Kernel Approximation on Compact Homogeneous Spaces

Carl Richard Odell.

Thesis submitted to The University of Leicester
for the degree of Doctor of Philosophy

January 2011

Dedicated to my Mum and Dad

Carl Odell: Kernel Approximation on Compact Homogeneous Spaces: Abstract

This thesis is concerned with approximation on compact homogeneous spaces. The first part of the research involves a particular kind of compact homogeneous space, the hypersphere, S^{d-1} embedded in $\mathbf{R}^d$. It is a calculation of three integrals associated with approximation using radial basis functions, calculating the Fourier-Gegenbauer coefficients for two such functions. The latter part of the research is a calculation of an error bound for compact homogeneous spaces when interpolating with a G-invariant kernel, a generalisation of a result already known for spheres.

Contents

CHAPTER 1

Introduction

This thesis is primarily concerned with approximation on compact homogeneous spaces. The most well-known example of these spaces is the unit sphere in d dimensions. We will generalise the error estimates for interpolation on spheres which were proved by Jetter et al. [**?**] and Light et al. [**?**], to the case of the compact homogeneous space. We take our inspiration for the development of the material here from the recent developments in the variational approach started by Duchon [**?**] for the approximation of scattered data in high dimensional Euclidean spaces. We will motivate what is to come later by considering approximation on spheres. In the next section we will give an outline of the key ideas that will be needed. We will then discuss more general compact homogeneous spaces, and will look in particular at the two-point spaces as a specific example.

1.1 Approximation on Spheres

Approximating accurately on the sphere is useful for a number of scientific disciplines, among them metereology and geology. We are interested in particular in interpolation on the sphere. Let xy denote the inner product of two points in $\mathbf{R}^d$. The length of the vector $x \in \mathbf{R}^d$ is $|x| = (xx)^{1/2}$. Then the sphere $S^{d-1} = \{x : |x| = 1\}$. Using the fact that $xy = |x||y| \cos\theta$ where $0 \leq \theta \leq \pi$ is the angle between x and y, and that $|x| = |y| = 1$ for $x, y \in S^{d-1}$, we define the geodesic distance $d(x, y)$ between $x, y \in S^{d-1}$ by $d(x, y) = \arccos(xy)$.

The interpolation problem is as follows: Given n distinct points on the sphere, Y (our 'knot set') and n pieces of data, $\{f_y, y \in Y\}$, we need a function s_Y, such that:

$$s_Y(y) = f_y, \quad y \in Y. \tag{1.1.1}$$

The function, s_Y, is then known as an interpolant to the data. Of course, there are many such potential candidates for s_Y. The set from which we will be drawing our interpolant consists of functions formed from translates of a fixed basis function. The form of such a function is as follows:

$$s_Y(x) = \sum_{y \in Y} c_y \phi(d(x, y)), \tag{1.1.2}$$

where ϕ is a univariate function.

Definition 1.1. *A kernel $\kappa : S^{d-1} \times S^{d-1} \to \mathbf{R}$ is called a* zonal kernel *if it has the form:*

$$\kappa(x, y) = \phi(d(x, y)),$$

where ϕ is a univariate function defined on $[-1, 1]$.

Remark 1.1. *Later in this chapter we will introduce the G-invariant kernels, which generalise zonal kernels. These kernels are invariant under the action of the symmetry group of the manifold in question. In this case we have the orthogonal group acting on the sphere, and functions of the geodesic distance are the only G-invariant ones. We will see in Subsection 1.3.1 that such manifolds are termed two-point homogeneous manifolds, or manifolds of rank 1, and that the analysis of approximation on such manifolds is more or less the same as that for the sphere.*

If we substitute an interpolant of the form (1.1.1) into (1.1.2) we end up with a set of linear equations for the unknown vector $\mathbf{c} = \{c_y : y \in Y\}$

$$A\mathbf{c} = \mathbf{f},$$

where

$$A_{yz} = \phi(d(y,z)), \quad y,z \in Y,$$

and $\mathbf{f} = \{f_y : y \in Y\}$ is the vector of data.

The matrix is known as the interpolation matrix and the interpolation problem is solvable if and only if the interpolation matrix is invertible. The ideal situation is that it is invertible for any set of n distinct knot points. One way to ensure invertibility is for the matrix to be positive definite, in other words, for the function to be strictly positive definite.

Definition 1.2. *A function ϕ:[0,2π)$\rightarrow$ $\mathbf{R}$ is strictly positive definite on S^{d-1} if for any finite set of points $Y \subset S^{d-1}$ of cardinality n,*

$$\mathbf{c}^t A\mathbf{c} > 0$$

for all $0 \neq \mathbf{c} \in \mathbf{R}^n$, where A is the associated matrix.

Remark 1.2. *Let Y be a finite point set on the sphere. Then the space $T(Y)$ of translates of a strictly positive definite function ϕ is:*

$$T(Y) = \left\{ \sum_{y \in Y} \lambda_y \phi(d(\cdot, y)),\ \lambda_y \in \mathbf{R} \right\}.$$

Our aim is to demonstrate that there exists a Hilbert space of continuous functions within which the interpolant is in some sense optimal. On $T(Y)$ we define an inner product of two functions

$$f(\cdot) = \sum_{y \in Y} \alpha_y \phi(d(\cdot, y)),$$

and

$$g(\cdot) = \sum_{z \in Y} \beta_z \phi(d(\cdot, z)),$$

in this space by

$$(f,g)_{T(Y)} = \sum_{y,z \in Y} \alpha_y \beta_z \phi(d(y,z)).$$

This in turn defines a norm $\|f\|_{T(Y)} = (f,f)^{1/2}_{T(Y)}$.

Definition 1.3. *(First definition)*
The completion of $T(Y)$ is the **native space** *of ϕ which we call $\mathcal{N}_\phi$, with associated norm $\|\cdot\|_\phi$. More details can be found in Schaback [27]. The native space is significant because it forms the natural space of functions which one can approximate using translates of the basis function ϕ.*

We can also define the native space via Fourier series. The strictly positive definite univariate function ϕ has an expansion in Gegenbauer polynomials

$$\phi(t) = \sum_{k=0}^{\infty} b_k P_k^{((d-3)/2)}(t), t \in [-1, 1],$$

where $b_k > 0, k = 0, 1, \cdots$, and

$$\sum_{k=0}^{\infty} b_k < \infty.$$

Remark 1.3. *The Gegenbauer polynomials $P_k^{(\lambda)}$ are orthogonal with respect to the weight $(1-t^2)^{\lambda-1/2}$. We will explain why this is the family of polynomials for the sphere S^{d-1} (when λ=(d-3)/2) in Chapter 2.*

Remark 1.4. *For our purposes we shall require that the coefficients b_k, $k = 0, 1, \cdots$, be all positive so that ϕ is strictly positive definite. Schoenberg [**?**] showed that nonnegative coefficients gave rise to a positive definite function. The proof uses the addition formula for Gegenbauer polynomials (see (1.1.3) below).*

Definition 1.4. *Let Δ^* be the Laplace-Beltrami operator, defined for all sufficiently smooth functions f on S^{d-1} as*

$$\Delta^* u(x) := \Delta \tilde{u}(x), x \in S^{d-1},$$

where Δ is the Laplacian in $\mathbf{R}^d$ and $\tilde{u}(x) := u(x/|x|), x \in \mathbf{R}^d/\{0\}$.

Definition 1.5. *The space of spherical harmonics of degree k consists of all infinitely differentiable functions that are eigenfunctions of Δ^* in the eigenvalue problem:*

$$\Delta^* u + \lambda u = 0,$$

corresponding to an eigenvalue $\lambda_k = k(k + d - 2)$, where k is a nonnegative integer. Let the dimension of this space be denoted by d_k.

Given an orthonormal basis $\{Y_{kl} : l = 1, \dots, d_k\}$ for the space of spherical harmonics of degree k, every square integrable function $u \in L_2(S^{d-1})$ has a Fourier series

$$u = \sum_{k=0}^{\infty} \sum_{l=1}^{d_k} u_{kl} Y_{kl},$$

with equality holding in $L_2(S^{d-1})$. The Fourier coefficients

$$u_{kl} = \int_{S^{d-1}} u(x) Y_{kl}(x) d\mu(x),$$

where μ is the normalised surface measure on S^{d-1}.

The addition formula for spherical harmonics is (see [**?**])

$$P_k^{((d-3)/2)}(xy) = \frac{P_k^{((d-3)/2)}(1)}{d_k} \sum_{l=0}^{d_k} Y_{kl}(x) Y_{kl}(y). \tag{1.1.3}$$

Definition 1.6. *(Second definition) The native space associated with the function ϕ is the space of all functions $u \in L_2(S^{d-1})$ for which*

$$\|u\|_\phi := \left(\sum_{k=0}^{\infty} \frac{d_k}{b_k \omega} \sum_{l=1}^{d_k} u_{kl}^2 \right)^{1/2} < \infty,$$

*where ω is the surface area of S^{d-1}. More details can be found for example in Morton and Neamtu [**?**].*

For a function in the native space we get an error estimate of a particular form

$$|s_Y(x) - f(x)| \leq P(x) \|f\|_\phi,$$

where

$$P(x) = \inf_{\{c_y : y \in \mathbf{R}\}} \left\{ \phi(0) - 2 \sum_{y \in Y} c_y \phi(d(x,y)) + \sum_{y,z \in Y} c_y c_z \phi(d(z,y)) \right\}^{1/2}$$

is termed the *Power Function*. The coefficients $\{c_y, y \in Y\}$ are free to be chosen, and are selected to annihilate spherical harmonics of some specific degree. More details can be found in, for example, Jetter et al. [**?**].

The research on which this thesis is based is concerned with some aspects of approximating on manifolds. The earlier part of the research addresses interpolation on spheres (hyperspheres) using radial basis functions. The result obtained is of a technical nature and is a calculation of the Fourier-Gegenbauer coefficients for three radial basis functions. The proofs use elementary techniques to calculate the integrals representing the corresponding Gegenbauer expansion coefficients. The latter part of the research is more theoretical in nature. It involves the generalisation to manifolds of a result specific to spheres. The new result is the calculation of an estimate of the convergence rate for interpolation on manifolds using G-invariant kernels; these are a generalisation of radial basis functions and zonal functions on the sphere.

