/4l} \\
&= \sum_{j,k=1}^{n} \alpha_j \overline{\alpha_k} e^{-i\omega^{\mathrm{T}}(x_j - x_k)} \widehat{\gamma}_l(\omega) \\
&= \sum_{j,k=1}^{n} \alpha_j \overline{\alpha_k} (\tau_{(x_j - x_k)} g_l)^{\wedge}(\omega) \\
&= \Big(\sum_{j,k=1}^{n} \alpha_j \overline{\alpha_k} g_l(\cdot - (x_j - x_k)) \Big)^{\wedge}(\omega).
\end{aligned}
$$

Thus, (2.5) follows from [25, p. 252]. Finally, we use Theorem A.4(iii), to get

$$
\begin{aligned}
\sum_{j,k=1}^{n} \alpha_j \overline{\alpha_k} \Phi(x_j - x_k) &= \sum_{j,k=1}^{n} \alpha_j \overline{\alpha_k} \lim_{l \to \infty} \int_{\mathbb{R}^m} \Phi(x) g_l(x - (x_j - x_k)) dx \\
&= \lim_{l \to \infty} \int_{\mathbb{R}^m} \Phi(x) \sum_{j,k=1}^{n} \alpha_j \overline{\alpha_k} g_l(x - (x_j - x_k)) dx \\
&= \lim_{l \to \infty} \int_{\mathbb{R}^m} \Phi(x) (\gamma_l * \widetilde{\gamma}_l)(x) dx \\
&= \lim_{l \to \infty} \int_{\mathbb{R}^m} \int_{\mathbb{R}^m} \Phi(x - y) \gamma_l(x) \overline{\gamma_l(y)} dx dy \geq 0.
\end{aligned}
$$

Hence, Φ is positive definite on $\mathbb{R}^m$. $\square$

2.3. Bochner's Characterization of Positive Definite Functions

Bochner's characterization of positive definite functions is based on Fourier transforms. More precisely, if Φ is continuous and integrable on $\mathbb{R}^m$ with an integrable Fourier transform $\widehat{\Phi}$, then the Fourier inversion formula (A.2) entails

$$\Phi(x) = (2\pi)^{-m/2} \int_{\mathbb{R}^m} \widehat{\Phi}(\omega) e^{ix^{\mathrm{T}}\omega} d\omega.$$

This means that a quadratic form involving Φ can be expressed as

$$\begin{aligned} \sum_{j,k=1}^{n} \alpha_j \overline{\alpha_k} \Phi(x_j - x_k) &= (2\pi)^{-m/2} \sum_{j,k=1}^{n} \alpha_j \overline{\alpha_k} \int_{\mathbb{R}^m} \widehat{\Phi}(\omega) e^{i\omega^{\mathrm{T}}(x_j - x_k)} d\omega \\ &= (2\pi)^{-m/2} \int_{\mathbb{R}^m} \widehat{\Phi}(\omega) \Big| \sum_{j=1}^{n} \alpha_j e^{ix_j^{\mathrm{T}}\omega} \Big|^2 d\omega. \end{aligned}$$

Hence, if $\widehat{\Phi}$ is nonnegative then the function Φ is positive definite. In fact, every positive definite and integrable function has an integrable Fourier transform, [26]. For a non-integrable function, we will characterize every continuous positive definite function as the Fourier transform of a nonnegative finite Borel measure μ.

Theorem 2.10 (Bochner)**.** *A continuous function $\Phi : \mathbb{R}^m \to \mathbb{C}$ is positive definite on $\mathbb{R}^m$ if and only if it is the Fourier transform of a finite nonnegative Borel measure μ on $\mathbb{R}^m$, i.e.,*

$$\Phi(x) = \widehat{\mu}(x) = (2\pi)^{-m/2} \int_{\mathbb{R}^m} e^{-ix^{\mathrm{T}}\omega} d\mu(\omega), \quad x \in \mathbb{R}^m.$$

We do not include the proof of this standart result here. The interested reader should consult on [27]for Bochner's original proof, and on [28] or [29] for alternative approaches. Several proofs for this theorem can be found in the literature. Bochner's original proof can be found in [27]. A proof that uses the Riesz representation theorem to interpret Borel measures as distributions, and that takes advantage of distributional Fourier transforms can be found in [1].

2.4. Schoenberg's Characterization of Positive Definite Functions

In the previous section we saw that translation invariant positive definite functions can be characterized via Fourier transforms. Since Fourier transforms are not always easy to compute, we now present an alternative criteria that allows to decide whether a function is positive definite and radial on $\mathbb{R}^m$. We start with a brief review of basic facts about completely monotone functions. For details, see [1].

Definition 2.11. *A function $\phi \in C^\infty(0,\infty)$ is called completely monotone on $(0,\infty)$ if*

$$(-1)^n \phi^{(n)}(r) \geq 0,$$

for all $n = 0, 1, 2...$ and all $r > 0$. If, in addition $\phi \in [0,\infty)$, then it is called completely monotone on $[0,\infty)$.

Example 2.12. Some standart examples of completely monotone functions on $[0,\infty)$ are

i. $\phi(r) = c, c \geq 0$,

ii. $\phi(r) = e^{-\alpha r}, \alpha \geq 0$, since for $n = 0, 1, 2...$

$$(-1)^n \phi^{(n)}(r) = \alpha^n e^{-\alpha r} \geq 0,$$

iii. $\phi(r) = \frac{1}{(1+r)^\beta}, \beta \geq 0$, since for $n = 0, 1, 2...$

$$(-1)^n \phi^{(n)}(r) = (-1)^{2n} \beta(\beta+1)\cdots(\beta+n-1)(1+r)^{-\beta-n} \geq 0.$$

A characterization of positive definite functions can be given based on iterated forward differences.

Definition 2.13. *Let $k \in \mathbb{N}$. Suppose that $\{f_j\}_{j\in\mathbb{N}}$ is a sequence of real numbers.*

The kth-order iterated forward difference is

$$\Delta^k\{f_j\}(n) \equiv \Delta^k f_n := \sum_{j=0}^{k}(-1)^{k-j}\binom{k}{j} f_{n+j}, \quad \textit{for} \quad n = 0, 1, 2.$$

For a function $\phi : [0, \infty) \to \mathbb{R}$ *we define the kth-order difference by*

$$\Delta_h^k \phi(r) := \sum_{j=0}^{k}(-1)^{k-j}\binom{k}{j}\phi(r + jh), \tag{2.6}$$

for any $r \geq 0$ *and* $h > 0$*. If* ϕ *is defined only on* $(0, \infty)$ *then we restrict* r *in (2.6) to* $r > 0$.

Now, we are ready to state our theorem.

Theorem 2.14. *For a function* $\phi : (0, \infty) \to \mathbb{R}$ *the following statements are equivalent:*

i. ϕ *is completely monotone on* $(0, \infty)$*;*

ii. ϕ *satisfies* $(-1)^n \Delta_h^n \phi(r) \geq 0$ *for all* $r, h > 0$ *and* $n = 0, 1, 2, ...$

Similar to positive definite functions, completely monotone functions have an integral representation.

Theorem 2.15 (Hausdorff-Bernstein-Widder, [1])**.** *A function* $\phi : [0, \infty) \to \mathbb{R}$ *is completely monotone on* $[0, \infty)$ *if and only if it is the Laplace transform of a nonnegative finite Borel measure* ν*, i.e. it is of the form*

$$\phi(r) = \mathcal{L}\nu(r) = \int_0^\infty e^{-rt} d\nu(t). \tag{2.7}$$

Widder's proof of this theorem can be found in ([30], p.160), where he reduces the proof of this theorem to another theorem by Hausdorff on completely monotone sequences. A detailed proof can also be found in [1].

Having established that completely monotone functions are nothing other than Laplace transforms of nonnegative finite Borel measures, we turn to the connection between positive definite radial functions and completely monotone functions.

Theorem 2.16 (Schoenberg, [24]). *A function ϕ is completely monotone on $[0,\infty)$ if and only if $\Phi := \phi(\| \cdot \|^2)$ is positive definite on every $\mathbb{R}^m$.*

Proof. If ϕ is completely monotone on $[0,\infty)$ then Theorem 2.15 implies that there exists a nonnegative finite Borel measure ν such that

$$\phi(r) = \mathcal{L}\nu(r) = \int_0^\infty e^{-rt} d\nu(t).$$

Therefore, $\Phi(x) := \phi(\| x \|^2)$ has the representation

$$\Phi(x) = \int_0^\infty e^{-t\|x\|^2} d\nu(t).$$

Then for arbitrary $x_1, ..., x_n \in \mathbb{R}^m$ and $\alpha \in \mathbb{R}^n$, we have

$$\sum_{j,k=1}^{n} \alpha_j \alpha_k \Phi(\| x_j - x_k \|^2) = \int_0^\infty \sum_{j,k=1}^{n} \alpha_j \alpha_k e^{-t\|x_j - x_k\|^2} d\nu(t) \geq 0,$$

because the Gaussians involved are positive definite on every $\mathbb{R}^m$ (see Example 2.8) and the measure is nonnegative and finite. Conversely, suppose that $\phi(\| \cdot \|^2)$ is positive definite on every $\mathbb{R}^m$. Since ϕ is continuous at zero, by Theorem 2.14, it suffices to show that $(-1)^k \Delta_h^k \phi(r) \geq 0$ for all $r, h > 0$ and $k = 0, 1, 2, ...$ where $\Delta_h^k \phi(r)$ is the kth-order difference defined in (2.6). This can be done by induction on k.

For $k = 0$ we have to show that $\phi(r) \geq 0$ for all $r \in (0,\infty)$. To this end, we choose $x_j = \sqrt{r/2} e_j$, $1 \leq j \leq n$, where e_j denotes the jth unit coordinate vector in

$\mathbb{R}^n$. Since $\phi(\| \cdot \|^2)$ is positive definite on every $\mathbb{R}^n$ we get

$$0 \leq \sum_{j,k=1}^{n} \phi(\| x_j - x_k \|^2) = n\phi(0) + n(n-1)\phi(r)$$

because $\| x_j - x_k \|^2 = r$ for $j \neq k$. Dividing by $n(n-1)$ and letting $n \to \infty$ allows us to conclude that $\phi(r) \geq 0$.

For the induction step, we assume $(-1)^k \Delta_h^k \phi(r) \geq 0$ for all $r, h > 0$ and $k = 0, 1, 2, ...$ and want to show that $(-1)^{k+1} \Delta_h^{k+1} \phi(r) \geq 0$. We know that a positive definite function is nonnegative. Thus, for the induction step, it suffices to show that if $\phi(\| r \|^2)$ is positive definite on every $\mathbb{R}^m$, then $-\Delta_h^1 \phi(r)$ is also positive definite on every $\mathbb{R}^m$. Since then we can conclude that $(-1)^{k+1} \Delta_h^{k+1} \phi(r)$ is positive definite. To do this, suppose that $x_1, ..., x_n \in \mathbb{R}^m$ and $\alpha \in \mathbb{R}^n$ are given. We take the x_j as elements of $\mathbb{R}^{m+1}$ and define

$$y_j = \begin{cases} x_j & \text{if} \quad 1 \leq j \leq n \\ x_{j-n} + \sqrt{h} e_{m+1} & \text{if} \quad n < j \leq 2n \end{cases}$$

and

$$\beta_j = \begin{cases} \alpha_j & \text{if} \quad 1 \leq j \leq n \\ -\alpha_{j-n} & \text{if} \quad n < j \leq 2n. \end{cases}$$

Since $\phi(\| \cdot \|^2)$ is also positive definite on $\mathbb{R}^{m+1}$ we have

$$
\begin{aligned}
0 &\leq \sum_{j,k=1}^{2n} \beta_j \beta_k \phi(\| y_j - y_k \|^2) \\
&= \sum_{j,k=1}^{n} \alpha_j \alpha_k \phi(\| x_j - x_k \|^2) - \sum_{j=1}^{n} \sum_{k=n+1}^{2n} \alpha_j \alpha_{k-n} \phi(\| x_j - x_{k-n} \|^2 + h) \\
&- \sum_{j=n+1}^{2n} \sum_{k=1}^{n} \alpha_{j-n} \alpha_k \phi(\| x_{j-n} - x_k \|^2 + h) + \sum_{j,k=n+1}^{2n} \alpha_{j-n} \alpha_{k-n} \phi(\| x_{j-n} - x_{k-n} \|^2 + h) \\
&= 2 \sum_{j,k=1}^{n} \alpha_j \alpha_k [\phi(\| x_j - x_k \|^2) - \phi(\| x_j - x_k \|^2 + h)] \\
&= -2 \sum_{j,k=1}^{n} \alpha_j \alpha_k \Delta_h^1 \phi(\| \cdot \|^2).
\end{aligned}
$$

Thus, $-\Delta_h^1 \phi(r)$ is also positive definite on every $\mathbb{R}^m$. $\square$

3. ULTRASPHERICAL POLYNOMIALS

In this chapter, our main concern is the addition formula for ultraspherical (or Gegenbauer) polynomials. This formula plays a fundamental role in Schoenberg's characterization of positive definite functions. Explicitly it is

$$\begin{aligned} P_n^{(\lambda)}(\cos\theta_1\cos\theta_2+\sin\theta_1\sin\theta_2\cos\phi) &= P_n^{(\lambda)}(\cos\theta_1)P_n^{(\lambda)}(\cos\theta_2) \\ &+\sum_{s=1}^{n} c_{s,n}^{\lambda} P_{n-s}^{(\lambda+s)}(\cos\theta_1)(\sin\theta_1)^s P_{n-s}^{(\lambda+s)}(\cos\theta_2)(\sin\theta_2)^s P_s^{(\lambda-\frac{1}{2})}(\cos\phi) \end{aligned} \tag{3.1}$$

with

$$c_{s,n}^{\lambda} = \frac{\Gamma(\lambda-\frac{1}{2})}{\Gamma(\lambda)}\frac{\Gamma(\lambda+s)\Gamma(2\lambda+2s)(n-s)!}{\Gamma(\lambda+s-\frac{1}{2}))\Gamma(2\lambda+s+n)} \tag{3.2}$$

where $0 \leq \theta_1, \theta_2, \phi\pi$, and $P_n^{(\lambda)}$ is the ultraspherical polynomial of degree n and order λ. The proof of the addition formula requires information about the theory of ultraspherical polynomials, so we firstly introduce them. Then, we give the proof following [22]. We conclude this chapter with classical results on ultraspherical expansions.