1.2 Differential Geometry

This thesis concerns itself with differentiable manifolds. Spheres are themselves a particular kind of manifold.

Definition 1.7. *A topological manifold M is a topological space which is*

(i) Hausdorff, i.e. any 2 points may be surrounded by disjoint open sets,

(ii) every open set is homeomorphic to an open subset of some Euclidean space.

Definition 1.8. *A mapping ϕ is a diffeomorphism if it is differentiable and has a differentiable inverse.*

Let M be a topological manifold, then any pair (U, ϕ), where U is an open set of a manifold M and ϕ is a homeomorphism of U to an open subset of $\mathbf{R}^n$, is called a coordinate neighbourhood. Two coordinate neighbourhoods , (U, ϕ) and (V, ψ) are C^∞-compatible if $U \bigcap V$ is non-empty and $\phi\psi^{-1}$ and $\psi\phi^{-1}$ are diffeomorphisms of the open subsets $\psi(U \bigcap V)$ and $\phi(U \bigcap V)$ of $\mathbf{R}^n$. A differentiable or C^∞(or smooth) structure on a topological manifold M is a family N=$\{U_\alpha, \phi_\alpha : \alpha \in I\}$ of coordinate neighbourhoods that

(1) the U_α cover M,

(2) for any $\alpha, \beta \in I$ the neighbourhoods U_α, ϕ_α and U_β, ϕ_β are C^∞-compatible,

(3) any coordinate neighbourhood (V, ψ) compatible with every $(U_\alpha, \phi_\alpha) \in N$ is itself in N.

To verify whether a cover of neighbourhoods constitutes a manifold, however, it is only necessary to determine whether a covering of neighbourhoods is compatible, as the following theorem pertains:

Theorem 1.1. *Let M be a Hausdorff space with a countable basis of open sets. If $\{V_\beta, \psi_\beta\}$ is a covering of M by C^∞-compatible neighbourhoods, then there is a unique differentiable structure on M containing these coordinate neighbourhoods.*

For further details see Helgason [11].

Here is a proof that a particular mathematical object is a manifold, the real projective space of dimension d, from Boothby [**?**].

Definition 1.9. *Let X be a topological space and $\sim$ an equivalence relation on X. Denote by $[x] = \{y \in X | y \sim x\}$ the equivalence class of x, and for a subset $A \subset X$, denote by $[A]$ the set $\bigcup_{a \in A}[a]$, that is, all x equivalent to some element of A. We let $X/\sim$ stand for the set of equivalence classes and denote by $\pi : X \to X/\sim$ the natural mapping (projection) taking each $x \in X$ to its equivalence class, $\pi(x) = [x]$. With these notations we define the standard quotient topology on $X/\sim$ as follows: $U \subset X/\sim$ is an open subset if $\pi^{-1}(U)$ is open; the projection π is then continuous. With this notation and toplology we shall call $X/\sim$ the quotient space of X relative to the relation $\sim$.*

Definition 1.10. *An equivalence relation $\sim$ on a space X is called open if whenever a subset $A \subset X$ is open, then $[A]$ is also open.*

Lemma 1.1. *An equivalence relation $\sim$ on X is open if and only if π is an open mapping. When $\sim$ is open and X has a countable basis of open sets, then $X/\sim$ has a countable basis also.*

Proof: Let $A \subset X$ be an open subset. Since $[A] = \pi^{-1}(\pi(A))$, we see by definition of the quotient topology on $X/\sim$ that $[A]$ is open if π is open and conversely $[A]$ open implies $\pi(A)$ is open. Now suppose $\sim$ is open and X has a countable basis $\{U_i\}$ of open sets. If W is an open subset of $X/\sim$, then $\pi^{-1}(W) = \bigcup_{j \in J} U_j$ for some subfamily of $\{U_i\}$ and $W = \pi(\pi^{-1}(W)) = \bigcup_{j \in J} \pi(U_j)$. It follows that $\{\pi(U_i)\}$ is a basis of open sets for $X/\sim$. ■

Lemma 1.2. *Let $\sim$ be an open equivalence relation on a topological space X. Then $R = \{(x,y)|x \sim y\}$ is a closed subset of the space $X \times X$ if and only if the quotient space $X/\sim$ is Hausdorff.*

Proof: Suppose $X/\sim$ is Hausdorff and suppose (x,y) is not a member of R, that is, it is not the case that $x \sim y$. Then there are disjoint neighbourhoods U of $\pi(x)$ and V of $\pi(y)$. We denote by U^* and V^* the open sets $\pi^{-1}(U)$ and $\pi^{-1}(V)$, which contain x and y, respectively. If the open set $U^* \times V^*$ containing (x,y) intersects R, then it must contain a point (x',y') for which $x' \sim y'$, so that $\pi(x') = \pi(y')$ contrary to the assumption that U and V are disjoint. This contradiction shows that $U^* \times V^*$ does not intersect R and that R is closed.

Conversely, suppose that R is closed, then given any distinct pair of points $\pi(x)$, $\pi(y)$ in $X/\sim$, there is an open set of the form $U^* \times V^*$ containing (x,y) and having no point in R. It follows that $U = \pi(U^*)$ and $V = \pi(V^*)$ are disjoint. The previous lemma and the hypothesis imply that U and V are open. Thus $X/\sim$ is Hausdorff. ■

Definition 1.11. *Let $X = \mathbf{R}^{d+1}/0$, all $(d+1)$-tuples of real numbers $x = (x^1, \ldots, x^{d+1})$ except 0=(0,...,0) and define $x \sim y$ if there is a real number $t \neq 0$ such that $y = tx$, that is,*

$$(y_1, \ldots, y_{d+1}) = (tx_1, \ldots, tx_{d+1}).$$

The equivalence classes $[x]$ may be visualised as lines through the origin. We denote the quotient space by $P^d(\mathbf{R})$; it is called real projective space.

Theorem 1.2. *$P^d(\mathbf{R})$ is a differentiable manifold of dimension d.*

Proof: To do so we first note that $\pi : X \to P^d(\mathbf{R})$ is an open mapping. If $t \neq 0$ is real number, let $\phi_1 : X \to X$ be the mapping defined by $\phi_t(x) = tx$. It is clearly a homeomorphism with $\phi_t^{-1} = \phi_{1/t}$. If $U \subset X$ is an open set, then $[U] = \bigcup \phi_t(U)$, the union being over all real $t \neq 0$. Since each $\phi_t(U)$ is open, $[U]$ is open and π is open by lemma . Next we apply lemma to prove $P^d(\mathbf{R})$ is Hausdorff. On the open submanifold $X \times X \subset \mathbf{R}^{d+1} \times \mathbf{R}^{d+1}$ we define a real-valued function $f(x,y)$ by

$$f(x^1,\ldots,x^{d+1};y^1,\ldots,y^{d+1}) = \sum_{i\neq j}(x^i y^j - x^j y^i)^2.$$

Then $f(x,y)$ is continuous and vanishes if and only if $y = tx$ for some real number $t \neq 0$, that is, if and only if $x \sim y$. Thus

$$R = \{(x,y)|x \sim y\} = f^{-1}(0)$$

is a closed subset of $X \times X$ and $P^d(\mathbf{R})$ is Hausdorff. ■

We define $d+1$ coordinate neighbourhoods $U_i, \phi_i i = 1,\ldots,d+1$, as follows: Let $U_i^* = \{x \in X|x^i \neq 0\}$ and $U_i = \pi(U_i^*)$. Then $\phi_i : U_i \to \mathbf{R}^n$ is defined by choosing any $x = (x^1,\ldots,x^{d+1})$ representing $[x] \in U_i$ and putting

$$\phi_i(x) = \left(\frac{x^1}{x^i},\ldots,\frac{x^{i-1}}{x^i},\frac{x^{i+1}}{x^i},\ldots,\frac{x^{d+1}}{x^i}\right).$$

It is seen that if $x \sim y$, then $\phi_i(x) = \phi_i(y)$; moreover $\phi_i(x) = \phi_i(y)$ implies $x \sim y$. Thus $\phi_i : U_i \to \mathbf{R}^d$ is properly defined, continuous, one-to-one, and even onto. For $z \in \mathbf{R}^d, \phi_i^{-1}(z)$ is given by composing a C^∞ map of $\mathbf{R}^d$ to $\mathbf{R}^{d+1}$ with π namely , $\phi_i^{-1}(z^1,\ldots,z^d) = \pi(z^1,\ldots,z^{i-1},+1,z^i,\ldots,z^d)$; therefore π_i^{-1} is continuous. Thus $P^d(\mathbf{R})$ is a topological manifold and is C^∞ as the coordinate neighbourhoods are C^∞-compatible, that is $\phi_i\phi_j^{-1}$ is C^∞ (where defined) for $1 \leq i,j \leq d+1$.

Definition 1.12. *The geodesic distance between $x,y \in M$ is the length of the shortest path on the manifold which connects x to y. We will denote this by $d(x,y)$.*

1.3 Compact Homogeneous Manifolds

Spheres are themselves examples of a more general object, compact homogeneous manifolds. The starting point for understanding these objects is the concept of the Lie group. Lie groups combine the notions of group and manifold in the following manner:

Definition 1.13. *Let G be a group with binary operation represented by multiplicative notation. Let G also be a manifold. Then it is a Lie group if the operations:*

$$x \to x^{-1}$$

and

$$(x, y) \to xy$$

are themselves C^∞ mappings.

Thus Lie groups combine the structures of group and manifold. The following examples are from Baker [**?**].