3.1. Definition and Basic Properties of Ultraspherical Polynomials

Ultraspherical polynomials are special cases of *Jacobi* polynomials and generalization of *Chebyshev* and *Legendre* polynomials. Jacobi polynomials $P_n^{(\alpha,\beta)}(x)$ can be defined on $[-1,1]$ by

$$(1-x)^{\alpha}(1+x)^{\beta}P_n^{(\alpha,\beta)}(x) = \frac{(-1)^n}{2^n n!}D_x^n[(1-x)^{n+\alpha}(1+x)^{n+\beta}]. \tag{3.3}$$

They are orthogonal on $[-1,1]$ with the weight function $w(x) = (1-x)^{\alpha}(1+x)^{\beta}$. Integrability of $w(x)$ is achieved by requiring that $\alpha > -1, \beta > -1$ [31].

For $\alpha = \beta$, Jacobi polynomials are called ultraspherical polynomials. The following is the customary notation and normalization [32]:

$$\begin{aligned}
P_n^{(\lambda)}(x) &= \frac{\Gamma(\alpha+1)}{\Gamma(2\alpha+1)} \frac{\Gamma(n+2\alpha+1)}{\Gamma(n+\alpha+1)} P_n^{(\alpha,\alpha)}(x) \\
&= \frac{\Gamma(\lambda+\frac{1}{2})}{\Gamma(2\lambda)} \frac{\Gamma(n+2\lambda)}{\Gamma(n+\lambda+\frac{1}{2})} P_n^{(\lambda-\frac{1}{2},\lambda-\frac{1}{2})}(x), \quad \alpha = \lambda - \frac{1}{2}.
\end{aligned} \tag{3.4}$$

Hence, for ultraspherical polynomials, (3.3) becomes

$$(1-x^2)^{\lambda-\frac{1}{2}} P_n^{(\lambda)}(x) = \frac{(-2)^n \Gamma(n+\lambda)\Gamma(n+2\lambda)}{n!\Gamma(\lambda)\Gamma(2n+2\lambda)} D_x^n[(1-x^2)^{\lambda+n-\frac{1}{2}}]. \tag{3.5}$$

where we have used the following well known property of the Gamma function (see (B.5))

$$\Gamma(\lambda)\Gamma(\lambda+\frac{1}{2}) = \sqrt{\pi} 2^{1-2\lambda} \Gamma(2\lambda). \tag{3.6}$$

Moreover, ultraspherical polynomials are orthogonal on $[-1, 1]$ with weight $(1-x^2)^{\lambda-\frac{1}{2}}$ for $\lambda > -\frac{1}{2}$ (see §3.4).

Alternatively, we can define the ultraspherical polynomials $P_n^{(\lambda)}(x)$ as the coefficients in the expansion

$$(1-2rx+r^2)^{-\lambda} = \sum_{n=0}^{\infty} P_n^{(\lambda)}(x) r^n, \quad (\lambda > 0, \quad x \in [-1,1], \quad r \in (-1,1)), \tag{3.7}$$

so that the function on the left is the generating function of the polynomials $P_n^{(\lambda)}(x)$. Several essential formulas concerning ultraspherical polynomials are given in Appendix C.

3.2. Special Cases of Ultraspherical Polynomials

The simplest case of ultraspherical polynomials is when $\lambda = 0$. In this case, ultraspherical polynomials are called Chebyshev polynomials of the first kind and they are denoted by $T_n(x)$. That is,

$$T_n(\cos\theta) = P_n^{(0)}(\cos\theta) = \cos n\theta. \tag{3.8}$$

They satisfy the orthogonality condition

$$\int_{-1}^{1} T_n(x)T_m(x)(1-x^2)^{-\frac{1}{2}}dx = \int_0^{\pi} \cos n\theta \cos m\theta d\theta = 0,$$

for $n \neq m$. For this case, the addition formula for ultraspherical polynomials (3.1) reduces to the well known cosine addition formula

$$\cos n(\theta_1 - \theta_2) = \cos n\theta_1 \cos n\theta_2 + \sin n\theta_1 \sin n\theta_2 \tag{3.9}$$

which can be seen by letting $\phi = 0$ and $\lambda \to 0$ in (3.1) [31]. Also, (3.9) explains why (3.1) is called the addition formula.

For $\lambda = \frac{1}{2}$, ultraspherical polynomials $P_n^{(\lambda)}(x)$ are called the Legendre polynomials which are denoted by $P_n(x)$.

All formulas and theorems concerning Legendre polynomials can be obtained immediately from the theory of ultraspherical polynomials by setting $\lambda = \frac{1}{2}$. A remarkable property of Legendre polynomials is the simple orthogonality relation

$$\int_{-1}^{1} P_n(x)P_m(x)dx = \frac{2}{2n+1}\delta_{nm}$$

where δ_{nm} is the Kronecker delta. In addition, (3.5) simplifies to the Rodrigues formula

$$P_n(x) = \frac{1}{2^n n!} D_x^n (x^2 - 1)^n$$

which, by integration by parts, yields the Rodrigues rule

$$\int_{-1}^{1} f(x) P_n(x) dx = \frac{1}{2^n n!} \int_{-1}^{1} f^{(n)}(x)(1 - x^2)^n dx$$

for every $f(x) \in C^{(n)}[-1, 1]$ [17].

For Legendre polynomials, the addition formula is obtained by taking limit as $\lambda \to \frac{1}{2}$ in (3.1).

$$\begin{aligned} P_n(\cos\theta_1 \cos\theta_2 + \sin\theta_1 \sin\theta_2 \cos\phi) &= P_n(\cos\theta_1) P_n(\cos\theta_2) \\ &\quad + 2 \sum_{s=1}^{n} \frac{(n-s)!}{(n+s)!} P_n^s(\cos\theta_1) P_n^s(\cos\theta_2) \cos(s\phi) \end{aligned} \tag{3.10}$$

where $P_n^s(x)$ is the associated Legendre function defined by

$$P_n^s(x) = (-1)^s (1 - x^2)^{s/2} D_x^s P_n(x), \quad -1 < x < 1.$$

Legendre polynomials are especially important since they appear in Schoenberg's characterization of positive definite functions on the 2-dimensional unit sphere $\mathcal{S}^2$. For the addition formula for Jacobi polynomials, we refer the reader to [33], for example.

3.3. Addition Formula for Ultraspherical Polynomials

In this section, we will prove the addition formula for ultraspherical polynomials (Theorem 3.2). For the proof, we need the following lemma:

Lemma 3.1. *Let $x, \alpha, \beta \in [-1, 1]$, and $\nu \geq 0$. Then for $\omega = \alpha\beta + x\sqrt{\alpha^2 - 1}\sqrt{\beta^2 - 1}$*

such that $\omega \in [-1,1]$, $y = P_n^{(\nu)}(\omega)$ *satisfies the partial differential equation*

$$(1-\alpha^2)\frac{\partial^2 y}{\partial \alpha^2}-(1+2\nu)\alpha\frac{\partial y}{\partial \alpha}+n(n+2\nu)y+\frac{1}{1-\alpha^2}\left((1-x^2)\frac{\partial^2 y}{\partial x^2}-2\nu x\frac{\partial y}{\partial x}\right)=0. \quad (3.11)$$

Proof. Differentiating $y = P_n^{(\nu)}(\omega)$ two times with respect to α and with respect to x, we have

$$\begin{aligned}\frac{\partial y}{\partial \omega} &= \frac{\partial y}{\partial \alpha}\frac{\sqrt{\alpha^2-1}}{\eta}, \quad \frac{\partial^2 y}{\partial \omega^2} = \frac{\alpha^2-1}{\eta^2}\frac{\partial^2 y}{\partial \alpha^2}+\frac{x\sqrt{\beta^2-1}}{\eta^3}\frac{\partial y}{\partial \alpha}, \\ \frac{\partial y}{\partial x} &= \frac{(\alpha^2-1)\sqrt{\beta^2-1}}{\eta}\frac{\partial y}{\partial \alpha}, \quad \frac{\partial^2 y}{\partial x^2} = (\alpha^2-1)(\beta^2-1)\left(\frac{\alpha^2-1}{\eta^2}\frac{\partial^2 y}{\partial \alpha^2}+\frac{x\sqrt{\beta^2-1}}{\eta^3}\frac{\partial y}{\partial \alpha}\right)\end{aligned} \quad (3.12)$$

where we put for abbreviation $\eta = \beta\sqrt{\alpha^2-1}+\alpha x\sqrt{\beta^2-1}$, which gives $1-\omega^2 = -\eta^2+(x^2-1)(\beta^2-1)$. Then, utilizing the Gegenbauer differential equation (C.7) for $y = P_n^{(\nu)}(\omega)$, we have

$$(1-\omega^2)\frac{\partial^2 y}{\partial \omega^2}-(1+2\nu)\omega\frac{\partial y}{\partial \omega}+n(n+2\nu)y = 0. \quad (3.13)$$

Inserting the differential formulas (3.12) into (3.13) gives immediately (3.11). $\square$

Theorem 3.2 (Addition Formula)**.** *Let* $x, \alpha, \beta \in [-1,1]$, *and* $\nu \geq 0$. *Then for* $\omega = \alpha\beta + x\sqrt{\alpha^2-1}\sqrt{\beta^2-1}$ *such that* $\omega \in [-1,1]$, *we have*

$$\begin{aligned}P_n^{(\nu+\frac{1}{2})}(\omega) &= \sum_{s=0}^{n}\frac{\Gamma(\nu)}{\Gamma(\nu+\frac{1}{2})}\frac{\Gamma(\nu+s+\frac{1}{2})\Gamma(2\nu+2s+1)(n-s)!}{\Gamma(\nu+s)\Gamma(2\nu+s+n+1)} \quad (3.14)\\ &\times (\alpha^2-1)^{s/2}P_{n-s}^{(\nu+s+\frac{1}{2})}(\alpha)(\beta^2-1)^{s/2}P_{n-s}^{(\nu+s+\frac{1}{2})}(\beta)P_s^{(\nu)}(x). \quad (3.15)\end{aligned}$$

Proof. Consider the ultraspherical function

$$y = P_n^{(\nu+\frac{1}{2})}(\omega), \quad \omega = \alpha\beta + x\sqrt{\alpha^2-1}\sqrt{\beta^2-1}. \quad (3.16)$$

We can write it as a linear combination of the ultraspherical polynomials having degree less or equal to n. That is,

$$\Gamma(\nu+\frac{1}{2})P_n^{(\nu+\frac{1}{2})}(\alpha\beta+x\sqrt{\alpha^2-1}\sqrt{\beta^2-1})=\Gamma(\nu)\sum_{s=0}^{n}A_s^{\nu,n}(\alpha,\beta)P_s^{(\nu)}(x), \tag{3.17}$$

for some $A_s^{\nu,n}(\alpha,\beta)$ which are symmetric with respect to two variables α and β, i.e.

$$A_s^{\nu,n}(\alpha,\beta)=A_s^{\nu,n}(\beta,\alpha). \tag{3.18}$$

We need to find the coefficients $A_s := A_s^{\nu,n}$. To this end, using the differential formula (C.5) we first differentiate both sides of (3.17) with respect to x. Then, we have

$$\sqrt{\alpha^2-1}\sqrt{\beta^2-1}\Gamma(\nu+\frac{1}{2})2(\nu+\frac{1}{2})P_{n-1}^{(\nu+\frac{3}{2})}(\omega)=\Gamma(\nu)2\nu\sum_{s=0}^{n}A_s^{\nu,n}(\alpha,\beta)P_{s-1}^{(\nu+1)}(x).$$

Since $\nu\Gamma(\nu)=\Gamma(\nu+1)$ and $P_{-1}^{(\nu+1)}(x)=0$, we have

$$\sqrt{\alpha^2-1}\sqrt{\beta^2-1}\Gamma(\nu+\frac{3}{2})P_{n-1}^{(\nu+\frac{3}{2})}(\omega)=\Gamma(\nu+1)\sum_{s=0}^{n-1}A_{s+1}^{\nu,n}(\alpha,\beta)P_s^{(\nu+1)}(x).$$

Also, putting $\nu+1$ instead of ν and $n-1$ instead of n in (3.17), we get

$$\Gamma(\nu+\frac{3}{2})P_{n-1}^{(\nu+\frac{3}{2})}(\omega)=\Gamma(\nu+1)\sum_{s=0}^{n-1}A_s^{\nu+1,n-1}(\alpha,\beta)P_s^{(\nu+1)}(x).$$

Now, combining these two identities, we obtain the recursive formula

$$A_s^{\nu,n}(\alpha,\beta)=\sqrt{\alpha^2-1}\sqrt{\beta^2-1}A_{s-1}^{\nu+1,n-1}(\alpha,\beta)$$

which gives immediately

$$A_s^{\nu,n}(\alpha,\beta)=(\alpha^2-1)^{s/2}(\beta^2-1)^{s/2}A_0^{\nu+s,n-s}(\alpha,\beta) \tag{3.19}$$

so that it remains only to determine the coefficients $A_0^{\nu,n}(\alpha,\beta)$. Now, by Lemma 3.1, we know that $P_n^{(\nu+\frac{1}{2})}(\omega)$ satisfies

$$(1-\alpha^2)D_\alpha^2 P_n^{(\nu+\frac{1}{2})}(\omega) - (2+2\nu)\alpha D_\alpha P_n^{(\nu+\frac{1}{2})}(\omega) + n(n+2\nu+1)P_n^{(\nu+\frac{1}{2})}(\omega)$$

$$+\frac{1}{1-\alpha^2}\Big((1-x^2)D_x^2 P_n^{(\nu+\frac{1}{2})}(\omega) - (2\nu+1)xD_x P_n^{(\nu+\frac{1}{2})}(\omega)\Big) = 0.$$

Writing the right hand side of (3.17) instead of $P_n^{(\nu+\frac{1}{2})}(\omega)$, we obtain

$$\sum_{s=0}^{n} P_s^{(\nu)}(x)\big((1-\alpha^2)D_\alpha^2 A_s - (2+2\nu)\alpha D_\alpha A_s + n(n+2\nu+1)A_s\big)$$

$$+\frac{A_s}{1-\alpha^2}\Big((1-x^2)D_x^2 P_s^{(\nu)}(x) - (2\nu+1)xD_x P_s^{(\nu)}(x)\Big) = 0.$$

Using (C.7) for $P_s^{(\nu)}(x)$, we have

$$(1-x^2)D_x^2 P_s^{(\nu)}(x) - (2\nu+1)xD_x P_s^{(\nu)}(x) = -s(s+2\nu)P_s^{(\nu)}(x),$$

thus we get

$$\sum_{s=0}^{n} P_s^{(\nu)}(x)\Big[(1-\alpha^2)D_\alpha^2 A_s - (2+2\nu)\alpha D_\alpha A_s + [n(n+2\nu+1) - \frac{s(s+2\nu)}{1-\alpha^2}]A_s\Big] = 0.$$

But this implies that for all $s = 0, 1, ..., n$,

$$(1-\alpha^2)D_\alpha^2 A_s - (2+2\nu)\alpha D_\alpha A_s + [n(n+2\nu+1) - \frac{s(s+2\nu)}{1-\alpha^2}]A_s = 0.$$

In particular, for $s = 0$,

$$(1-\alpha^2)D_\alpha^2 A_0 - (2+2\nu)\alpha D_\alpha A_0 + n(n+2\nu+1)A_0 = 0.$$

By symmetry, (3.18), we also have

$$(1-\beta^2)D_\beta^2 A_0 - (2+2\nu)\beta D_\beta A_0 + n(n+2\nu+1)A_0 = 0.$$

Now, observe that $P_n^{(\nu+\frac{1}{2})}(\alpha)$ satisfies the first equation while $P_n^{(\nu+\frac{1}{2})}(\beta)$ satisfies the second one. Hence, A_0 should be in the form

$$A_0^{\nu,n}(\alpha,\beta) = a_0^{\nu,n} P_n^{(\nu+\frac{1}{2})}(\alpha) P_n^{(\nu+\frac{1}{2})}(\beta) \tag{3.20}$$

where the coefficients $a_0^{\nu,n}$ are independent of α and β. Now we need to find $a_0^{\nu,n}$, so let $\alpha=\beta=1$. Since (3.19) implies that $A_s^{\nu,n}(1,1)=0$ if $s>0$, by (3.17) we get

$$\Gamma(\nu+\frac{1}{2})P_n^{(\nu+\frac{1}{2})}(1) = \Gamma(\nu)A_0^{\nu,n}(1,1)P_0^{(\nu)}(x).$$

Then, it follows from (C.1), (3.20), and (C.4) that

$$\begin{aligned} a_0^{\nu,n} &= \frac{\Gamma(\nu+\frac{1}{2})}{\Gamma(\nu)P_n^{(\nu+\frac{1}{2})}(1)} \\ &= \frac{\Gamma(\nu+\frac{1}{2})\Gamma(2\nu+1)n!}{\Gamma(\nu)\Gamma(2\nu+n+1)}. \end{aligned}$$

Now, inserting (3.20) in the recursive formula (3.19), we get

$$\begin{aligned} A_s^{\nu,n}(\alpha,\beta) &= (\alpha^2-1)^{s/2}(\beta^2-1)^{s/2}A_0^{\nu+s,n-s}(\alpha,\beta) \\ &= (\alpha^2-1)^{s/2}(\beta^2-1)^{s/2}a_0^{\nu+s,n-s}P_{n-s}^{(\nu+s+\frac{1}{2})}(\alpha)P_{n-s}^{(\nu+s+\frac{1}{2})}(\beta) \\ &= \frac{\Gamma(\nu+s+\frac{1}{2})\Gamma(2\nu+2s+1)(n-s)!}{\Gamma(\nu+s)\Gamma(2\nu+s+n+1)} \\ &\times (\alpha^2-1)^{s/2}P_{n-s}^{(\nu+s+\frac{1}{2})}(\alpha)(\beta^2-1)^{s/2}P_{n-s}^{\nu+s+\frac{1}{2}}(\beta). \end{aligned}$$

Writing this in (3.17), we finally obtain (3.14). □

3.4. Orthogonality of Ultraspherical Polynomials

In this section, we will show that ultraspherical polynomials are orthogonal with respect to the weight function

$$h(x) = (1-x^2)^{\lambda-1/2}$$

on the interval $[-1,1]$ for $\lambda > -\frac{1}{2}$. We begin with a more general result that shall also needed later on in §3.5.

Proposition 3.3. *Let $f(x)$ be an n-times continuously differentiable function on $[-1,1]$. For $\lambda > -\frac{1}{2}$, we have*

$$\int_{-1}^{1} f(x)P_n^{(\lambda)}(x)(1-x^2)^{\lambda-\frac{1}{2}}dx = \frac{2^n\Gamma(n+\lambda)\Gamma(n+2\lambda)}{n!\Gamma(\lambda)\Gamma(2n+2\lambda)} \int_{-1}^{1} f^{(n)}(x)(1-x^2)^{\lambda+n-\frac{1}{2}}dx. \tag{3.21}$$

Proof. Using (3.5), we have

$$\begin{aligned}\int_{-1}^{1} & f(x)P_n^{(\lambda)}(x)(1-x^2)^{\lambda-\frac{1}{2}}dx \\ &= \frac{(-2)^n\Gamma(n+\lambda)\Gamma(n+2\lambda)}{n!\Gamma(\lambda)\Gamma(2n+2\lambda)} \int_{-1}^{1} f(x)D_x^n[(1-x^2)^{\lambda+n-\frac{1}{2}}]dx \\ &= \frac{(-2)^n\Gamma(n+\lambda)\Gamma(n+2\lambda)}{n!\Gamma(\lambda)\Gamma(2n+2\lambda)} \int_{-1}^{1} f^{(n)}(x)(1-x^2)^{\lambda+n-\frac{1}{2}}dx\end{aligned}$$

where the last equality follows by integration by parts n-times. □

Now, taking $f(x) = P_n^{(\lambda)}(x)$ in (3.21), by (C.6) we obtain

Corollary 3.4.

$$\int_{-1}^{1} P_n^{(\lambda)}(x)P_r^{(\lambda)}(x)(1-x^2)^{\lambda-\frac{1}{2}}dx = \begin{cases} 0 & \text{if} \quad r \neq n \\ \frac{\pi\Gamma(n+2\lambda)}{2^{2\lambda-1}(\lambda+n)n![\Gamma(\lambda)]^2} & \text{if} \quad r = n. \end{cases} \tag{3.22}$$

This shows that ultraspherical polynomials are indeed orthogonal with respect to the weight function $h(x) = (1-x^2)^{\lambda-1/2}$ on the interval $[-1,1]$ for $\lambda > -1/2$.

Next, let

$$f(x) = \sum_{k=0}^{n} a_k P_k^{(\lambda)}(x).$$

Then, by the orthogonality property (3.22), we have

$$\int_{-1}^{1} f(x) P_s^{(\lambda)}(x)(1-x^2)^{\lambda-\frac{1}{2}} dx = \frac{\pi\Gamma(s+2\lambda)}{2^{2\lambda-1}(\lambda+s)s![\Gamma(\lambda)]^2} a_s,$$

where

$$a_s = \frac{2^{s+2\lambda-1}\Gamma(s+\lambda+1)\Gamma(\lambda)}{\pi\Gamma(2s+2\lambda)} \int_{-1}^{1} f^{(s)}(x)(1-x^2)^{\lambda+n-\frac{1}{2}} dx$$

can be obtained using (3.21).

In particular, if $a_0 = 1$ and $a_k = 0$ for $1 \le k \le n$, then $f(x) = 1$ since $P_0^{(\lambda)}(x) = 1$. Hence,

Corollary 3.5.

$$\int_{-1}^{1} P_s^{(\lambda)}(x)(1-x^2)^{\lambda-\frac{1}{2}} dx = \begin{cases} 0 & \textit{if} \quad s > 0 \\ \frac{\pi\Gamma(2\lambda)}{2^{2\lambda-1}(\lambda)[\Gamma(\lambda)]^2} & \textit{if} \quad s = 0. \end{cases} \tag{3.23}$$

3.5. Expansion of Functions in Series of Ultraspherical Polynomials

Suppose we are given a function f(x) which is continuous in the closed interval $[-1,1]$ and we want to expand this in an infinite series of $P_n^{(\lambda)}(x)$. If

$$f(x) = \sum_{n=0}^{\infty} a_n P_n^{(\lambda)}(x),$$

then from the orthogonality property (3.22), we have

$$a_n = \frac{\Gamma(2\lambda)(n+\lambda)n!\Gamma(\lambda)}{\sqrt{\pi}\Gamma(n+2\lambda)\Gamma(\lambda+\frac{1}{2})}\int_{-1}^{1} f(x)P_n^{(\lambda)}(x)(1-x^2)^{\lambda-\frac{1}{2}}dx. \tag{3.24}$$

Thus, if the integral on the right hand side of (3.24) exists, we can determine the coefficients a_n. For instance, if the given function $f(x)$ is analytic on $[-1,1]$, then Lemma 3.21 implies that

$$\begin{aligned} a_n &= \frac{\Gamma(2\lambda)(n+\lambda)n!\Gamma(\lambda)}{\sqrt{\pi}\Gamma(n+2\lambda)\Gamma(\lambda+\frac{1}{2})}\int_{-1}^{1} f(x)P_n^{(\lambda)}(x)(1-x^2)^{\lambda-\frac{1}{2}}dx \\ &= \frac{\Gamma(2\lambda)(n+\lambda)}{\sqrt{\pi}\Gamma(\lambda+\frac{1}{2})}\frac{2^n\Gamma(n+\lambda)}{\Gamma(2n+2\lambda)}\int_{-1}^{1} f^{(n)}(x)(1-x^2)^{\lambda+n-\frac{1}{2}}dx. \end{aligned}$$

For further details about the summability conditions see [32, p.243] and references therein.

In this section, we try to find ultraspherical expansion of a real continuous function $F(p)$ on $\mathcal{S}^m$. We do this by using Poisson integral which is defined as follows:

Definition 3.6 ([9])**.** *For an arbitrary point p of $\mathcal{S}^m$ and $0 \le r < 1$, let p_r be that point on the radius $\overline{op}$ such that $op_r = r$. Furthermore, let $F(p)$ be a real and continuous point function defined in $\mathcal{S}^m$. Then the Poisson integral is defined as*

$$F(p_r) = \frac{1}{\omega_m}\int_{\mathcal{S}^m} \frac{1-r^2}{(1-2r\cos pp' + r^2)^{\frac{1}{2}(m+1)}} F(p')d\omega_{p'}. \tag{3.25}$$

A classical result concerning the Poisson integral is

$$\lim_{r\to 1^-} F(p_r) = F(p). \tag{3.26}$$

The proof can be found in [34], for example.

Now, we first find ultraspherical development of $F(p_r)$, then using the definition of Abel summability given below and (3.26) we obtain ultraspherical development of $F(p)$.