Example 1.1. *$Gl(n, \mathbf{R})$, the set of nonsingular $n \times n$ matrices, is a submanifold of $M_n(\mathbf{R})$, the set of $n \times n$ real matrices, identified with $\mathbf{R}^{n^2}$. However, $Gl(n, \mathbf{R})$ is also a group, with respect to matrix multiplication, the real general linear group of order n. In fact, an $n \times n$ matrix A is nonsingular if and only if $\det A \neq 0$. However, $\det(AB) = (\det A)(\det B)$, so if A and B are nonsingular AB is also. An $n \times n$ matrix A is nonsingular, that is, $\det A \neq 0$, if and only if A has a multiplicative inverse. Matrix multiplication is associative and the multiplicative identity I_n is the group identity. Thus $Gl(n, \mathbf{R})$ is a group as well as a manifold.*

Both the maps $(A, B) \to AB$ and $A \to A^{-1}$ are C^∞. The product has entries which are polynomials in the entries of A and B, and thus the product mapping is C^∞. The inverse of $A = (a_{ij})$ may be written as $A^{-1} = (\frac{1}{\det A})(a^*_{ij})$, where the (a^*_{ij}) are the cofactors of A (and thus polynomials in the entries of A) and where $\det A$ is a polynomial in these entries which does not vanish on $Gl(n, \mathbf{R})$. Thus the entries of A^{-1} are rational functions on $Gl(n, \mathbf{R})$ with nonvanishing denominators, and hence C^∞. Therefore $Gl(n, \mathbf{R})$ is a Lie group.

A special case is $Gl(1, \mathbf{R})$=$\mathbf{R}^*$, the multiplicative group of nonzero real numbers.

Example 1.2. *Let $\mathbf{C}^*$ be the nonzero complex numbers. Then $\mathbf{C}^*$ is a group with respect to the multiplication of complex numbers, the inverse being $z^{-1} = \frac{1}{z}$. Also $\mathbf{C}^*$ is a two-dimensional C^∞ manifold covered by a single coordinate neighbourhood*

$U = \mathbf{C}^*$ with coordinate map $z \to \phi(z)$ given by $\phi(x+iy) = (x,y)$ for $z = x+iy$. Using these coordinates, the product $w = zz', z = x+iy$ and $z' = x'+iy'$ is given by

$$((x,y)(x',y')) \to (xx' - yy', xy' + yx')$$

and the mapping $z \to z^{-1}$ by

$$(x,y) \to \left(\frac{x}{x^2+y^2}, \frac{-y}{x^2+y^2}\right).$$

Thus the product map and the inverse map are C^∞ and therefore $\mathbf{C}^*$ is a Lie group.

Definition 1.14. *A subgroup $G \leq Gl_n(\mathbf{R})$ which is also a closed subspace is a matrix group over $\mathbf{R}$ or a real matrix group.*

The following examples are due to Baker [**?**].

Example 1.3. *For $n \geq 1$, an $n \times n$ real matrix A for which $A^T A = I_n$ is called an orthogonal matrix; here A^T is the transpose of $A = [a_{ij}]$, whose entries are given by*

$$(A^T)_{ij} = a_{ji}.$$

Such an orthogonal matrix has an inverse, namely A^T, and the product of two orthogonal matices A, B is orthogonal since

$$(AB)^T(AB) = B^T A^T AB = B^T I_n B = B^T B = I_n.$$

I_n is an orthogonal matrix. Thus the subset

$$O(n) = \{A \in GL_n(\mathbf{R}) : A^T A = I_n\} \subset M_n(\mathbf{R})$$

is a subgroup of $GL_n(\mathbf{R})$ and is called the $n \times n$ real orthogonal group. The single matrix equation $A^T A = I_n$ is equivalent to the n^2 polynomial equations:

$$\sum_{k=1}^{n} a_{ki} a_{kj} = \delta_{ij}$$

for the n^2 real numbers a_{ij}, where the Kronecker symbol δ_{ij} is defined by

$$\delta_{ij} = 1 \text{ if } i = j, \delta_{ij} = 0 \text{ if } i \neq j.$$

*Since polynomial functions are continuous we can express $O(n)$ as a closed subset of $GL_n(\mathbf{R})$. Hence $O(n)$ is a matrix group. For further details see [**?**].*

Example 1.4. *Consider the determinant function restricted to $O(n)$. For $A \in O(n)$,*

$$(\det A)^2 = \det A^T \det A = \det(A^T A) = \det I_n = 1,$$

which implies that $\det A$*=+1 or det A=-1. Thus we have*

$$O(n) = O(n)^+ \bigcup O(n)^-,$$

where

$$O(n)^+ = \{A \in O(n) : \det A = 1\}, O(n)^- = \{A \in O(n) : det A = -1\}.$$

Notice that

$$O(n)^+ \bigcap O(n)^- = \emptyset,$$

so $O(n)$ is the disjoint union of the subsets $O(n)^+$ and $O(n)^-$. The subgroup

$$SO(n) = O(n)^+ \leq O(n)$$

is the $n \times n$ special orthogonal group.

Example 1.5. *For $A = [a_{ij}] \in M_n(\mathbf{C})$,*

$$A^* = (\overline{A})^T,$$

is the hermitian conjugate of A, i.e.,$(A^*)_{ij} = \bar{a}_{ji}$. The $n \times n$ unitary group is the subgroup

$$U(n) = \{A \in Gl_n(\mathbf{C}) : A^* A = I\} \leq Gl_n(\mathbf{C}).$$

Theorem 1.3. *Let $G \leq Gl_n(\mathbf{R})$ be a matrix subgroup. Then G is a Lie subgroup of $Gl_n(\mathbf{R})$.*

Proof: See [**?**]. ■

The following is also from [**?**]:

Definition 1.15. *Let M be a smooth manifold of dimension d and $p \in M$. Let $\gamma : (a, b) \to M$ be a continuous curve with $a < 0 < b$. γ is differentiable at $t \in (a, b)$ if for every chart $\phi : U \to V$ is differentiable at $t \in (a, b)$, i.e., $(\phi \circ \gamma)'(t)$ exists. The curve γ is smooth at $t \in (a, b)$ if all the derivatives of $\phi \circ \gamma$ exist at t. The curve γ is differentiable if it is differentiable at all points in (a, b). Similarly γ is smooth if it is smooth at all points in (a, b).*

We will now change the notation to $fg = f \circ g$ for the composition of functions f and g when no confusion seems likely to result.

Lemma 1.3. *Let $\phi_0 : U_0 \to V_0$ be a chart with $\gamma(t) \in U_0$ and suppose that*

$$\phi_0 \circ \gamma : (a, b) \bigcap \gamma^{-1}\phi_0^{-1}V_0 \to V_0$$

is differentiable (respectively smooth) at t. Then for any chart $\phi : U \to V$ with $\gamma(t) \in U$,

$$\phi \circ \gamma : (a, b) \bigcap \gamma^{-1}\phi^{-1}V \to V$$

is differentiable (respectively smooth) at t.

Proof: See [**?**]. ■

The chain rule for the derivative applies here:

$$(\phi\gamma)'(t) = J_{\phi\phi_0^{-1}}(\phi_0\gamma(t))(\phi_0\gamma)'(t).$$

Here, for a differentiable function

$$h : W_1 \to W_2; h(x) = [h_1(x) \ldots h_{m_2}(x)],$$

where $W_1 \subset \mathbf{R}^{m_1}$ and $W_2 \subset \mathbf{R}^{m_2}$ are open subsets, $x \in W_1$,

$$J_h(x) = \left[\frac{\partial h_i}{\partial x_j}(x)\right] \in M_{m_2,m_1}(\mathbf{R})$$

is the Jacobian matrix of h at x.

If $\gamma(0) = p$ and γ is differentiable at 0 then for any (and hence every) chart $\phi_0 : U_0 \to V_0$ with $\gamma(0) \in U_0$, there is a derivative vector $v_0 = (\phi_0\gamma)'(0) \in \mathbf{R}^n$. In passing to another chart $\phi : U \to V$ with $\gamma(0) \in U$ we have

$$(\phi\gamma)'(0) = J_{\phi\phi_0}(\phi_0\gamma(0))(\phi_0\gamma)'(0).$$

In order to define the tangent space T_pM to the manifold M at p, we consider all pairs of the form

$$((\phi\gamma)'(0), \phi : U \to V)$$

where $\gamma(0) = p \in U$, and then impose an equivalence relation $\sim$ under which

$$((\phi_1\gamma)'(0), \phi_1 : U_1 \to V_1) \sim ((\phi_2\gamma)'(0), \phi_2 : U_2 \to V_2).$$

Since

$$(\phi_2\gamma)'(0) = J_{\phi_2\phi_1^{-1}}(\phi_1\gamma(0))(\phi_1\gamma)'(0),$$

we can also write this as

$$(v, \phi_1 : U_1 \to V_1) \sim (J_{\phi_2\phi_1^{-1}}(\phi_1(p)v, \phi_2 : U_2 \to V_2),$$

whenever there is a curve γ in M for which

$$\gamma(0) = p, (\phi_1\gamma)'(0) = v.$$

The set of equivalence classes is the tangent space T_pM.

Proposition 1.1. *For $p \in M$, T_pM is a real vector space of dimension d.*

Proof: See [?]. ■

The following is also from [?]. Let $h : M \to M'$ be a smooth map between manifolds of dimension n, n'. For $p \in M$ consider a pair of charts with $p \in U_\alpha$ and $h(p) \in U'_{\alpha'}$. Since $h_{\alpha',\alpha} = \phi'_{\alpha'} \circ h \circ \phi_\alpha^{-1}$ is differentiable, the Jacobian matrix $J_{h_{\alpha'\alpha}}(\phi_\alpha(p))$ has an associated real linear transformation

$$\underset{\cdot}{h}_{\alpha'\alpha} : \mathbf{R}^n \to \mathbf{R}^{n'}; \underset{\cdot}{h}_{\alpha'\alpha}(x) = J_{h_{\alpha'\alpha}}(\phi_\alpha(p))x.$$

This passes to equivalence classses to give a well-defined real linear transformation

$$\underset{\cdot}{h}_p : T_pM \to T_{h(p)}M'.$$

Definition 1.16. *An action μ of a group G on a set X is a function $\mu{:}G \times X \to X$ for which we usually write $\mu(g, x) = gx$ if there is no danger of ambiguity, satisfying the following conditions for all $g, h \in G$ and $x \in X$ and with ι being the identity element of G:*

(i) $(gh)x = g(hx)$, i.e., $\mu(g, \mu(h, x))$;

(ii) $\iota x = x$.