Definition 3.7 ([25])**.** *If $\sum_{k=0}^{\infty} a_k$ is a series of complex numbers, for $0 < r < 1$ its r^{th} Abel mean is the series $\sum_{k=0}^{\infty} r^k a_k$. If the latter series converges for $r < 1$ to the sum $S(r)$ and the limit $S = \lim_{r\to 1^-} S(r)$ exists, then the series $\sum_{k=0}^{\infty} a_k$ is said to be Abel summable to S, and denoted as $S \sim \sum_{k=0}^{\infty} a_k$.*

Theorem 3.8. *The ultraspherical development of $F(p)$ is given by*

$$F(p) \sim \sum_{n=0}^{\infty} \frac{n+\lambda}{\lambda\omega_m} \int_{\mathcal{S}^m} F(p')P_n^{(\lambda)}(\cos pp')d\omega_{p'}. \tag{3.27}$$

Proof. Writing $x = \cos pp'$ in (3.7) and differentiating with respect to r, we get

$$\frac{1-r^2}{(1-2r\cos pp' + r^2)^{\lambda+1}} = \sum_{n=0}^{\infty}(\frac{n}{\lambda}+1)r^n P_n^{(\lambda)}(\cos pp'). \tag{3.28}$$

Let $\lambda = \frac{1}{2}(m-1)$. Multiplying (3.28) by $F(p')/\omega_m$ and integrating both sides over $\mathcal{S}^m$, in view of (3.25), we now get the expansion

$$\begin{aligned} F(p_r) &= \frac{1}{\omega_m}\int_{\mathcal{S}^m} F(p')\sum_{n=0}^{\infty}\frac{n+\lambda}{\lambda\omega_m} r^n P_n^{(\lambda)}(\cos pp')d\omega_p' \\ &= \sum_{n=0}^{\infty}\frac{n+\lambda}{\lambda\omega_m} r^n \int_{\mathcal{S}^m} F(p')P_n^{(\lambda)}(\cos pp')d\omega_p'. \end{aligned} \tag{3.29}$$

Hence, by (3.26), we have

$$F(p) = \lim_{r\to 1^-} F(p_r) = \lim_{r\to 1^-} \sum_{n=0}^{\infty} \frac{n+\lambda}{\lambda\omega_m} r^n \int_{\mathcal{S}^m} F(p') P_n^{(\lambda)}(\cos pp') d\omega_p'$$
$$= \sum_{n=0}^{\infty} \frac{n+\lambda}{\lambda\omega_m} \int_{\mathcal{S}^m} F(p') P_n^{(\lambda)}(\cos pp') d\omega_{p'}. \tag{3.30}$$

Hence, the ultraspherical expansion (3.30) is Abel summable at every point p of $\mathcal{S}^m$ to the sum $F(p)$. □

Now let F(p) be a real continuous function on $\mathcal{S}^m$. Assume further that F is a *zonal* function. That is, for an arbitrary point p' of $\mathcal{S}^m$, $F(p) = f(p \cdot p') = f(\cos pp')$ for some $f : [-1, 1] \to \mathbb{R}$ where pp' is the spherical distance between p and p'. Then, it is possible to find the ultraspherical expansion of such a function:

Theorem 3.9. *Let the function F given in Definition 3.6 be zonal. Consider the special case $F(p) = f(\cos\theta)$. Then the ultraspherical expansion of $f(\cos\theta)$ is given by*

$$f(\cos\theta) \sim \sum_{n=0}^{\infty} \frac{(n+\lambda)\Gamma(\lambda)\Gamma(n+1)\Gamma(2\lambda)}{\Gamma(\lambda+\frac{1}{2})\Gamma(\frac{1}{2})\Gamma(n+2\lambda)} P_n^{(\lambda)}(\cos\theta) \tag{3.31}$$
$$\times \int_0^{\pi} P_n^{(\lambda)}(\cos\theta') f(\cos\theta') \sin^{m-1}\theta' d\theta'.$$

Proof. Writing $f(\cos\theta)$ instead of $F(p')$ in (3.27), in view of (B.13), the integral on the right hand side of (3.27) becomes

$$I_n = \int_0^{\pi} \cdots \int_0^{\pi} \int_0^{2\pi} f(\cos\theta') P_n^{(\lambda)}(\cos pp') sin^{m-1}\theta' sin^{m-2}\theta_1' \ldots \sin\theta_{m-2}' d\phi' d\theta_1' \ldots d\theta_{m-2}' d\theta' \tag{3.32}$$

which can be written as

$$I_n = \int_0^{\pi} f(\cos\theta') J_n sin^{m-1}\theta' d\theta' \tag{3.33}$$

where we have set

$$J_n = \int_0^{\pi} \cdots \int_0^{\pi} \int_0^{2\pi} P_n^{(\lambda)}(\cos pp') \sin^{m-2}\theta_1' ... \sin\theta_{m-2}' d\phi' d\theta_1' ... d\theta_{m-2}'. \tag{3.34}$$

This integral is now readily reducible to a simple integral as follows. Consider the two points p_1 and p_1' given in polar coordinates

$$p_1 = (\frac{1}{2}\pi, \theta_1, ..., \theta_{m-2}, \phi), \quad p_1' = (\frac{1}{2}\pi, \theta_1', ..., \theta_{m-2}', \phi') \tag{3.35}$$

both lying in the unit sphere $\mathcal{S}^{m-1}$ defined by $\theta = \frac{1}{2}\pi$. Then,

$$\begin{aligned} \cos pp' = p \cdot p' &= \cos\theta\cos\theta' + \sin\theta\sin\theta' p_1 \cdot p_1' \\ &= \cos\theta\cos\theta' + \sin\theta\sin\theta'\cos p_1p_1'. \end{aligned} \tag{3.36}$$

We notice that J_n as given by (3.34) amounts to an integration of $P_n^{(\lambda)}(\cos pp')$ over $\mathcal{S}^{m-1}$. If we take in $\mathcal{S}^{m-1}$ a new system of polar coordinates ($\zeta = p_1p_1', \zeta_1, ..., \zeta_{m-3}, \psi$), of pole p_1, we obtain

$$J_n = \int_0^{\pi} \cdots \int_0^{\pi} \int_0^{2\pi} P_n^{(\lambda)}(\cos pp') \sin^{m-2}\zeta \sin^{m-3}\zeta_1 ... \sin\zeta_{m-3} d\psi d\zeta_1 ... d\zeta_{m-3} d\zeta. \tag{3.37}$$

Since pp' depends only on ζ, as shown by (3.36), the remaining integrations may be carried out leading to the expression

$$\begin{aligned} J_n &= \omega_{m-2}\int_0^{\pi} P_n^{(\lambda)}(\cos pp') \sin^{m-2}\zeta d\zeta \\ &= \omega_{m-2}\int_0^{\pi} P_n^{(\lambda)}(\cos pp') \sin^{2\lambda-1}\zeta d\zeta \\ &= \omega_{m-2}\frac{\Gamma(\lambda)\Gamma(\frac{1}{2})\Gamma(n+1)\Gamma(2\lambda)}{\Gamma(\lambda+\frac{1}{2})\Gamma(n+2\lambda)} P_n^{(\lambda)}(\cos\theta) P_n^{(\lambda)}(\cos\theta') \end{aligned}$$

where the last equality follows from the addition formula and (3.23). Hence, J_n is explicitly computed. Substituting its value into (3.33) we find that the general expansion (3.27) reduces to (3.31). This expression is also Abel-summable as shown in Theorem 3.8. $\square$

4. POSITIVE DEFINITE FUNCTIONS ON SPHERES

Radial basis function approximation is well suited in many branches of geosciences such as geodesy, meteorology, navigation etc. [17]. In such applications, we can use the unit sphere as a model for earth. When we approximate on $\mathcal{S}^m$ within $\mathbb{R}^{m+1}$, we no longer use the Euclidean norm in connection with a univariate radial function but apply the so called spherical (or geodesic) distances which is the length of the shorter path of the great circle joining the two points on the sphere.

Therefore, the characterizations of positive definite functions given in Chapter 2 no longer apply, and one has to study alternative concepts of positive definite functions on $\mathcal{S}^m$. These functions are often called *zonal* rather than *radial* basis functions since $\mathcal{S}^m$ is invariant under all orthogonal transformations. In this chapter, following Schoenberg's classical paper [9], we give characterizations of positive definite functions on finite and infinite dimensional unit spheres.

4.1. Positive Definite Functions on $\mathcal{S}^m$

In this section we present Schoenberg's characterization of positive definite functions on $\mathcal{S}^m$ as those ones whose expansions in series of ultraspherical polynomials always have nonnegative coefficients. We begin with two auxiliary results.

Lemma 4.1. *Let $f(\cos\theta)$ be real and continuous for $\theta \in [0,\pi]$ and such that $f(\cos\theta)$ is positive definite on $\mathcal{S}^m$. Then,*

$$\int_{\mathcal{S}_m} f(\cos pp')d\omega_{p'} \geq 0. \tag{4.1}$$

Proof. Positive definiteness of $f(\cos t)$ on $\mathcal{S}^m$ implies that

$$\sum_{j,k=1}^{n} f(\cos p_j p_k)\alpha_j\alpha_k \geq 0, \qquad (p_j \in \mathcal{S}^m, \quad \alpha_j \quad real).$$

By (2.4), this quadratic form is equivalent to the integral inequality

$$I(h) = \int_{\mathcal{S}_m} \int_{\mathcal{S}_m} f(\cos pp')h(p)h(p')d\omega_p d\omega_{p'} \geq 0, \tag{4.2}$$

for an arbitrary continuous function $h(p)$ in $\mathcal{S}^m$. For $h(p) \equiv 1$, (4.2) becomes

$$I(1) = \int_{\mathcal{S}_m} \left(\int_{\mathcal{S}_m} f(\cos pp')d\omega_{p'} \right) d\omega_p = \omega_m \int_{\mathcal{S}_m} f(\cos pp')d\omega_{p'}.$$

since the last integral is independent of p. Since $I(1) \geq 0$, the result follows. □

Lemma 4.2. *The ultraspherical polynomials*

$$P_n^{(\lambda)}(\cos t), \quad (n = 0, 1, 2, ...; \quad \lambda = \frac{1}{2}(m-1)) \tag{4.3}$$

are all positive definite on $\mathcal{S}^m$.

Proof. We proceed by induction on m. For $m = 1$, $\lambda = 0$, the statement follows from (3.9). Let $m \geq 2$, and assume that $P_n^{(\lambda-\frac{1}{2})}(\cos t)$ is positive definite on $\mathcal{S}^{m-1}$ for all $n = 0, 1, 2....$ Suppose $p_i \in \mathcal{S}^m (i = 1, ..., N)$ and associate with the point p_i a point p_i', on the "equator" of $\mathcal{S}^{m-1}$ of equation $\theta = \frac{1}{2}\pi$, such that the last $m-1$ polar coordinates $\theta_1, ..., \phi$ of both points p_i and p_i' agree. By (3.36), we have

$$\cos p_i p_k = \cos \theta^i \cos \theta^k + \sin \theta^i \sin \theta^k \cos p_i' p_k'.$$

An equivalent representation for addition formula for ultraspherical polynomials given in Theorem 3.2 may be obtained by letting

$$\nu + \frac{1}{2} = \lambda, \quad \omega = \cos p_i p_k, \quad \alpha = \cos \theta^i, \quad \beta = \cos \theta^k,$$

$$x = -\cos p_i' p_k', \quad (\alpha^2 - 1)^{s/2} P_{n-s}^{(\lambda+s)}(t) =: P_n^{\lambda,s}(t).$$

Then, we have

$$P_n^{(\lambda)}(\cos p_i p_k) = \sum_{s=0}^{n} c_{s,n}^{\lambda} P_n^{\lambda,s}(\cos\theta^i) P_n^{\lambda,s}(\cos\theta^k) P_s^{(\lambda-\frac{1}{2})}(\cos p_i' p_k') \tag{4.4}$$

where $c_{s,n}^{\lambda}$'s are positive coefficients given in (3.2). But then,

$$\sum_{i,k=1}^{N} P_n^{(\lambda)}(\cos p_i p_k)\alpha_i\alpha_k = \sum_{s=0}^{n} c_{n,\lambda,s} \sum_{i,k=1}^{N} P_s^{(\lambda-\frac{1}{2})}(\cos p_i' p_k')\eta_i\eta_k \geq 0 \tag{4.5}$$

where $\eta_i = P_n^{\lambda,s}(\cos\theta^i)\alpha_i$. The expression is indeed non-negative since $P_n^{(\lambda-\frac{1}{2})}(\cos t)$ was assumed to be positive definite on $\mathcal{S}^{m-1}$ for all $n = 0, 1, 2, \ldots$. This completes our proof. □

Theorem 4.3. *A necessary and sufficient condition in order that a continuous function $f(\cos\theta)$ be positive definite on $\mathcal{S}^m$ is that the ultraspherical expansion (3.31) have nonnegative coefficients in which case the series (3.31) converges throughout $0 \leq \theta \leq \pi$ absolutely and uniformly to the sum $f(\cos\theta)$. The most general $f(\cos\theta)$ which is positive definite on $\mathcal{S}^m$ is therefore given by the expansion*

$$f(\cos\theta) = \sum_{n=0}^{\infty} a_n P_n^{(\lambda)}(\cos\theta), \quad (a_n \geq 0, \lambda = \frac{1}{2}(m-1)) \tag{4.6}$$

provided the series converges for $\theta = 0$.