There are two important notions associated to such an action.

Definition 1.17. *For $x \in X$, the stabilizer of x is*

$$\mathrm{Stab}_G(x) = g \in G : gx = x \subset G,$$

while the orbit of x is

$$\mathrm{Orb}_G(x) = gx \in X : g \in G \subset X.$$

The following is from [**?**]. Let f be a real-valued function defined on an open subset W_f of a C^∞ manifold M, possibly all of M ; in brief, $f : W_f \to \mathbf{R}$. If U, ϕ is a coordinate neighbourhood such that $W_f \bigcap U \neq$ and if $x^1 \ldots x^n$ denotes the local coordinates, then f corresponds to a function $f^*(x^1, \ldots x^n)$ on $\phi(W_f \bigcap U)$ defined by $f^* = f \circ \phi^{-1}$, that is, so that $f(p) = f^*(x^1(p) \ldots x^n(p)) = f^*(\phi(p))$ for all $p \in W_f \bigcap U$.

Definition 1.18. *Using the notation above, $f : W_f \to \mathbf{R}$ is a C^∞ function if each $p \in W_f$ lies in a coordinate neighbourhood U, ϕ such that $f \circ \phi^{-1}(x^1 \ldots x^n = f^*(x^1, \ldots, x^n)$ is infinitely differentiable on $\phi(W_f \bigcap U)$.*

Definition 1.19. *Suppose that M and N are C^∞ manifolds, $W \subset M$ is an open subset, and $F : W \to N$ is a mapping. F is a C^∞ mapping of W into N if for every $p \in W$ there exist coordinate neighbourhoods U, ϕ of p and V, ψ of $F(p)$ with $F(U) \subset V$ such that $\psi \circ F \circ \phi^{-1} : \phi(U) \to \psi(V)$ is infinitely differentiable.*

Definition 1.20. *Let G be a Lie group and let $g \in G$. Then left multiplication by g is a function defined by:*

$$L_g : G \to G; L_g(x) = gx.$$

Proposition 1.2. *For each $g \in G$, the maps L_g are diffeomorphisms with inverse*

$$L_g^{-1} = L_{g^{-1}}.$$

Proof: See [**?**]. ■

Definition 1.21. *A vector field X on a manifold M is a function assigning to each point p of M a vector $v_p \in T_pM$ whose components in the bases of any local coordinates (or patch) U, ϕ are C^∞ functions on the domain U of the coordinates.*

For further details see [**?**]. The following is from [**?**].

Definition 1.22. *If $F : M \to M$ is a diffeomorphism and X is a vector field on M such that $dF(X) = X$, then X is said to be invariant with respect to F. If X is invariant with respect to left translations then it is said to be left-invariant.*

Definition 1.23. *A real vector space, L is a (real) Lie algebra if it possesses an additional product, a map $L \times L \to L$, taking the pair (X, Y) to the element $[X, Y]$ of L, which has the following properties:*

(1) it is bilinear over $\mathbf{R}$

$$[\alpha_1 X_1 + \alpha_2 X_2, Y] = [\alpha_1 X_1, Y] + [\alpha_2 X_2, Y],$$

$$[X, \alpha_1 Y_1 + \alpha_2 Y_2] = [X, \alpha_1 Y_1] + [X, \alpha_2 Y_2];$$

(2) it is skew-commutative

$$[X, Y] = -[Y, X];$$

(3) it satisfies the Jacobi identity:

$$[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0.$$

Example 1.6. *The vector space $\mathbf{R}^3$, of dimension 3 over $\mathbf{R}$, with the usual vector product of vector calculus is a Lie algebra.*

Example 1.7. *Let $M_n(\mathbf{R})$ denote the algebra of $n \times n$ matrices over $\mathbf{R}$ with XY denoting the usual matrix product of X and Y. Then $[X, Y] = XY - YX$, the 'commutator' of X and Y, defines a Lie algebra structure on $M_n(\mathbf{R})$.*

Definition 1.24. *The exponential* $\exp X$ *of a matrix X is defined to be the matrix given by*

$$\exp X = I + X + \frac{X^2}{2!} + \frac{X^3}{3!} + \dots$$

if the series converges.

Theorem 1.4. *If G is a Lie group, then the left-invariant vector fields on G form a Lie algebra g and* $\dim g$=$\dim G$. *(This is the Lie algebra g of a Lie group G.)*

Proof: See [?]. ■

Theorem 1.5. *For a matrix group $G \leq GL_n(\mathbf{R})$ the exponential map* $\exp : g \to M_n(\mathbf{R})$ *has image in G,* $\operatorname{Im}\exp \leq G$.

Proof: See [?]. ■

Definition 1.25. *Consider a manifold M which is the orbit of a fixed point $\eta \in \mathbf{R}^m$ for some m, under the action of a Lie group G, i.e. $M = \{g \circ \eta : g \in G\}$. There is a (possibly trivial) subgroup of symmetries $H \leq G$ such that $H\eta = \eta$. Then we can identify*

$$M \equiv \frac{G}{H}.$$

It is clear from this definition that each point on the manifold can be mapped to any other point using G. We call this property transitivity.

Example 1.8. *Spheres are simple examples of this structure. Consider the unit sphere embedded in 3-dimensions, S^2. Take an element of the sphere $\eta = (0,0,1)$, which we call the north pole, and act on this element with the group of rotations of $\mathbf{R}^3$,* $\mathrm{SO}(3)$. *Then this action will generate the whole sphere. In other words*

$$S^2 = \{g\eta : g \in \mathrm{SO}(3)\}.$$

However, rotations with axis through η fix η so that $H = \mathrm{SO}(2)$, and we can identify

$$S^2 \equiv \frac{\mathrm{SO}(3)}{\mathrm{SO}(2)}.$$

For the next three results you can find proofs in [?].

Definition 1.26. *Let X be a topological space. X is connected if whenever $X = U \bigcup V$ with $U, V \neq \emptyset$ both open subsets, then $U \bigcap V \neq \emptyset$. X is path connected if whenever $x, y \in X$, there is a continuous path $p : [0,1] \to X$ with $p(0) = x$ and $p(1) = y$.*

Proposition 1.3. *If X is a path connected topological space then X is connected.*

Definition 1.27. *Let G be a Lie group. Two elements x, y are said to be connected by a path in G if there is a continuous path $p[0,1] \to G$ with $p(0) = x$ and $p(1) = y$; we will then write $x \sim_G y$.*

Lemma 1.4. *$\sim_G$ is an equivalence relation on G.*

For $g \in G$, we can consider the equivalence class of g, the path component of g in G,

$$G_g = \{x \in G : x \sim_G y\}.$$

Proposition 1.4. *The path component of the identity is a closed and open normal subgroup of G; hence it is a closed Lie subgroup of dimension $dimG$. It is also known as the connected component of the identity.*

In the next example we consider the complex projective space

$$\mathbf{P}^d(\mathbf{C}) = \{v \in \mathbf{C}^{d+1} \ : \ \|v\| = 1\}/\{f \in \mathbf{C} \ : \ |f| = 1\}.$$

This is not a homogeneous embedding since this space has edges in a manner analogous to the real projective space. In the real case the component equivalence classes are composed of opposite points on a unit sphere, thus an orbit can never be smooth as such as there are continual 'discontinuities' as the orbit proceeds. Below we view $\mathbf{P}^d(\mathbf{C})$ as the orbit under conjugation.

Example 1.9. *Let $\mathcal{H}_{d+1}$ be the square Hermitian matrices of size $d+1$ over $\mathbf{C}$, i.e. for $H \in \mathcal{H}$, $H_{i,j} = H^*_{j,i}$, $i, j = 1, \ldots, d+1$. Let $\mathcal{U}_{d+1}$ be the space of unitary matrices over $\mathbf{C}$, i.e. for $U \in \mathcal{U}_{d+1}$, $UU^* = I_{d+1}$. Let $\mathcal{U}^0_{d+1}$ be the connected component of the identity in $\mathcal{U}_{d+1}$. Then $\mathcal{U}^0_{d+1}$ acts on $\mathcal{H}_{d+1}$ via conjugation $U \circ H = UHU^*$, since*

$$(U \circ H)^* = (UHU^*)^* = UH^*U^* = UHU^* = U \circ H.$$

The inner product of $U, V \in \mathcal{H}_{d+1}$ is $\mathrm{Re}\,(\mathrm{Trace}\,(UV^*))$. If we set $\eta_{1,1} = 1$, $\eta_{i,j} = 0$ otherwise, then the north pole $\eta \in \mathcal{H}_{d+1}$ and $U(\eta) = \{vv^* : v \text{ is the first column of } U \in \mathcal{U}^0_{d+1}\}$. Since any vector of length 1 could potentially be the first column in a matrix from $\mathcal{U}^0_{d+1}$ we have $\mathcal{U}^0_{d+1}(\eta) = \{vv^* : |v| = 1\}$. Two vectors v and u give rise to the same point in $\mathcal{U}^0_{d+1}(\eta)$ if $v = \alpha u$ for some $\alpha \in \mathbf{C}$, with $|\alpha| = 1$.

The subgroup which leaves the north pole fixed is the set of unitary matrices U with $U_{11} = 1$, and $U_{1j} = U_{j1} = 0$, $j = 2, \cdots, d+1$.

Example 1.10. *The flat torus d-dimensional torus $T^d = S^1 \times S^1 \times \cdots S^1$ (d times). Since each S^1 is a subset of $\mathbf{R}^2$ we have a homogeneous embedding in $\mathbf{R}^{2d}$. We remark here that the usual embedding of T^2 in $\mathbf{R}^3$ is not homogeneous since there is clearly an inside and outside of the torus.*

As outlined in the section on the sphere, the fundamental process underlying this thesis is that of interpolation. The basic problem of interpolation on a compact manifold M is as follows: Given a set of nodes, $Y \subset M$, and a set of data, $\{f_y, y \in Y\}$, find a function s_Y, from a prescribed set of functions such that:

$$s_Y(y) \quad = \quad f_y, \quad y \in Y.$$

As was proven by Mairhuber [?], it is not possible in general to interpolate from a fixed finite subspace in more than one dimension. Thus we would like the space of interpolants to depend upon the set of data points. As we will see, interpolation is then a reliable process.