Proof. First, let the function $f(\cos\theta)$ be continuous for $\theta \in [0, \pi]$ and positive definite on $\mathcal{S}^m$. The coefficient of $P_n^{(\lambda)}(\cos\theta)$ in (3.31) may be written as

$$\int_0^{\pi} P_n^{(\lambda)}(\cos\theta') f(\cos\theta') \sin^{m-1}\theta' d\theta' = \frac{1}{\omega_{m-1}} \int_{\mathcal{S}_m} P_n^{(\lambda)}(\cos ap') f(\cos ap') d\omega_{p'} \tag{4.7}$$

where a is the point of $\mathcal{S}^m$ of coordinates $x = 1$, $x_1 = \ldots = x_m = 0$, (See Appendix B.10). But then the integral on the right hand side of (4.7) is positive for the following reason. Since $P_n^{(\lambda)}(\cos t)$ and $f(\cos t)$ are both positive definite on $\mathcal{S}^m$, their product also enjoys this property. Now Lemma 4.1 shows that the integral on the right hand side of (3.31) is non-negative. Also, by the deginiton of the Gamma function, we get that all coefficients of $P_n^{(\lambda)}(\cos\theta)$ in (3.31) are non-negative. We may therefore rewrite (3.31) in the form

$$f(\cos\theta) \sim \sum_{n=0}^{\infty} a_n P_n^{(\lambda)}(\cos\theta), \qquad (a_n \geq 0). \tag{4.8}$$

On the other hand, we know from Theorem 3.8 that this series is Abel-summable for all θ, hence, in particular, for $\theta = 0$. Thus, by (C.9), we have

$$\sum_{n=0}^{k} a_n |P_n^{(\lambda)}(\cos\theta)| \leq \sum_{n=0}^{k} a_n P_n^{(\lambda)}(1) \leq \lim_{r\to 1^-} \sum_{n=0}^{\infty} a_n r^n P_n^{(\lambda)}(1) = f(1).$$

This shows that the series (4.6) is absolutely and uniformly convergent for all θ $(0 \leq \theta \leq \pi)$, hence convergent to its Abel-sum which is $f(\cos\theta)$.
Conversely, Theorem (2.2) implies that the convergent series (4.6) defines a positive definite $f(\cos\theta)$. Indeed, $f(\cos\theta)$ is continuous because the series (4.6) must converge uniformly. Now, $f(\cos\theta)$ being the continuous limit of positive definite functions, it is positive definite itself. $\square$

4.2. Positive Definite Functions in $\mathcal{S}^\infty$

In Chapter 1, we have denoted the class of positive definite functions as $\mathfrak{B}(M)$. An obvious property of $\mathfrak{B}(M)$ is as follows. If $M \subset N$, then $\mathfrak{B}(M) \supset \mathfrak{B}(N)$. By denoting the infinite dimensional sphere as $\mathcal{S}^\infty$, as we may assume

$$\mathcal{S}^1 \subset \mathcal{S}^2 \subset \cdots \subset \mathcal{S}^\infty,$$

it follows that

$$\mathfrak{B}(\mathcal{S}^1) \supset \mathfrak{B}(\mathcal{S}^2) \supset \cdots \supset \mathfrak{B}(\mathcal{S}^\infty).$$

In fact, $\mathfrak{B}(\mathcal{S}^\infty)$ is identical with the intersection of all classes $\mathfrak{B}(\mathcal{S}^m)$ $(m = 1, 2, ...)$. In Example 2.7, we have shown that $\cos t \in \mathfrak{B}(\mathcal{S}^m)$ for all m. Hence, we have $\cos t \in \mathfrak{B}(\mathcal{S}^\infty)$. By Theorem 2.2(i) we also have that $(\cos t)^n \in \mathfrak{B}(\mathcal{S}^\infty)$. Theorem 2.2(i) and (iii) show that

$$f(t) = \sum_{n=0}^{\infty} a_n (\cos t)^n \in \mathfrak{B}(\mathcal{S}^\infty), \tag{4.9}$$

provided $a_n \geq 0$ and $\sum_{n=0}^{\infty} a_n$ converges. In this section, we will show that the functions $f(t)$ of the form (4.9) exhaust the class $\mathfrak{B}(\mathcal{S}^\infty)$.

Lemma 4.4. *For $x \in [-1, 1]$ and $n = 0, 1, 2, ...,$ we have*

$$\lim_{\lambda \to \infty} \frac{P_n^{(\lambda)}(x)}{P_n^{(\lambda)}(1)} = x^n.$$

Proof. Let $p_n^\lambda(x) := \frac{P_n^{(\lambda)}(x)}{P_n^{(\lambda)}(1)}$. Utilizing (C.1) and (C.4), we have

$$p_n^\lambda(x) = \frac{n!\Gamma(2\lambda)}{\Gamma(n + 2\lambda)\Gamma(\lambda)} \sum_{s=0}^{[\frac{n}{2}]} \frac{(-1)^s \Gamma(n + \lambda - s)}{s!(n - 2s)!} (2x)^{n-2s}.$$

Then,

$$\lim_{\lambda \to \infty} p_n^\lambda(x) = \sum_{s=0}^{[\frac{n}{2}]} \frac{(-1)^s (2x)^{n-2s} n!}{s!(n - 2s)!} \lim_{\lambda \to \infty} \frac{\Gamma(n + \lambda - s)\Gamma(2\lambda)}{\Gamma(\lambda)\Gamma(n + 2\lambda)}. \tag{4.10}$$

Now, using (B.8), we have

$$\begin{aligned}\frac{\Gamma(n + \lambda - s)\Gamma(2\lambda)}{\Gamma(\lambda)\Gamma(n + 2\lambda)} &= \frac{(n + \lambda - s)^{n+\lambda-s-\frac{1}{2}} e^{-(n+\lambda-s)} (2\pi)^{1/2} e^{\theta_1/12(n+\lambda-s)}}{\lambda^{\lambda-\frac{1}{2}} e^{-\lambda} (2\pi)^{1/2} e^{\theta_3/12\lambda}} \\ &\times \frac{(2\lambda)^{2\lambda-\frac{1}{2}} e^{-(2\lambda)} (2\pi)^{1/2} e^{\theta_2/12(2\lambda)}}{(n + 2\lambda)^{n+2\lambda-\frac{1}{2}} e^{-(n+2\lambda)} (2\pi)^{1/2} e^{\theta_4/12(n+2\lambda)}}\end{aligned}$$

where $0 < \theta_j < 1$ for $j = 1, ..., 4$. Therefore,

$$\begin{aligned}
\lim_{\lambda\to\infty} \frac{\Gamma(n+\lambda-s)\Gamma(2\lambda)}{\Gamma(\lambda)\Gamma(n+2\lambda)} &= \lim_{\lambda\to\infty} \frac{[(n+\lambda-s)^{n+\lambda-s}e^{-(n+\lambda-s)}][(2\lambda)^{2\lambda}e^{-(2\lambda)}]}{[\lambda^{\lambda}e^{-\lambda}][(n+2\lambda)^{n+2\lambda}e^{-(n+2\lambda)}]} \\
&= \lim_{\lambda\to\infty} e^{s}\Big(\frac{n+\lambda-s}{\lambda}\Big)^{\lambda}\frac{(n+\lambda-s)^{n-s}}{(n+2\lambda)^{n}}\Big(\frac{2\lambda}{n+2\lambda}\Big)^{2\lambda} \\
&= \lim_{\lambda\to\infty} \Big(1+\frac{n-s}{\lambda}\Big)^{\lambda}\Big(\frac{n+\lambda-s}{n+2\lambda}\Big)^{n}\frac{e^{s}}{(n+\lambda-s)^{s}}\Big(\frac{2\lambda}{n+2\lambda}\Big)^{2\lambda} \\
&= e^{s}e^{n-s}2^{-n}e^{-n}\lim_{\lambda\to\infty}\frac{1}{(n+\lambda-s)^{s}} = 2^{-n}
\end{aligned}$$

if $s = 0$, and zero otherwise. Now, the lemma follows by inserting this result into (4.10). □

Theorem 4.5. *A continuous function $f(\cos\theta)$ which is positive definite on $\mathcal{S}^{\infty}$ is necessarily of the form*

$$f(\cos\theta) = \sum_{n=0}^{\infty} a_n \cos^n\theta, \quad (a_n \geq 0) \tag{4.11}$$

where $\sum_{n=0}^{\infty} a_n < \infty$.

For the proof we need the following lemma:

Lemma 4.6. *For all $\theta \in (0,\pi)$, and $\epsilon > 0$, there exists $L(\theta,\epsilon)$ such that for all $\lambda > L(\theta,\epsilon)$*

$$|p_n^{(\lambda)}(\cos\theta) - \cos^n\theta| < \epsilon \tag{4.12}$$

holds for all $n = 0, 1, 2, ...,$ where $p_n^{(\lambda)}(\cos\theta) := \frac{P_n^{(\lambda)}(\cos\theta)}{P_n^{(\lambda)}(1)}$.

Proof. From (C.8), we have

$$\begin{aligned}
\Delta_n^{\lambda} := p_n^{\lambda}(\cos\theta) - \cos^n\theta &= \frac{P_n^{(\lambda)}(\cos\theta) - \cos^n\theta P_n^{(\lambda)}(1)}{P_n^{(\lambda)}(1)} \\
&= \int_0^{\pi} F_n(\theta,\phi)\sin^{(2\lambda-1)}\phi d\phi \Big/ \int_0^{\pi} sin^{(2\lambda-1)}\phi d\phi,
\end{aligned} \tag{4.13}$$

where

$$F_n(\theta,\phi) = (\cos\theta + i\sin\theta\cos\phi)^n - \cos^n\theta. \tag{4.14}$$

We observe that

$$\begin{aligned} |F_n(\theta,\phi)| &\leq |(\cos\theta + i\sin\theta\cos\phi)^n| + |\cos^n\theta| \\ &= (\cos^2\theta + \sin^2\theta\cos^2\phi)^{n/2} + |\cos\theta|^n \\ &\leq (\cos^2\theta + \sin^2\theta)^{n/2} + 1 = 2. \end{aligned} \tag{4.15}$$

Now we choose a $\delta = \delta(\theta)$ such that

$$0 < \delta < \frac{1}{2}\pi, \qquad \cos^2\theta + \sin^2\theta\sin^2\delta < 1. \tag{4.16}$$

Evidently, for $\frac{\pi}{2} - \delta \leq \phi \leq \frac{\pi}{2} + \delta, \quad \cos\phi \leq \cos(\frac{\pi}{2} - \delta) = \sin\delta$. Therefore,

$$|F_n(\theta,\phi)| \leq |(\cos^2\theta + \sin^2\theta\sin^2\delta)^{n/2}| + |\cos\theta|^n.$$

Thus, using (4.13), (4.14), (4.15) and (4.16) we get

$$\begin{aligned}
|\Delta_n^\lambda| &\leq \int_0^\pi |F_n(\theta,\phi)||\sin^{2\lambda-1}\phi|d\phi \Big/ \int_0^\pi sin^{(2\lambda-1)}\phi d\phi \\
&\leq 2\int_0^{\frac{1}{2}\pi-\delta} sin^{(2\lambda-1)}\phi d\phi \Big/ \int_0^\pi sin^{(2\lambda-1)}\phi d\phi \\
&+ \int_{\frac{1}{2}\pi-\delta}^{\frac{1}{2}\pi+\delta} |F_n(\theta,\phi)|\sin^{(2\lambda-1)}\phi d\phi \Big/ \int_0^\pi sin^{(2\lambda-1)}\phi d\phi, \\
&+ 2\int_{\frac{1}{2}\pi+\delta}^{\pi} sin^{(2\lambda-1)}\phi d\phi \Big/ \int_0^\pi sin^{(2\lambda-1)}\phi d\phi \\
&= 4\int_0^{\frac{1}{2}\pi-\delta} sin^{(2\lambda-1)}\phi d\phi \Big/ \int_0^\pi sin^{(2\lambda-1)}\phi d\phi \\
&+ (\cos^2\theta+\sin^2\theta\sin^2\delta)^{\frac{1}{2}n} + |\cos\theta|^n.
\end{aligned}$$

Now let $\epsilon > 0$ be given. Let $n_0 = n_0(\theta,\epsilon)$ be such that $n > n_0$ implies

$$(\cos^2\theta+\sin^2\theta\sin^2\delta)^{\frac{1}{2}n} + |\cos\theta|^n < \frac{1}{2}\epsilon.$$

The existence of such an n_0 is assured by (4.16). Furthermore, suppose $\lambda_0 = \lambda_0(\theta,\epsilon)$ is such that $\lambda > \lambda_0$ implies

$$4\int_0^{\frac{1}{2}\pi-\delta} sin^{(2\lambda-1)}\phi d\phi \Big/ \int_0^\pi sin^{(2\lambda-1)}\phi d\phi < \frac{1}{2}\epsilon.$$