In the sphere example above we have the so-called zonal kernel and we wish to generalise this with the G-invariant kernel.

Definition 1.28. *The function $\kappa : M \times M \to \mathbf{R}$ is called a G-invariant kernel if $\kappa(gx, gy) = \kappa(x, y)$ for all $x, y \in M$ and $g \in G$.*

We will construct our approximants from translates of a G-invariant kernel:

$$s_Y(x) = \sum_{y \in Y} c_y \kappa(x, y).$$

When the interpolation conditions are imposed for data $\{f_y : y \in Y\}$ we obtain the following set of linear equations

$$f_z = \sum_{y \in Y} c_y \kappa(z, y), \quad z \in Y,$$

which is a system of n equations in n unknowns. As before, we can write this in matrix form as the equation $A\mathbf{c} = \mathbf{f}$, where

$$A_{zy} = \kappa(z, y),$$

is known as the interpolation matrix.

Of course, we require the interpolation matrix to be nonsingular in order to solve the interpolation problem. One way of doing this involves the notion of positive definiteness.

Definition 1.29. *The kernel $\kappa : M \times M \to \mathbf{R}$ is said to be positive definite on M if for any finite set of points $Y \subset M$*

$$\mathbf{c}^t A\mathbf{c} = \sum_{y,z \in Y} c_z c_y \kappa(z, y) \geq 0$$

for all $\mathbf{c} \in \mathbf{R}^n$. If $\mathbf{c}^t A\mathbf{c} > 0$ whenever the points in Y are distinct and $\mathbf{c} \neq 0$, then we say that f is strictly positive definite. In these equations $\mathbf{c}^t$ denotes the transpose of $\mathbf{c}$.

As in Section 1.1 we will need to determine what the spaces of functions are which can be approximates by such s_Y and this will be explored in the next chapter.

1.3.1 Two point spaces

The sphere is the most familiar example of the compact two point homogeneous spaces. These are spaces with the property that for any two pairs of points x, y and w, z in M, with $d(x, y) = d(w, z)$ (where d is the geodesic distance on M), there is an element of the group G for which $gx = w$ and $gy = z$. The main simplification we achieve approximating on such spaces is that the G-invariant kernels are all univariate functions of distance alone. This fact is what is behind the addition formula (1.1.3) for the sphere.

Lemma 1.5. *Let κ be a G-invariant kernel of a two point homogeneous space M. Then*

$$\kappa(x, y) = \phi(d(x, y)),$$

for some univariate function ϕ.

Proof: Let $d(x,y) = d(w,z)$ for $x,y,w,z \in M$. Then, since M is two point, there is a $g \in G$ such that $gx = w$ and $gy = z$. Since κ is G-invariant,

$$\kappa(w,z) = \kappa(gx,gy) = \kappa(x,y).$$

Thus we see that κ is constant on pairs of points the same distance apart. Hence

$$\kappa(x,y) = \phi(d(x,y)),$$

for some $\phi : \mathbf{R}_+ \to \mathbf{R}$ with the obvious notation for the non-negative real numbers. ∎

1.4 Orthogonal Polynomials

As we saw in the example of the sphere above, orthogonal polynomials feature prominently in this research. In this section we give an introduction to the salient facts about such polynomials. These will be used in the next chapter.

Definition 1.30. *A set S in an inner product space X, with inner product $\langle \cdot, \cdot \rangle$ is said to be orthogonal if $\langle x,y \rangle = 0$ for all elements $x,y \in S$ whenever $x \neq y$. The set is called orthonormal if $\langle x,y \rangle = 1$ whenever $x = y$.*

Example 1.11. *The most familiar examples of an orthonormal set of functions is the trigonometric polynomials in the periodic functions $C[-\pi,\pi]$, with*

$$\langle f,g \rangle = \int_{-\pi}^{\pi} f(x)g(x)dx.$$

The polynomials $(2\pi)^{-1/2}$, $\pi^{-1/2}\cos nx$, $\pi^{-1/2}\sin nx$, $n = 1,2\cdots$.

Example 1.12. *Our second example is one of the family of Gegenbauer polynomials mentioned in Section 1.1. In C[-1,1] with*

$$\langle f,g \rangle = \int_{-1}^{1} \frac{f(x)g(x)}{(1-x^2)^{1/2}}dx,$$

the polynomials $T_n(x) = \cos(n\cos^{-1}x)$, $n = 0,1,2,\ldots$, form an orthogonal set. These are the Chebyshev polynomials (the Gegenbauer polynomials for the circle).

In the previous example we see functions which are orthogonal with respect to a weight function $w(t) > 0$. For the Chebyshev polynomials the weight is $(1-t^2)^{-1/2}$. More generally we have an inner product

$$\langle f, g \rangle = \int_a^b f(t)g(t)w(t)dt.$$

The orthogonal polynomials have an interesting and very useful property; they obey a three-term recurrence relation.

Theorem 1.6. *[?] The sequence of polynomials defined by $q_0(t) = 1, q_1(t) = t - a_1$*

$$a_n = \frac{\langle q_n, tq_{n-1}(t)\rangle}{\langle q_{n-1}(t), q_{n-1}(t)\rangle},$$

$$b_n = \frac{\langle tq_{n-1}(t), q_{n-1}(t)\rangle}{\langle q_{n-2}(t), q_{n-1}(t)\rangle}$$

and

$$q_n(t) = (t - a_n)q_{n-1}(t) - b_n q_{n-2}(t),$$

are orthogonal on $[-1, 1]$ with respect to $\langle \cdot, \cdot \rangle$.

Theorem 1.7. *[?] Let $q_0, q_1, \ldots$ be a sequence of monic polynomials orthogonal with respect to $\langle \cdot, \cdot \rangle$. Then $q_0(t) = 1, q_1(t) = t - a_1$ and*

$$q_n(t) = (t - a_n)q_{n-1}(t) - b_n q_{n-2}(t),$$

$n = 2, 3, \cdots$, where $a_n, b_n, \ldots$ are given in the statement of the theorem.

The material presented in this chapter is prefatory to the rest of this thesis. More material on Gegenbauer polynomials as special functions on spheres is to be found in the following chapter. Interpolation on spheres using radial basis functions is dealt with in Chapter 3 where the Fourier-Gegenbauer coefficients for two radial basis functions are calculated. Interpolation features again in Chapter 4 at a higher level of generality and it is here that the material presented in this introduction on differential geometry comes in useful where an error bound is found for compact homogeneous spaces.

CHAPTER 2

Harmonic Analysis and Special Functions

2.1 Introduction

In this section we will develop the harmonic analysis needed to investigate the interpolation problem which we consider in Chapter 4 and will follow Levesley and Ragozin [?] in this. We will show that on a compact homogeneous manifolds, an arbitrary G-invariant kernel can be decomposed into special G-invariant polynomials, which are the reproducing kernels for certain subspaces of polynomials. This development is analogous to the sphere case. For the two point homogeneous spaces we will see that such reproducing kernels are univariatate functions of the geodesic distance.

2.2 Polynomials

Let μ be an appropriately normalised G-invariant 'surface' measure on M. We can view this as the restirction of the Lebesgue measure from the ambient Euclidean space to the manifold, or the lifting of the Haar measure from the group G to the manifold. For more information on the Haar measure see, for instance, Hewitt and Ross [?]. Then for $f, g \in L_2(M)$ the following inner product may be defined

$$(f,g) = \int_M fg d\mu.$$

Definition 2.1. *Let*

$$\|f\|_p^p = \begin{cases} \int_M |f|^p d\mu, & 1 \leq p < \infty \\ \text{ess sup}|f|, & p = \infty. \end{cases}$$

We define the Lebesgue spaces of integrable functions

$$L_p(M) = \{f : \|f\|_p \leq \infty\}, \quad 1 \leq p \leq \infty.$$

Definition 2.2. *Let Π_n be the space of polynomials up to degree n in m variables (the manifold is homogeneously embedded in $\mathbf{R}^m$) , and $P_n = \Pi_n|_M$. Then the harmonic polynomials of degree n on M are defined to be $H_n = P_n \bigcap P_{n-1}^{\perp}$.*

Definition 2.3. *Let G be a matrix group, $g \in G$ and let M be a manifold on which G acts. Also let $f \in L_p(M)$ and V be a vector space . Then V is G-invariant if given $f \in V$ then $f(gx) \in V$ for all $g \in G$ and for all $x \in M$.*

Proposition 2.1. *P_n is G-invariant.*

Proof: The proof is by induction on n.

For the initial step let $n = 1$. Let $p(x) = ax + b$ be some arbitrarily chosen linear polynomial and let G be a matrix group, $g \in G$.

Then

$$p(gx) = a(gx) + b$$

and since gx also is in M, then this is simply another linear polynomial. So $p(gx) \in P_1$. This establishes the initial step.

Now suppose the proposition is true for P_{n-1}. Using the division algorithm we see that $p \in P_n$ can be written in the form

$$p(x) = xq(x) + r(x)$$

where $q, r \in P_{n-1}$. Then

$$p(gx) = (gx)q(gx) + r(gx).$$

However, by the initial step and the induction hypothesis,

$$gx \in P_1$$

and

$$q(gx), r(gx) \in P_{n-1}.$$

Thus

$$p(gx) \in P_n.$$

Therefore P_n is G-invariant. Since P_{n-1} is also G-invariant $H_n = P_n \bigcap P_{n-1}^{\perp}$ is also G-invariant. ■

Since H_n is finite dimensional we can decompose it into two finite subspaces in the following manner: Choose $q \in H_n$. Let $Q = \text{span}\{q(g\cdot) : g \in G\}$ and let P be the orthogonal complement of Q. Then P and Q are G-invariant subspaces. Let $q_1 \in P$ and iterate the process to give a decomposition into G-invariant subspaces. As H_n is finite dimensional at some point this process must terminate. Thus H_n can be uniquely decomposed into irreducible G-invariant subspaces $H_{nk}, k = 1, ..., \nu_n$ (subspaces with no proper G-invariant subspace).

Let $Y_{nk}^1, \ldots, Y_{nk}^{d_{nk}}$ be any orthonormal basis for H_{nk} and set

$$p_{nk}(x, y) = \sum_{j=1}^{d_{nk}} Y_{nk}^j(x) Y_{nk}^j(y).$$

Lemma 2.1. *p_{nk} is independent of the choice of Y_{nk}^j.*

Proof: Let $Z_{nk}^j, j = 1, \cdots, d_{nk}$ be a second orthonormal basis for H_{nk}. For this proof we will drop the nk subscript as there is no chance of confusion. Then

$$Z^j = \sum_{l=1}^{d} a_{jl} Y^l, \quad j = 1, \cdots, d.$$

Since Z^j are orthonormal we have

$$\begin{aligned} \delta_{jk} = \int_M Z^j Z^k d\mu &= \int_M \{\sum_{l=1}^{d} a_{jl} Y^l\}\{\sum_{m=1}^{d} a_{km} Y^m\} d\mu \\ &= \sum_{l=1}^{d} a_{jl} a_{kl}, \quad j, k = 1, \cdots, d, \end{aligned}$$

using orthonormality. Hence

$$\begin{aligned}\sum_{j=1}^{d} Z^j(x)Z^j(y) &= \sum_{j=1}^{d}\{\sum_{l=1}^{d} a_{jl}Y^l(x)\}\{\sum_{m=1}^{d} a_{jm}Y^m(y)\} \\ &= \sum_{l=1}^{d}\sum_{m=1}^{d} Y^l(x)Y^m(y)\sum_{j=1}^{d} a_{jl}a_{jm} \\ &= \sum_{l=1}^{d} Y^l(x)Y^l(y).\end{aligned}$$

Thus p_n is independent of the choice of basis. ■

The kernel p_{nk} is the unique kernel for the orthogonal projection T_{nk} onto H_{nk}:

$$T_{nk}f(x) = \int_M p_{nk}(x,y)f(y)d\mu(y).$$

We can restate this in terms of the so called reproducing kernel property.

Definition 2.4. *Let H be a Hilbert space of continuous functions defined on a set X with inner product $\langle \cdot, \cdot \rangle$. Then the kernel $\kappa : H \times H \to \mathbf{R}$ is called a reproducing kernel if*

$$f(x) = \langle f, \kappa(\cdot, x)\rangle$$

for all $x \in X$ and $f \in H$.

Thus p_{nk} is the reproducing kernel for H_{nk}. Additionally we have

Lemma 2.2. *p_{nk} is G-invariant.*

Proof: Since, for any $g \in G$ $\{Y^j_{nk}(gx), j = 1, \cdots, d_{nk}\}$ is another orthonormal basis for H_{nk} we have, using the previous lemma, for any $x, y \in M$ and $g \in G$,

$$\begin{aligned}p_n(gx, gy) &= \sum_{j=1}^{d_{nk}} Y^j_{nk}(gx)Y^j_{nk}(gy) \\ &= \sum_{j=1}^{d_{nk}} Y^j_{nk}(x)Y^j_{nk}(y) \\ &= p_n(x,y). \qquad ■\end{aligned}$$

Lemma 2.3. *For any $x, y \in M, n > 0$, and $1 \le k \le \nu_n$,*

$$p_{nk}(y,y) = \sum_{j=1}^{d_{nk}} (Y^j_{nk}(y))^2 = d_{nk}, \tag{2.2.1}$$

$$|p_{nk}(x,y)| \le d_{nk}. \tag{2.2.2}$$

Proof: By G-invariance and transitivity

$$p_{nk}(y,y) = p_{nk}(gx,gx) = p_{nk}(x,x)$$

for all $x \in M$ and $g \in G$. Thus

$$p_{nk}(y,y) = \int_M p_{nk}(y,y)d\mu(y) = \int_M \sum_{j=1}^{d_{nk}} (Y_{nk}^j(y))^2 d\mu(y) = d_{nk},$$

by orthonormality. Using the Cauchy-Schwartz inequality

$$|p_{nk}(x,y)| = \left|\sum_{j=1}^{d_{nk}} Y_{nk}^j(x) Y_{nk}^j(y)\right| \leq \left\{\sum_{j=1}^{d_{nk}} (Y_{nk}^j(x))^2\right\}^{1/2} \left\{\sum_{j=1}^{d_{nk}} (Y_{nk}^j(y))^2\right\}^{1/2} = d_{nk},$$

for $1 \leq k \leq \nu_n$ and $n > 0$. ■

Corollary 2.1.

$$\frac{\|T_{nk}f\|_\infty}{\|T_{nk}f\|_2} \leq (d_{nk})^{1/2}.$$

Proof: If we let

$$a_{nk}^j(f) = \int_M f Y_{nk}^j d\mu,$$

be the Fourier coefficient for Y_{nk}^j, using the Cauchy-Schwartz inequality we have

$$\begin{aligned} \|T_{nk}f\|_\infty &= \max_{y \in M} \left|\sum_{j=1}^{d_{nk}} a_{nk}^j(f) Y_{nk}^j(y)\right| \leq \left\{\sum_{j=1}^{d_{nk}} (a_{nk}^j(f))^2\right\}^{1/2} \left\{\sum_{j=1}^{d_{nk}} (Y_{nk}^j)^2\right\}^{1/2} \\ &= (d_{nk})^{1/2} \|T_{nk}f\|_2, \end{aligned}$$

for $1 \leq k \leq \nu_n$. ■

In order to prove an important spectral decomposition result for G-invariant kernels we first need some prelinary lemmas.

Lemma 2.4. *Let us fix $n \in \mathbf{N}$, $1 \leq k \leq \nu_n$ and normal $Y \in H_{nk}$. Then for any normal $Y^* \in H_{nk}$,*

$$\int_M \kappa(x,y) Y^*(x) Y^*(y) d\mu(x) d\mu(y) = \int_M \kappa(x,y) Y(x) Y(y) d\mu(x) d\mu(y).$$

Proof: First we have

$$H_{nk} = \text{span}\,\{Y(g\cdot) : g \in G\}$$

since span $\{Y(g\cdot) : g \in G\} \subset H_{nk}$ and is clearly a G-invariant subspace. However, H_{nk} is the smallest G-invariant subspace containing Y so they must be equal. Therefore, there exist $g_1, \cdots, g_{\nu_{nk}}$, such that

$$H_{nk} = \text{span}\,\{Y(g_l\cdot) : l = 1, \cdots, \nu_{nk}\}.$$

Suppose this lemma is not true. Then, there exists $Z \in H_{nk}$ such that $Z(x)Z(y)$ is orthogonal to $Y(gx)Y(gy)$ for all $g \in G$. Thus,

$$\int_M Z(x)Y(gx)Z(y)Y(gy)d\mu(x)d\mu(y) = 0$$

for all $g \in G$, so that

$$\int_M Z(x)Y(gx)d\mu(x) \int Z(y)Y(gy)d\mu(y) = 0$$

for all $g \in G$. This is clearly impossible as $H_{nk} = \text{span}\,\{Y(g\cdot) : g \in G\}$. Thus, for some $m \in \mathbf{N}$, and $h_1, \cdots, h_m \in G$,

$$Y^*(x)Y^*(y) = \sum_{l=1}^{m} \alpha_l Y(g_l x)Y(g_l y).$$

Setting $x = y$ in the last equation and integrating over M we see that

$$\sum_{l=1}^{m} \alpha_l = 1. \tag{2.2.3}$$

Thus, for any normal $Y^* \in H_{nk}$

$$\begin{aligned}
\int_M \kappa(x,y)Y^*(x)Y^*(y)d\mu(x)d\mu(y) &= \int_M \kappa(x,y) \sum_l \alpha_l Y(g_l x)Y(g_l y)d\mu(x)d\mu(y) \\
&= \sum_l \alpha_l \int_M \kappa(x,y)Y(g_k x)Y(g_k y)d\mu(x)d\mu(y) \\
&= \sum_l \alpha_l \int_M \kappa(x,y)Y(x)Y(y)d\mu(x)d\mu(y) \\
&= \int_M \kappa(x,y)Y(x)Y(y)d\mu(x)d\mu(y).
\end{aligned}$$

using the G-invariance of the measure on G, and (2.2.3). ■

Using essentially the same argument we can show the following.

Lemma 2.5. *Let us fix $n \in \mathbf{N}$, $1 \leq k \leq \nu_n$, and $X(x), \tilde{X}(y) \neq X(y) \in H_{nk}$. Then for any $Y(x), \tilde{Y}(y) \neq Y(y) \in H_{nk}$*

$$\int_M \kappa(x,y)Y(x)\tilde{Y}(y)d\mu(x)d\mu(y) = \int_M \kappa(x,y)X(x)\tilde{X}(y)d\mu(x)d\mu(y).$$

We now prove our theorem.

Theorem 2.1. *Every G-invariant kernel κ on a compact homogeneous manifold M has a spectral decomposition*

$$\kappa(x,y) = \sum_{n=0}^{\infty}\sum_{k=1}^{\nu_n} a_{nk}p_{nk}(x,y).$$

Proof: Let us fix $n \in \mathbf{N}$, $1 \leq k \leq \nu_n$. In order to prove (2.1) we first show that for $Y^j_{nk} \in H_{nk}$,

$$\int_M \kappa(x,y)Y^j_{nk}(x)dx \in H_{nk}.$$

Let $y = g_y\eta$ then

$$\begin{aligned} \int_M \kappa(x,g_y\eta)Y^j_{nk}(x)d\mu(x) &= \int_M \kappa(g_y^{-1}x,\eta)Y^j_{nk}(x)d\mu(x) \\ &= \int_M \kappa(x,\eta)Y^j_{nk}(g_yx)d\mu(x). \end{aligned}$$

This is a linear combination of polynomials in H_{nk}, which is G-invariant, and hence is in H_{nk}.