Hence, $|\Delta_n^\lambda| < \epsilon$, provided $\lambda > \lambda_0(\theta,\epsilon)$ and $n > n_0(\theta,\epsilon)$. On the other hand, by Lemma 4.4 we have that $|\Delta_n^\lambda| < \epsilon \quad for \quad n = 0,1,...,n_0(\theta,\epsilon)$, provided $\lambda > \lambda_1(\theta,\epsilon)$. Now taking $L(\theta,\epsilon)$ as the larger of the two numbers λ_0 and λ_1, we get the result. □

Proof of Theorem 4.5. If $f(\cos\theta)$ is positive definite on $\mathcal{S}^\infty$ it is also positive definite on $\mathcal{S}^m$. By Theorem 4.3 we are therefore assured to have an expansion with nonnegative coefficients

$$f(\cos\theta) = \sum_{n=0}^{\infty} a_n(\lambda) P_n^{(\lambda)}(\cos\theta), \quad (a_n(\lambda) \geq 0, \quad 0 \leq \theta \leq \pi) \tag{4.17}$$

which is valid for all values of λ of the form $\lambda = \frac{1}{2}(m-1)$, $(m = 1, 2, 3, ...)$. Setting

$$p_n^{\lambda}(\cos\theta) = \frac{P_n^{(\lambda)}(\cos\theta)}{P_n^{(\lambda)}(1)}, \tag{4.18}$$

we have a similar expansion

$$f(\cos\theta) = \sum_{n=0}^{\infty} b_n(\lambda) p_n^{\lambda}(\cos\theta) \quad (b_n(\lambda) \geq 0, 0 \leq \theta \leq \pi). \tag{4.19}$$

Since $p_n^{\lambda}(1) = 1$, and hence

$$f(1) = \sum_{n=0}^{\infty} b_n(\lambda),$$

we see that the coefficients $b_n(\lambda)$ are uniformly bounded by $f(1)$ for all n and the range of values of λ. By Cantor's diagonalization argument, we may find a subsequence $\lambda_\nu \to \infty$ such that

$$\lim_{\nu\to\infty} b_n(\lambda_\nu) = a_n \geq 0, \quad (n = 0, 1, 2...). \tag{4.20}$$

Now, let θ have a fixed value between 0 and π, and write the relation (4.19) in the form

$$f(\cos\theta) = \sum_{0}^{\infty} b_n(\lambda_\nu) cos^n\theta + \sum_{0}^{\infty} b_n(\lambda_\nu)[p_n^{\lambda_\nu}(\cos\theta) - cos^n\theta].$$

By Lemma 4.6, we have

$$\big|\sum_{0}^{\infty} b_n(\lambda_\nu)[p_n^{\lambda_\nu}(\cos\theta) - cos^n\theta]\big| < \epsilon \sum_{0}^{\infty} b_n(\lambda_\nu) = \epsilon f(1),$$

provided λ_ν is sufficiently large, thus we may write

$$f(\cos\theta) = \sum_0^\infty b_n(\lambda_\nu) cos^n\theta + \sigma, \tag{4.21}$$

where $|\sigma| < \epsilon f(1)$ for sufficiently large λ_ν. However, the series

$$\sum_0^\infty b_n(\lambda_\nu) cos^n\theta \tag{4.22}$$

converges uniformly with respect to the variable λ_ν because it is majorized by the convergent series with constant terms

$$\sum_0^\infty f(1)|cos\theta|^n.$$

But now the limiting relations (4.20) imply that the series (4.22) will tend to $\sum a_n \cos^n\theta$ as $\lambda_\nu \to \infty$. Thus, (4.21) may be written as

$$f(\cos\theta) = \sum_0^\infty a_n cos^n\theta + \sigma'$$

where $|\sigma'| < \epsilon f(1)$. Now letting $\epsilon \to 0$, we conclude that $\sigma' = 0$ and hence that

$$f(\cos\theta) = \sum_0^\infty a_n cos^n\theta. \tag{4.23}$$

It remains to show that (4.23) is also valid for $\theta = 0$ and $\theta = \pi$. This last point, however, is readily settled. Indeed, (4.23) implies the convergence of the series $\sum a_n$ by letting $\theta \to 0$. Now the continuity of both sides of the relation (4.23) at both ends of the interval $0 \le \theta \le \pi$ implies its validity throughout this closed interval. $\square$

In the last part of his article, Schoenberg uses positive definite functions for embedding problems which is a quite different concept from our motivation of studying positive definite functions. Hereafter, we mention his results very briefly. We begin with a definition.

Definition 4.7. *Let $F(t)$,$(0 \leq t \leq \pi)$, is a continuous function with $F(0) = 0$, $F(t) \geq 0$ if $0 < t \leq \pi$. We can remetrize the metric space $\mathcal{S}^\infty$ from the original distance function pq to the new distance function $F(pq)$. The new semi-metric space thus obtained is called the metric transform of $\mathcal{S}^\infty$ by the function $F(t)$ and denoted by the symbol $F(\mathcal{S}^\infty)$.*

Theorem 4.8. *The metric transform $F(\mathcal{S}^\infty)$ is isometrically imbeddable in the real Hilbert space iff*

$$F^2(t) = \sum_{n=1}^{\infty} a_n(1 - \cos^n t), \qquad (a_n \geq, \quad 0 \leq t \leq \pi).$$

In order to prove the above theorem, Schoenberg uses a lemma stating that the metric transform $F(\mathcal{S}^\infty)$ is isometrically imbeddable in real Hilbert space iff the function $\exp[-\lambda F^2(t)]$ is positive definite on $\mathcal{S}^\infty$ for all $\lambda > 0$. The proof of this lemma is given in ([35], p. 527).

Moreover, as Askey [36] observed, the fact that the isometric imbedding of a metric space in another one gives rise to a reverse inclusion in their positive definite functions can be used to obtain "a couple of interesting results" when combined with the work of Bochner [37]. For example, $P^d(R)$ can be isometrically imbedded in $P^{2d}(C)$, but $P^{d+1}(R)$ cannot be isometrically imbedded in $P^{2d}(C)$. Here, $P^d(R)$ and $P^d(R)$ are real and complex projective spaces with dimension d, respectively. For further examples and details, we refer the reader to [31].

5. EXTENSIONS OF POSITIVE DEFINITE FUNCTIONS

In this chapter, our main concern will be strictly positive definite functions and conditionally (strictly) positive definite functions. The results presented in this chapter are mostly based on the results which we have obtained for positive definite functions in previous chapters.

5.1. Strictly Positive Definite Functions

In Chapter 2, we have shown that the interpolation matrix $A_{\Phi,X} = (\phi(\| x_j - x_k \|^2))$ (1.5) is positive semi-definite if ϕ is completely monotone on $[0,\infty)$ (see Theorem 2.16). Recall that a strictly positive definite function leads to a positive definite interpolation matrix. Hence, for the interpolation theorem to be uniquely solvable we should work with strictly positive definite functions rather than positive definite ones. In this section, we will present several characterizations of strictly positive definite functions. The first characterization is due to Micchelli [38].

Theorem 5.1 ([1]). *For a function $\phi : [0,\infty) \to \mathbb{R}$, the corresponding multivariate function $\Phi(\cdot) := \phi(\| \cdot \|)$ is strictly positive definite on every $\mathbb{R}^n$ iff $\phi(\sqrt{\cdot})$ is completely monotone on $[0,\infty)$ but not constant.*

Based on Schoenberg's result Theorem 4.3, Cheney and Xu [39] gave a sufficient condition for strictly positive definite functions on spheres.

Theorem 5.2 ([39]). *Let m be a positive integer. Set $\lambda = (m-1)/2$. Let*

$$g(t) = \sum_{n=0}^{\infty} a_n P_n^{(\lambda)}(\cos\theta), \quad a_n > 0, \quad \sum_{n=0}^{\infty} a_n P_n^{(\lambda)}(1) < \infty. \tag{5.1}$$

Let $x_1, x_2, ..., x_N$ be N distinct points on $\mathcal{S}^m$. In order that the $N \times N$ matrix $A_{jk} = g(x_j x_k)$ be positive definite it is sufficient that the coefficients a_n be positive *for* $0 \leq n < N$.

This means that in order that the function g be strictly positive definite on $\mathcal{S}^m$, it is sufficient that all coefficients a_n be positive.

Later on, Screiner [40] improved Cheney and Xu's result by showing that if only finitely many of the coefficients in the expansion (5.1) are zero, then f is still strictly positive definite.

It is readily seen that the strict-positive-definiteness of a function f given in (5.1) depends only on the set

$$K_{m,f} = \{n \in \mathbb{Z}_+ : a_n > 0\},$$

but not on the actual values of the coefficients a_n [41]. Motivated with this fact we have the following definition.

Definition 5.3 ([41]). *A subset K of $\mathbb{Z}_+$ is said to induce strict positive definiteness on $\mathcal{S}^m$ if the function*

$$\theta \longmapsto \sum_{k \in K} P_k^{(\lambda)}(\cos\theta)$$

is strictly positive definite on $\mathcal{S}^m$. Here we assume that the series converges uniformly.

Example 5.4 ([47]). The following functions are strictly positive definite on $\mathcal{S}^m$.

i. $f(t) = (1 + r^2 - 2rt)^{-1/2}$, where $a_n = h^n$, for $0 < r < 1$,
ii. $f(t) = (1 - r^2)(1 + r^2 - 2rt)^{-3/2}$, where $a_n = (2n + 1)h^n$, for $0 < r < 1$,
iii. $f(t) = 1 - \sqrt{\frac{1-t}{2}}$, where $a_0 = 1/3$ and $a_n = \frac{2}{(2n-1)(2n+3)}$, $n \geq 1$.

Ron and Sun [42] showed that if K contains arbitrarily long sequences of consecutive even integers and of consecutive odd integers, then K induces strict positiveness on $\mathcal{S}^m$.

Menegatto et. al. [41] proved that in order for a subset K of $\mathbb{Z}_+$ to induce strict positive definiteness on $\mathcal{S}^m$, $(m \geq 2)$, it is necessary and sufficient that K contains infinitely many odd integers as well as infinitely many even ones. This characterization, however, fails to stand for strictly positive definite functions on the unit circle, which corresponds to the case $m = 1$. Giving a full characterization for strictly positive definite functions on the unit circle is an open problem [43].

In chapter 4, we have presented Schoenberg's characterization of the family of all positive definite functions on the unit sphere of a real Hilbert space. Menegatto [44] showed necessary and sufficient conditions for the strictly positive definite functions on the unit sphere of a real and complex Hilbert space [45].

5.2. Conditionally Positive Definite Functions

Our investigation of positive definite functions was motivated by the interpolation problem (1.4), using an interpolant of the form (1.3). This particular choice was very convenient for analysis because everything could be restricted to the investigation of a single function Φ. But of course (1.3) is not the only possible approach to the interpolation problem. We will show in this section that we can interpolate the data uniquely without using strictly positive definite radial basis functions. Our motivation is the following theorem:

Theorem 5.5 ([38])**.** *Let $\phi \in C^\infty[0,\infty)$ be such that ϕ' is completely monotone but not constant. Suppose further that $\phi(0) \geq 0$. Then $A_{\Phi,X}$ is non-singular for $\Phi(x) = \phi(\| x \|^2)$.*

This is a weaker requirement because if ϕ is completely monotone, then all of its derivatives are, subject to a suitable sign change in the original function. Thus, we may consider ϕ such that ϕ' is completely monotone, or so as to weaken the requirement further, by demanding that, for some k, $(-1)^k \phi^{(k)}$ is completely monotone. In this case, positive definiteness of the interpolation matrix can no longer be expected. As an example, consider the thin plate spline

$$\Phi(x) = \parallel x \parallel^2 \log \parallel x \parallel, \quad x \in \mathbb{R}^n.$$

Obviously, it is a radial function so let $r = \parallel x \parallel^2$. Then,

$$\phi(r) = \frac{1}{2} r \log r.$$

Then, $\phi'(r) = \frac{1}{2}\log r + \frac{1}{2}$ *and* $\phi''(r) = \frac{1}{2r}$ which is completely monotone on $(0, \infty)$. Now, let the centers be the vertices of a regular simplex whose edges are all of unit length, and $N = n + 1$. Then, all entries $\Phi(x_j - x_k)$ of the interpolation matrix are zero meaning that the thin-plate spline is not strictly positive definite. We will later see that a slight change in the definition of interpolant (1.3) ensures solvability of the interpolation problem with the thin plate spline.

Hence, we should generalize the notion of positive definite functions in a way that covers all the relevant possibilities for basis functions. The notion of *conditional positive definiteness* serves for this purpose.

Definition 5.6. *A continuous function* $\Phi : \mathbb{R}^n \to \mathbb{C}$ *is said to be* conditionally positive definite of order m *if, for all* $N \in \mathbb{N}$*, all pairwise distinct centers* $x_1, ..., x_N \in \mathbb{R}^n$*, and all* $\alpha \in \mathbb{C}^N$ *satisfying*

$$\sum_{j=1}^{N} \alpha_j p(x_j) = 0 \tag{5.2}$$

for all complex-valued polynomials of degree less than m*, the quadratic form*

$$\sum_{j,k=1}^{N} \alpha_j \overline{\alpha_k} \Phi(x_j - x_k) \tag{5.3}$$

is nonnegative, Φ *is said to be* strictly conditionally positive definite of order m *if* *the quadratic form is positive unless α is zero.*

The conditional positive definiteness of order m can be interpreted as the positive definiteness of the matrix $A_{\Phi,X} = (\Phi(x_j - x_k))$ on the space of vectors α such that

$$\sum_{j=1}^{N} \alpha_j p_l(x_j) = 0, \quad 1 \leq l \leq Q = dim(\pi_{m-1}(\mathbb{R}^n)),$$

where $\pi_{m-1}(\mathbb{R}^n)$ denotes the space of polynomials of total degree at most $(m-1)$ in n unknowns. Thus, in this sense, $A_{\Phi,X}$ is positive definite on the space of vectors α "perpendicular" to polynomials. We should reflect on this subject for a moment. For each pair consisting of a vector $\alpha \in \mathbb{C}^N$ and a set of distinct points $X = \{x_1, ..., x_N\}$ that together satisfy (5.2) for all polynomials in $\pi_{m-1}(\mathbb{R}^n)$, we define a linear functional

$$\lambda_{\alpha,X} := \sum_{j=1}^{N} \alpha_j \delta_{x_j}$$

where δ_x denotes the point evaluation functional at x, and we denote the space of all such functionals by $\pi_{m-1}(\mathbb{R}^n)^{\perp}$. Then α is admissible in the definition of a conditionally positive definite function iff $\lambda_{\alpha,X} \in \pi_{m-1}(\mathbb{R}^n)^{\perp}$.

Since the matrix $A_{\Phi,X}$ is conditionally positive definite of order m, it is positive definite on a subspace of dimension $N-Q$, $Q = dim(\pi_{m-1}(\mathbb{R}^n))$. Courant-Fischer theorem, ([46],p.136), implies that at least $N-Q$ of its eigenvalues are positive. In the case $m=1$, we have a stronger statement:

Theorem 5.7. *Suppose that Φ is conditionally positive definite of order 1 and that $\Phi(0) \leq 0$. Then the matrix $A_{\Phi,X} \in \mathbb{R}^{N \times N}$ has one negative and $N-1$ positive eigenvalues. In particular, it is invertible.*

Proof. From the Courant-Fischer theorem we conclude that $A_{\Phi,X}$ has at least $N-1$

positive eigenvalues. But since $0 \geq N\Phi(0) = tr(A_{\Phi,X}) = \sum_{i=1}^{n} \lambda_i$ where the λ_i denote the eigenvalues of $A_{\Phi,X}$ and $tr(A_{\Phi,X})$ its trace, $A_{\Phi,X}$ must also have at least one negative eigenvalue. □

Theorem 5.8. *Let a function $\phi \in C[0,\infty) \cap C^{\infty}(0,\infty)$ be given. Then the function $\Phi(x) = \phi(\| x \|^2)$ is conditionally positive definite of order m on every $\mathbb{R}^n$ iff $(-1)^m \phi^{(m)}$ is completely monotone on $(0,\infty)$.*

It follows immediately from the preceding theorem that the thin plate spline is conditionally positive definite of order 2.

The most important observation we make now is that we can always interpolate uniquely with strictly conditionally positive definite functions if we define the interpolant to a function f at the centers $X = \{x_1, ..., x_N\}$ as

$$s_{f,X}(x) = \sum_{j=1}^{N} \alpha_j \Phi(x - x_j) + \sum_{k=1}^{Q} \beta_k p_k(x).$$

Here, Q is again the dimension of the polynomial space $\pi_{m-1}(\mathbb{R}^n)$ and $\{p_1, ..., p_Q\}$ is a basis of $\pi_{m-1}(\mathbb{R}^n)$. To cope with the additional degrees of freedom, the interpolation conditions

$$s_{f,X}(x_j) = f(x_j), \quad 1 \leq j \leq N,$$

are completed by the additional conditions

$$\sum_{j=1}^{N} \alpha_j p_k(x_j) = 0, \quad 1 \leq k \leq Q.$$

Solvabilitiy of this system is therefore equivalent to solvability of the system

$$\begin{pmatrix} A_{\Phi,X} & P \\ P^{\mathrm{T}} & 0 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} f|_X \\ 0 \end{pmatrix} \tag{5.4}$$

where $A_{\Phi,X} = (\Phi(x_j - x_k)) \in \mathbb{R}^{N\times N}$ and $P = (p_k(x_j)) \in \mathbb{R}^{N\times Q}$. This last system is obviously solvable if the matrix on the left hand side, which we will denote by $\widetilde{A}_{\Phi,X}$, is invertible.

Theorem 5.9 ([1]). *Suppose that Φ is conditionally positive definite of order m and X is a $\pi_{m-1}(\mathbb{R}^n)-$unisolvent set of centers. Then the system (5.4) is uniquely solvable.*

Note that the points $X = \{x_1, ..., x_N\} \subseteq \mathbb{R}^n$ with $N \geq Q = dim\pi_{m-1}(\mathbb{R}^n)$ are called $\pi_{m-1}(\mathbb{R}^n)$ $-$ *unisolvent* if the zero polynomial is the only polynomial from $\pi_{m-1}(\mathbb{R}^n)$ that vanishes on all of them.

The method described in this section can be generalized by using arbitrary linearly independent functions $p_1, ..., p_Q$ on $\mathbb{R}^n$ instead of polynomials. Moreover, the conditionally positive definite function can be replaced by a conditionally positive definite kernel $\Phi : \Omega \times \Omega \to \mathbb{C}$.

For zonal functions on the unit sphere S^n, we have the following definition.

Definition 5.10 ([47]). *A function $\Phi : [0, \pi] \to \mathbb{R}$ is said to be conditionally strictly positive of order m on S^n whenever its associated interpolation matrix $A_{\Phi,X} = \Phi(\arccos(x_j^{\mathrm{T}} x_k))$ is positive definite on the subspace of $\mathbb{R}^N$ given by*

$$W_{m-1} = \{\alpha \in \mathbb{R}^N : \sum_{j=1}^{N} \alpha_j Y(x_j) = 0 \quad \textit{for all} \quad Y \in \pi_{m-1}(S^n)\}$$

for all distinct set of centers $X = x_1, ..., x_N$ on S^n. Here, $\pi_{m-1}(S^n) = \pi_{m-1}(\mathbb{R}^{n+1})|_{S^n}$ denotes the space of all spherical harmonics on S^n of order at most $m-1$. Detailed information on spherical harmonics can be found in [34].

Based on Schoenberg's result on positive definite functions, for zonal functions on the unit sphere S^n, we have the following characterization of conditional strict positive definiteness:

Theorem 5.11 ([47])**.** *If $f(t)$ is conditionally strictly positive definite of order m on S^n, then it has the following form*

$$f(t) = \sum_{k=0}^{\infty} a_k P_k^{(\lambda)}(t) \tag{5.5}$$

where

$$a_k \geq 0 \quad \textit{for} \quad k \geq m \quad \textit{and} \quad \sum_{k=0}^{\infty} a_k P_k^{(\lambda)}(1) < \infty. \tag{5.6}$$

Here, $P_k^{(\lambda)}$ are the ultraspherical polynomials and $\lambda = (n-1)/2$.

Remark 5.12. *The case that $f(t)$ is strictly positive definite of order on S^n is covered by setting $m = 0$ in the above theorem.*

However, the complete characterization of the class of functions of the form (5.5) satisfying (5.6) that are conditionally positive definite on S^n remains an open problem. See [47] and [17] for further information.

APPENDIX A: The Fourier Transform on $\mathbb{R}^m$

For the proofs of the results given in this section, one may consult on [26], for example.

Definition A.1. *For $f \in L_1(\mathbb{R}^m)$ we define its Fourier transform by*

$$\widehat{f}(x) = (2\pi)^{-m/2} \int_{\mathbb{R}^m} f(\omega) e^{-ix^{\mathrm{T}}\omega} d\omega$$

and its inverse Fourier transform by

$$f^{\curlyvee}(x) = (2\pi)^{-m/2} \int_{\mathbb{R}^m} f(\omega) e^{ix^{\mathrm{T}}\omega} d\omega, \quad x \in \mathbb{R}^m. \tag{A.1}$$

Theorem A.2. *If $f \in L_1(\mathbb{R}^m)$ is continuous and has a Fourier transform $\widehat{f} \in L_1(\mathbb{R}^m)$ then f can be recovered from its Fourier transform:*

$$f(x) = (2\pi)^{-m/2} \int_{\mathbb{R}^m} \widehat{f}(\omega) e^{ix^{\mathrm{T}}\omega} d\omega, x \in \mathbb{R}^m. \tag{A.2}$$

Theorem A.3. *Suppose $f, g \in L_1(\mathbb{R}^m)$. Then,*

i. $\int_{\mathbb{R}^m} \widehat{f}(x) g(x) dx = \int_{\mathbb{R}^m} f(x) \widehat{g}(x) dx$

ii. The Fourier transform of the convolution

$$f * g(x) = \int_{\mathbb{R}^m} f(y) g(x-y) dy$$

is given by

$$f * g = (2\pi)^{m/2} \widehat{f}\widehat{g}.$$

iii. For $\tau_a f(x) = f(x-a), \quad a \in \mathbb{R}^m$, we have $\widehat{\tau_a f}(x) = e^{-ix^{\mathrm{T}}a} \widehat{f}(x)$.

iv. With $\tilde{f}(x) = \overline{f(-x)}$, we find that $f * \tilde{f} = (2\pi)^{m/2} |\widehat{f}|^2$.

Theorem A.4. *Define $g_l(x) = (l/\pi)^{m/2} e^{-l\|x\|^2}, l \in \mathbb{N}, x \in \mathbb{R}^m$. Then the following* *hold true.*

i. $\int_{\mathbb{R}^m} g_l(x)dx = 1$,

ii. $\widehat{g}_l(x) = (2\pi)^{-m/2} e^{-\|x\|_2^2/4l}$,

iii. $\Phi(x) = \lim_{l\to\infty} \int_{\mathbb{R}^m} \Phi(\omega) g_l(\omega - x) d\omega$ *provided that* $\Phi(x) = \mathcal{O}(\| x \|^n)$ *for* $\| x \| \to \infty$, $n \in \mathbb{N}$.

APPENDIX B: The Gamma Function and Integration on Spheres

The material given in this appendix is mostly taken from [48, 32, 34].

The Gamma function can be defined as

$$\Gamma(z) = \int_0^\infty e^{-t} t^{z-1} dt \tag{B.1}$$

for $Re(z) > 0$. It follows from the definition that

$$\Gamma(\frac{1}{2}) = \sqrt{\pi} \quad and \quad \Gamma(1) = 1. \tag{B.2}$$

A useful property of the Gamma function follows from the definition applying integration by parts:

$$\Gamma(z+1) = z\Gamma(z). \tag{B.3}$$

Another important formula is

$$\Gamma(z)\Gamma(z+\frac{1}{n})\cdots\Gamma(z+\frac{n-1}{n}) = n^{\frac{1}{2}-nz}(2\pi)^{(n-1)/2}\Gamma(nz) \tag{B.4}$$

where n is a positive integer.
The special case n=2 of (B.4) is

$$\Gamma(z)\Gamma(z+\frac{1}{2}) = 2^{1-2z}\sqrt{\pi}\Gamma(2z) \tag{B.5}$$

which is known as the *duplication* formula [48].

The Euler integral of the first kind

$$B(p,q) = \int_0^1 x^{p-1}(1-x)^{q-1}dx = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}, \quad p > 0, \quad q > 0 \tag{B.6}$$

can be expressed in terms of the Gamma function as

$$B(p,q) = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}. \tag{B.7}$$

Another essential identity for the Gamma function is

$$\Gamma(x) = x^{x-\frac{1}{2}}e^{-x}(2\pi)^{\frac{1}{2}}e^{\theta/12x} \tag{B.8}$$

where $x > 0$ and $0 < \theta < 1$ (for details see [48, p. 253]).

From now on, we summarize the necessary information about integration on spheres following [34].