Hence, for any $1 \leq j \leq d_{nk}$,

$$\int_M \kappa(x,y)Y^j_{nk}(x)d\mu(x) = \sum_{l=1}^{d_{nk}} \gamma_l Y^l_{nk}(y).$$

If we multiply this equation by Y^l_{nk} and integrate this equation over M we see that

$$\int_M\int_M \kappa(x,y)Y^j_{nk}(x)Y^l_{nk}(y)d\mu(x)d\mu(y),$$

which equals α_{nk} say if $l = j$ from Lemma 2.4, and β_{nk} say, otherwise, from Lemma 2.5.

Therefore,

$$\int_M \kappa(x,y)Y^j_{nk}(x)d\mu(x) = \alpha_{nk}Y^j_{nk}(y) + \sum_{l\neq j}\beta_{nk}Y^l_{nk}(y). \tag{2.2.4}$$

Thus

$$\begin{aligned}
\int_M \kappa(x,y)p_{nk}(x,z)d\mu(x) &= \sum_{j=1}^{d_{nk}} \int_M \kappa(x,y)Y_{nk}^j(x)d\mu(x)Y_{nk}^j(z) \\
&= \alpha_{nk}\sum_{j=1}^{d_{nk}} Y_{nk}^j(y)Y_{nk}^j(z) + \beta_{nk}\sum_{j=1}^{d_{nk}}\sum_{l\neq j} Y_{nk}^l(y)Y_{nk}^j(z) \\
&= \alpha_{nk}p_{nk}(y,z) + \beta_{nk}\sum_{j=1}^{d_{nk}}\sum_{l\neq j} Y_{nk}^l(y)Y_{nk}^j(z) \qquad (2.2.5)
\end{aligned}$$

As a kernel of y and w the left hand side is G-invariant, since

$$\begin{aligned}
\int_M \kappa(x,gy)p_{nk}(x,gz)d\mu(x) &= \int_M \kappa(g^{-1}x,y)p_{nk}(g^{-1}x,z)d\mu(x) \\
&= \int_M \kappa(x,y)p_{nk}(x,z)d\mu(x).
\end{aligned}$$

Therefore the right hand side of (2.2.5) must also be G-invariant. The first term of the right hand side is G-invariant because it is the reproducing kernel p_{nk}. Thus,

$$\beta_{nk}\sum_{j=1}^{d_{nk}}\sum_{l\neq j} Y_{nk}^l(y)Y_{nk}^j(z)$$

is G-invariant. Set $z = y$ above. The sum

$$\beta_{nk}\sum_{j=1}^{d_{nk}}\sum_{l\neq j} Y_{nk}^l(gy)Y_{nk}^j(gy)$$

is constant for $g \in G$. If we integrate over G by the orthogonality of the basis we get

$$\beta_{nk}\sum_{j=1}^{d_{nk}}\sum_{l\neq j} Y_{nk}^l(gy)Y_{nk}^j(gy) = 0.$$

Now, we can choose a basis which is positive at some fixed point, hence

$$\sum_{j=1}^{d_{nk}}\sum_{l\neq j} Y_{nk}^l(gy)Y_{nk}^j(gy) \neq 0,$$

so that $\beta_{nk} = 0$.

Hence, from (2.2.4),

$$\int_M \kappa(x,y)p_{nk}(x,z)d\mu(x) = \alpha_{nk}p_{nk}(y,z),$$

which completes the proof. ■

2.3 Native Spaces

It was shown in Levesley and Ragozin [?] that for a G-invariant kernel with spectral decomposition

$$\kappa(x,y) = \sum_{n=0}^{\infty} \sum_{k=1}^{\nu_n} a_{nk}(\kappa) p_{nk}(x,y)$$

then we get positive definiteness if the coefficients a_{nk} are all positive. This is only a necessary condition. For the sphere there are a series of papers starting with Schoenberg [?], and developed in [?], and [?], which give more refined conditions for the strict positive definiteness of zonal kernels on spheres.

Using the positivity of the coefficients we can define the following inner product

$$\langle f,g\rangle_\kappa = \sum_{n=0}^{\infty} \sum_{k=1}^{\nu_n} (a_{nk}(\kappa))^{-1} (T_{nk}f, T_{nk}g).$$

The associated norm is denoted by $\|f\|_\kappa = \langle f,f\rangle_\kappa^{1/2}$

Definition 2.5. *Let*

$$\mathcal{N}_\kappa = \{f \in L_2(M) : \|f\|_\kappa < \infty\},$$

be the native space *associated with the kernel κ.*

Remark 2.1. *This definition follows that of the Definition 1.7, rather than the earlier definition. We can define the native space either way but we think this is more straightforward.*

The key property of the inner product which is given above is that the kernel κ becomes a reproducing kernel for the native space.

Proposition 2.2. *The kernel κ is the reproducing kernel for $\mathcal{N}_\kappa$. In other words, for $f \in \mathcal{N}_\kappa$,*

$$f(x) = \langle f, \kappa(\cdot, x)\rangle.$$

Proof: Let $f \in \mathcal{N}_\kappa$. Then

$$f = \sum_{n=0}^{\infty} \sum_{k=1}^{\nu_n} T_{nk}f,$$

and

$$\kappa(\cdot, x) = \sum_{n=0}^{\infty} \sum_{k=1}^{\nu_n} a_{nk} p_{nk}(\cdot, x).$$

Then, using the definition of inner product in $\mathcal{N}_\kappa$, we have

$$\begin{aligned}
\langle f, \kappa(\cdot, x)\rangle &= \sum_{n=0}^{\infty} \sum_{k=1}^{\nu_n} a_{nk}^{-1} \langle T_{nk} f, T_{nk}\kappa(\cdot, x)\rangle_\kappa \\
&= \sum_{n=0}^{\infty} \sum_{k=1}^{\nu_n} a_{nk}^{-1} \langle T_{nk} f, a_{nk} p_{nk}(\cdot, x)\rangle_\kappa \\
&= \sum_{n=0}^{\infty} \sum_{k=1}^{\nu_n} T_{nk} f(x) \\
&= f(x). \qquad \blacksquare
\end{aligned}$$

2.4 Special Functions

In this section we will look more carefully at the special functions which arise in our study. We will see why they arise, and give some of their fundamental properties. These can be found in [**?**]. In the first chapter we introduced the Gegenbauer polynomials. These are a subfamily of the Jacobi polynomials.

Definition 2.6. *The Jacobi polynomials $P_n^{(\alpha,\beta)}$ are the family of orthogonal polynomials which are orthogonal on $[-1, 1]$ with respect to the weight function:*

$$(1 - x)^\alpha (1 + x)^\beta.$$

As we observed in Theorem 1.3, these orthogonal polynomials satisfy a three term recurrence relation, which for the Gegenbauer polynomials is given by:

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

We also have the Rodriguez formula for the Gegenbauer polynomials:

$$(1 - x^2)^{\lambda-1/2} P_n^{(\lambda)}(x) = \gamma_n^\lambda \frac{d^n}{dx^n}(1 - x^2)^{n+\lambda-1/2}. \tag{2.4.1}$$

The Gegenbauer polynomials have a natural connection to the spheres. If we consider a zonal function, $\phi(xy) = \phi(\cos\theta)$, and integrating this over the sphere, thinking of y as the north pole and θ the polar angle, we obtain

$$\int_{S^{d-1}} \phi(xy)d\mu(x) \quad = \quad \omega_{d-2}\int_0^{\pi} \phi(\cos\theta)\sin^{d-2}\theta d\theta,$$

since the intersection of the hyperplane $xy = \cos\theta$ for fixed θ intersects the sphere in a sphere of one lower dimension, of radius $\sin\theta$, and hence of volume $\omega_{d-2}(\sin\theta)^{d-2}$ where ω is the volume of a hypersphere of one dimension less. Substituting $t = \cos\theta$ gives:

$$\int_{S^{d-1}} \phi(xy)d\mu(x) \quad = \quad \omega_{d-2}\int_{-1}^{1} \phi(t)(1-t^2)^{\frac{d-3}{2}}\omega dt.$$

Thus the weight function for the Gegenbauer polynomials appears naturally when integrating over a sphere.

A similar (but more complex) process performed over the two point homogeneous spaces leads to Jacobi polynomials of the form $P_n^{(d-2,0)}$ for complex projective spaces, and $P_n^{(d-2,1)}$ for quaternionic projective spaces.

The material here on special functions finds its application in the next chapter on Gegenbauer expansions for radial basis functions. The material on harmonic analysis is needed for the final chapter on compact homogeneous spaces.

CHAPTER 3

Gegenbauer Expansions

The integrals we compute below arise from the desire to approximate on the sphere using functions defined on the ambient Euclidean space. We consider two families of radial basis functions on $\mathbf{R}^d$: for fixed $y \in \mathbf{R}^d$, $\phi(x) = \|x - y\|^{2\alpha}$ and $\phi(x) = \|x-y\|^{2\alpha} \log \|x-y\|$, where $\|\cdot\|$ is the Euclidean norm in d–dimensional space. If we consider the restriction of these functions to the unit sphere we become interested in the following functions on S^{d-1}: for fixed $y \in S^{d-1}$, $\phi(x) = (1-xy)^\alpha$ and $\phi(x) = (1-xy)^\alpha \log(1-xy)$, where xy denotes the inner product of vectors $x, y \in S^{d-1}$.

The purpose of this chapter is to compute the Fourier–Gegenbauer series for the functions $\phi(t) = (1-t)^\alpha$ and $\phi(t) = (1-t)^\alpha \log(1-t)$; see Theorems 3.2 and 3.3. These expansions may be used to classify the space of functions which can be approximated using the above Euclidean radial basis functions for approximation on S^{d-1}. These integrals and others have been computed, using different techniques, by Baxter and Hubbert [**?**]. The results presented here have appeared in [**?**].