Let $\mathbb{R}^{m+1}$ be the usual $(m+1)$-dimensional Euclidean space, let

$$\xi \cdot \eta = x_1y_1 + ... + x_my_m + x_{m+1}y_{m+1}$$

denotes the Euclidean inner product of two vectors $\xi = (x_1, ..., x_{m+1})^{\mathrm{T}}$ and $\eta = (y_1, ..., y_{m+1})^{\mathrm{T}}$ in $\mathbb{R}^{m+1}$ and let

$$\parallel \xi \parallel = \sqrt{\xi \cdot \xi} = \sqrt{x_1^2 + ... + x_{m+1}^2}$$

denotes the Euclidean norm of ξ. The m-dimensional unit sphere $\mathcal{S}^m$ is given by

$$\mathcal{S}^m = \{\xi \in \mathbb{R}^{m+1} : \quad \parallel \xi \parallel = 1\}.$$

We denote its surface element by $d\omega_m$ and the total surface by ω_m, where this surface is given by

$$\omega_m = \int_{\mathcal{S}^m} d\omega_m.$$

By definition we set $\omega_0 = 2$. Then we have

$$\omega_1 = 2\pi \quad and \quad \omega_2 = 4\pi.$$

If the vectors $\varepsilon_1, ..., \varepsilon_{m+1}$ are an orthonormal system in $\mathbb{R}^{m+1}$, we may represent the points on $\mathcal{S}^m$ by

$$\xi_{m+1} = t\varepsilon_{m+1} + \sqrt{1-t^2}\xi_m; \quad -1 \leq t \leq 1; \quad t = \varepsilon_{m+1} \cdot \xi_{m+1} \tag{B.9}$$

where ξ_m is a unit vector in the space spanned by $\varepsilon_1, ..., \varepsilon_m$. The surface element of the unit sphere then can be written as

$$d\omega_m = (1-t^2)^{(m-2)/2} dt d\omega_{m-1}, \tag{B.10}$$

and we have from above

$$\omega_m = \int_{\mathcal{S}_{m-1}} \int_{-1}^{1} (1-t^2)^{(m-2)/2} dt d\omega_{m-1}. \tag{B.11}$$

Making the transformation $u = t^2$, we have

$$\int_{-1}^{1} (1-t^2)^{(m-2)/2} dt = \int_{0}^{1} (1-u)^{(m-2)/2} u^{-1/2} du = \frac{\Gamma(\frac{1}{2})\Gamma(\frac{m}{2})}{\Gamma(\frac{m+1}{2})}$$

where the last equality follows from (B.6) and (B.7). Hence, for $m = 2, 3, ...$ we have

$$\omega_m = \frac{\sqrt{\pi}\Gamma(\frac{m}{2})}{\Gamma(\frac{m+1}{2})}\omega_{m-1} = \frac{(\sqrt{\pi})^{m-1}}{\Gamma(\frac{m+1}{2})}\omega_1 = \frac{2\pi^{\frac{m+1}{2}}}{\Gamma(\frac{m+1}{2})}. \tag{B.12}$$

Definition B.1. *A point* $p = (x, x_1, ..., x_m) \in \mathcal{S}^m$ *is expressed in terms of hyper-spherical coordinates as*

$$\begin{aligned}
x &= \cos\theta, \\
x_1 &= \sin\theta\sin\theta_1\cos\theta_2, \\
&\cdot \\
&\cdot \\
&\cdot \\
x_{m-2} &= \sin\theta\sin\theta_1...\cos\theta_{m-2}, \\
x_{m-1} &= \sin\theta\sin\theta_1...\sin\theta_{m-2}\cos\phi, \qquad (0 \leq \theta_j \leq \pi, 0 \leq \phi \leq 2\pi) \\
x_m &= \sin\theta\sin\theta_1...\sin\theta_{m-2}\sin\phi, \qquad (j = 0, 1, ..., m-2).
\end{aligned} \tag{B.13}$$

APPENDIX C: Further Properties of Ultraspherical Polynomials

In addition to the definitions in Chapter 3, we add here more information about the ultraspherical polynomials following [32], [22].

Ultraspherical polynomials have the following explicit expression

$$P_n^{(\lambda)}(x) = \sum_{s=0}^{[n/2]} \frac{(-1)^s \Gamma(\lambda + n - s)}{\Gamma(\lambda) s!(n-2s)!} (2x)^{n-2s} \tag{C.1}$$

for $n \geq 0$; and

$$P_{-n}^{(\lambda)}(x) = 0 \tag{C.2}$$

for all $n = 1, 2, ...$

From (C.1), we can easily read off the first few ultraspherical polynomials:

$$\begin{aligned}
P_0^{(\lambda)}(x) &= 1, \\
P_1^{(\lambda)}(x) &= 2\lambda x, \\
P_2^{(\lambda)}(x) &= -\lambda + 2\lambda(1+\lambda)x^2, \\
P_3^{(\lambda)}(x) &= -2\lambda(1+\lambda)x + \frac{4}{3}\lambda(1+\lambda)(2+\lambda)x^3.
\end{aligned}$$

Moreover, writing $x = 0$ in (C.1), we obtain

$$P_{2n}^{(\lambda)}(0) = \frac{(-1)^n \Gamma(n+\lambda)}{n!\Gamma(\lambda)}, \quad P_{2n+1}^{(\lambda)}(0) = 0 \tag{C.3}$$

and using (3.7) in $x = 1$, we have

$$P_n^{(\lambda)}(1) = \frac{\Gamma(n+2\lambda)}{n!\Gamma(2\lambda)}. \tag{C.4}$$

The following property of ultraspherical polynomials follows easily from (C.1)

$$D_x P_n^{(\lambda)}(x) = 2\lambda P_{n-1}^{(\lambda+1)}(x) \tag{C.5}$$

which implies that

$$D_x^n P_n^{(\lambda)}(x) = \frac{2^n \Gamma(\lambda+n)}{\Gamma(\lambda)}. \tag{C.6}$$

Other well known properties of ultraspherical polynomials $P_n^{(\lambda)}(x)$ is that they satisfy the following recurrence relation

$$nP_n^{(\lambda)}(x) = 2(n+\lambda-1)xP_{n-1}^{(\lambda)}(x) - (n+2\lambda-2)P_{n-2}^{(\lambda)}(x), \quad n = 2, 3, ...$$

and the following equation named as Gegenbauer differential equation

$$(1-x^2)y'' - (2\lambda-1)xy' + n(n+2\lambda)y = 0, \quad y = P_n^{(\lambda)}(x). \tag{C.7}$$

Laplaces's integral representation for ultraspherical polynomials is

$$P_n^{(\lambda)}(\cos\theta) = \frac{2^{1-2\lambda}\Gamma(n+2\lambda)}{\Gamma(\lambda)n!} \int_0^\pi (\cos\theta + \sin\theta\cos\phi)^n \sin^{(2\lambda-1)}\phi d\phi. \tag{C.8}$$

It immediately follows from the Laplace's integral representation that

$$|P_n^{(\lambda)}(\cos\theta)| \le P_n^{(\lambda)}(1). \tag{C.9}$$

REFERENCES

1. Wendland, H., *Scattered Data Approximation*, Cambridge Monographs on Applied and Computational Mathematics-17, Cambridge University Press, Cambridge, 2005.

2. Buhmann, M. D., *Radial Basis Functions: Theory and Implementations*, Cambridge Monographs on Applied and Computational Mathematics-17, Cambridge University Press, Cambridge, 2003.

3. Kincaid, D. and E. W. Cheney, *Numerical Analysis:Mathematics of scientific computing*, 3rd Edition, The Brooks/Cole, Pacific Grove, California, 2001.

4. Stewart, J., *Positive Definite Functions and Generalizations, A Historical Survey*, Rocky Mountain Journal of Mathematics, 6, pp. 409-434, 1976.

5. Bochner, S., *Harmonic Analysis and the Theory of Probability*, Dover, New York, 2005.

6. Berg, C., J. P. R. Christensen and P. Ressel, *Harmonic Analysis on Semigroups:Theory of Positive Definite and Related Functions*, Springer-Verlag, New York, 1984.

7. Akhiezer, I. N., *The Classical Moment Problem and Some Related Questions in Analysis*, Oliver&Boyd, 1965.

8. Mercer, J., *Functions of positive and negative type and their connection with the theory of integral equations*, Philos. Trans. Roy. Soc. London, 209, pp. 415-446, 1909.

9. Schoenberg, I. J., *Positive Definite Functions on Spheres*, Duke Math. J., 9, pp. 96-108, 1942.

10. Schölkopf, B. and A. J. Smola, *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*, MIT Press Cambridge, MA, 2001.

11. Mathias, M., *Uber positive Fourier-integrale*, Math. Zeit., 16, pp. 103-125, 1923.

12. Aronszajn, N., *Theory of Reproducing Kernels*, Trans. Am. Math. Soc., 68, pp. 337-404, 1950.

13. Schaback, R., *Native Hilbert Spaces for Radial Basis Functions I*,Int.Ser.Numer.Math., Birkhauser Verlag, Basel, 132, pp. 255-282, 1999.

14. Paulsen, V. I., *An Introduction to the Theory of Reproducing Kernel Hilbert Spaces*, Unpublished Lecture Notes, 2009.

15. Narcowich, F., *Generalized Hermite Interpolation by Positive Definite Kernels on a Riemannian Manifold*, J. Math Anal. Appl., 190, pp. 165-193, 1995.

16. Dyn, N., F. Narcowich, and J. Ward, *Variational Principles and Sobolev-type Estimates for Generalized Interpolation on a Riemannian Manifold*, Constr. Approx., 15, pp. 175-208, 1999.

17. Freeden, W., T. Gervens and M. Schreiner, *Constructive Approximation on the Sphere with Applications to Geomathematics*, Oxford University Press, Oxford, 1998.

18. Freeden, W., *On Spherical Spline Interpolation and Approximation*, Math. Methods Appl. Sci., 3, pp. 551–575, 1981.

19. Wahba, G., *Spline Interpolation and Smoothing on the Sphere*, SIAM J. Sci. Stat. Comput., 3, pp. 385-386, 1982.

20. Wahba, G., *Surface Fitting with scattered noisy data on Euclidean d-space and on the sphere*, Rocky Mountain J. Math., 14, pp. 281-299, 1984.

21. Jetter, K., J. Stöckler, and J. Ward, *Error estimates for scattered data interpolation on spheres*, Math. Comput., 68, pp. 733-747, 1999.

22. Nielsen, N., *Les Fonctions Metaspheriques*, Ecole Polytechnique University Press, Paris, 1911.

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

24. Schoenberg, I. J., *Metric Spaces and Completely Monotone Functions*, Ann. Math., 39, pp. 811-841, 1938.

25. Folland, G. B., *Real Analysis:Modern Techniques and Their Applications*, 2^{nd} ed., Wiley, 1999.

26. Stein, E. M. and G. Weiss, *Fourier Analysis in Euclidean Spaces*, Princeton University Press, Princeton, New Jersey, 1971.

27. Bochner, S., *Monotone Funktionen, Stieltjes Integrale und Harmonische Analyse*, Math. Ann., 108, pp. 378–410, 1933.

28. Cuppens, R., *Decomposition of Multivariate Probability*, Academic Press, New York, 1975.

29. Gelfand, I. M. and N. Y. Vilenkin, *Generalized Functions*, Vol.4, Academic Press, New York, 1964.

30. Widder, D. V., *The Laplace Transform*, Princeton University Press, Princeton, New Jersey, 1941.

31. Askey, R., *Orthogonal Polynomials and Related Functions*, Odyssey Press, Dover, New Hampshire, 1994.

32. Szegö, G., *Orthogonal Polynomials*, 4^{th} ed., Amer. Math. Soc.,Providence, 1975.

33. Koornwinder, T. H., *Yet Another Proof of the Addition Formula for Jacobi Polynomials*, J. Math. Anal. Appl., 61, pp. 136-141, 1977.

34. Müller, C., *Spherical Harmonics*, Lecture Notes in Mathematics, 17, Springer, 1966.

35. Schoenberg, I. J., *Metric Spaces and Positive Definite Functions*, Trans. Amer. Math. Soc., 44, pp. 522-536, 1938.

36. Askey, R., *Jacobi Polynomial Expansions with Positive Coefficients and Imbeddings of Projective Spaces*, Bull. Amer. Math. Soc., 74, pp. 301-304, 1968.

37. Bochner, S., *Hilbert Distances and Positive Definite Functions*, Ann. of Math., 42, pp. 647-656, 1941.

38. Micchelli, C. A., *Interpolation of Scattered Data:Distance Matrices and Conditionally Positive Definite Functions*, Const. Approx., 2, pp. 11-22, 1986.

39. Xu Y. and E. W. Cheney, *Strictly Positive Definite Functions on Spheres*, Proc. Amer. Math. Soc., 116, pp. 977-981, 1992.

40. Schreiner, M., *On a new condition for strictly positive definite functions on spheres*, Proc. Amer. Math. Soc., 125, pp. 531-539, 1997.

41. Menegatto, V. A., D. Chen, and X. Sun, *A necessary and Sufficient Condition on Strictly Positive Definte Functions on Spheres*, Proc. Amer. Math. Soc., 131, pp. 2733-2740, 2003.

42. Ron, A. and X. Sun, *Strictly Positive Definite Functions on Spheres in Euclidean Spaces*, Math. Comp., 65, 1513-1530, 1996.

43. Sun, X., *Strictly Positive Definite Functions on the Unit Circle*, Math. Comp., 74, pp. 709-721, 2004.

44. Menegatto, V. A., *Strictly positive definite kernels on the Hilbert sphere*, Appl. Anal., 55, pp. 91-101, 1994.

45. Menegatto, V. A., *Interpolation on the complex Hilbert sphere*, Approx. Theory Appl., 11, pp. 1-9, 1995.

46. Kress, R., *Numerical Analysis*, Springer-Verlag, New York, 1998.

47. Baxter, B. J. C. and S. Hubbert, *Radial Basis Functions for the Sphere*, Recent progress in multivariate approximation, Birkhauser, Berlin, pp. 33-47, 2001.

48. Whittaker, E. T., and G. N. Watson, *A Course of Modern Analysis*, Cambridge University Press, Cambridge, 1965.