Let $P_n^\lambda(x)$ be the Gegenbauer polynomial of degree n, orthogonal on $[-1,1]$ with respect to the weight $(1-x^2)^{\lambda-1/2}$. These polynomials are normalised so that

$$\begin{aligned} h_n^\lambda &:= \int_{-1}^{1} P_n^\lambda(x)(1-x^2)^{\lambda-1/2}\,dx \\ &= \frac{\pi 2^{1-2\lambda}\Gamma(n+2\lambda)}{n!(n+\lambda)(\Gamma(\lambda))^2}. \end{aligned} \tag{3.0.1}$$

3.1 Results

Theorem 3.1. *For $\alpha > 0$ and $\lambda > -1/2$, let*

$$I_n^{\alpha,\lambda} = \int_{-1}^{1} (1-x^2)^\alpha P_n^{(\lambda)}(x)\,dx.$$

If n is even then

$$I_n^{\alpha,\lambda} = \frac{\Gamma(\lambda+n/2)2^{-2\alpha}\pi\Gamma(2\alpha+1)\Gamma(\lambda-\alpha-1/2+n/2)}{(n/2)!\Gamma(\lambda)\Gamma(\lambda-\alpha-1/2)\Gamma(\alpha+n/2+3/2)\Gamma(\alpha+1/2)}.$$

However, if n is odd then $I_n^{\alpha,\lambda} = 0$.

Proof: The case of n odd is obvious.

From Section 2.3, we have the following three term recurrence relation for orthogonal polynomials in its form for the Gegenbauer polynomials:

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

Substituting this recurrence relation into the above integral gives:

$$\begin{aligned} I_n^{\alpha,\lambda} &= \int_{-1}^{1} (1-x^2)^\alpha \frac{2(n+\lambda-1)xP_{n-1}^{(\lambda)}(x)}{n}\,dx \\ &\qquad - \int_{-1}^{1} (1-x^2)^\alpha \frac{(n+2\lambda-2)}{n} P_{n-2}^{(\lambda)}(x)\,dx \\ &= A - B, \end{aligned}$$

with the obvious definitions of A and B. Then, dealing with A first and using integration by parts,

$$A = \int_{-1}^{1} \frac{(1-x^2)^{\alpha+1}(n+\lambda-1)2\lambda P_{n-2}^{(\lambda+1)}\,dx}{(\alpha+1)n},$$

where we have used the fact that

$$\frac{d}{dx}P_{n-1}(\lambda)(x) = 2\lambda P_{n-2}^{(\lambda+1)}(x);$$

see [? , Page 82]. Thus, putting this back into the equation gives a new recurrence relation, this time between integrals:

$$I_n^{\alpha,\lambda} = \frac{2(n+\lambda-1)\lambda}{n(\alpha+1)}I_{n-2}^{\lambda+1} - \frac{(n+2\lambda-2)}{n}I_{n-2}^{\alpha,\lambda}.$$

We now proceed by induction. Assume the truth of the induction hypothesis for the case $n-2$ for all λ and for all α. Under the induction hypothesis we have:

$$I_{n-2}^{\alpha,\lambda} = \frac{\Gamma(\lambda+n/2-1)2^{-2\alpha}\pi\Gamma(2\alpha+1)\Gamma(\lambda-\alpha-3/2+n/2)}{(n/2-1)!\Gamma(\lambda)\Gamma(\lambda-\alpha-1/2)\Gamma(\alpha+n/2+1/2)\Gamma(\alpha+1/2)},$$

and

$$I_{n-2}^{\alpha+1,\lambda+1} = \frac{\Gamma(\lambda+n/2-)2^{-2\alpha-2}\pi\Gamma(2\alpha+3)\Gamma(\lambda-\alpha-3/2+n/2)}{(n/2-1)!\Gamma(\lambda+1)\Gamma(\lambda-\alpha-1/2)\Gamma(\alpha+n/2+3/2)\Gamma(\alpha+3/2)}.$$

Using the recurrence relation just derived one obtains, by substituting for $I_{n-2}^{\alpha+1,\lambda+1}$ and $I_{n-2}^{\alpha,\lambda}$, we obtain

$$\begin{aligned} I_n^{\alpha,\lambda} &= \frac{2(n+\lambda-1)\lambda\Gamma(\lambda+n/2)2^{-2\alpha}\pi}{4n(\alpha+1)\Gamma(\lambda+1)\Gamma(\alpha+3/2)} \\ &\quad \times\frac{(\Gamma(2\alpha+3)\Gamma(\lambda-\alpha-3/2+n/2)}{\Gamma(\lambda-\alpha-1/2)\Gamma(\alpha+n/2+3/2)(n/2-1)!} \\ &\quad -\frac{(n+2\lambda-2)\Gamma(\lambda+n/2-1)2^{-2\alpha}\pi\Gamma(2\alpha+1)\Gamma(\lambda-\alpha-3/2+n/2)}{(n/2-1)!n\Gamma(\lambda-\alpha-1/2)\Gamma(\alpha+n/2+1/2)\Gamma(\lambda)\Gamma(\alpha+1)\Gamma(\alpha+1/2)}. \end{aligned}$$

Using $\Gamma(x+1) = x\Gamma(x)$, applying this to $\Gamma(2\alpha+3)$ gives $\Gamma(2\alpha+3) = (2\alpha+1)(2\alpha+2)\Gamma(2\alpha+1)$ and substituting this into the above gives:

$$\begin{aligned} I_n^{\alpha,\lambda} &= \frac{2(n+\lambda-1)\lambda\Gamma(\lambda+n/2)2^{-2\alpha}\pi}{4n(\alpha+1)\Gamma(\lambda+1)\Gamma(\alpha+3/2)} \\ &\quad \times\frac{(2\alpha+2)(2\alpha+1)\Gamma(2\alpha+1)\Gamma(\lambda-\alpha-3/2+n/2)}{\Gamma(\lambda-\alpha-1/2)\Gamma(\alpha+n/2+3/2)(n/2-1)!} \\ &\quad -\frac{(n+2\lambda-2)\Gamma(\lambda+n/2-1)2^{-2\alpha}\pi\Gamma(2\alpha+1)\Gamma(\lambda-\alpha-3/2+n/2)}{(n/2-1)!n\Gamma(\lambda-\alpha-1/2)\Gamma(\alpha+n/2+1/2)\Gamma(\lambda)\Gamma(\alpha+1)\Gamma(\alpha+1/2)} \\ &= \frac{\Gamma(\lambda+n/2-1)2^{-2\alpha}\pi\Gamma(2\alpha+1)\Gamma(\lambda-\alpha-3/2+n/2)}{(n/2-1)!\Gamma(\lambda)\Gamma(\alpha+1/2)\Gamma(\lambda-\alpha-1/2)\Gamma(\alpha+n/2+1/2)} \\ &\quad \times\left[\frac{(n+\lambda-1)2\lambda(\lambda+n/2-1)(2\alpha+2)(2\alpha+1)}{4n(\alpha+1)\lambda(\alpha+1/2)(\alpha+n/2+1/2)} - \frac{(n+2\lambda-2)}{n}\right], \end{aligned}$$

by substituting in $\Gamma(\lambda+1) = \lambda\Gamma(\lambda)$, $\Gamma(\alpha+3/2) = (\alpha+1/2)\Gamma(\alpha+1/2)$ and $\Gamma(\alpha+n/2+3/2) = (\alpha+n/2+1/2)\Gamma(\alpha+n/2+1/2)$. However,

$$\begin{aligned}
\frac{(n+2\lambda-2)}{n}\left(\frac{2(n+\lambda-1)}{2\alpha+n+1}-1\right) &= \frac{(n+2\lambda-2)}{n}\left(\frac{2(n+\lambda-1)-(2\alpha+n+1)}{2\alpha+n+1}\right)\\
&= \frac{(n+2\lambda-2)}{n}\left(\frac{2n+2\lambda-2-2\alpha-n-1}{2\alpha+n+1}\right)\\
&= \frac{n+2\lambda-2}{n(2\alpha+n+1)}(2\lambda-2\alpha+n-3)\\
&= \frac{n+2\lambda-2}{n/2(2\alpha+n+1)}(\lambda-\alpha+n/2-3/2).
\end{aligned}$$

Therefore

$$\begin{aligned}
I_n^{\alpha,\lambda} &= \frac{\Gamma(\lambda+n/2-1)2^{-2\alpha}\pi\Gamma(2\alpha+1)\Gamma(\lambda-\alpha-3/2+n/2)}{(n/2-1)!\Gamma(\lambda)\Gamma(\alpha+1/2)\Gamma(\lambda-\alpha-1/2)\Gamma(\alpha+n/2+1/2)}\\
&\qquad\times\frac{(\lambda-\alpha+n/2-3/2)(\lambda+n/2-1)}{n/2(\alpha+n/2-1)}\\
&= \frac{\Gamma(\lambda+n/2)2^{-2\alpha}\pi\Gamma(2\alpha+1)\Gamma(\lambda-\alpha-1/2+n/2)}{(n/2)!\Gamma(\lambda)\Gamma(\alpha+1/2)\Gamma(\lambda-\alpha-1/2)\Gamma(\alpha+n/2+3/2)},
\end{aligned}$$

as required.

We complete the proof by noting that, from (3.0.1) with $\lambda = \alpha+1/2$ and $n=0$,

$$I_0^{\alpha,\lambda} = \frac{2^{-2\alpha}\pi\Gamma(2\alpha+1)}{(\Gamma(\alpha+1/2))^2(\alpha+1/2)}$$

and by substituting n=0 and $\lambda = \alpha+1/2$ into our formula gives

$$I_0^{\alpha,\lambda} = \frac{\Gamma(\alpha+1/2)2^{-2\alpha}\pi\Gamma(2\alpha+1)\Gamma(0)}{0!(\Gamma(\alpha+1/2))^2\Gamma(0)\Gamma(\alpha+3/2)}.$$

Cancelling and using $\Gamma(x+1) = x\Gamma(x)$ to give $\Gamma(\alpha+3/2) = (\alpha+1/2)\Gamma(\alpha+1/2)$ gives the result. ∎

Theorem 3.2. *Let $\alpha > 0$ and $\lambda > -1/2$. Then*

$$\begin{aligned}
J_n^{\alpha,\lambda} &:= \int_{-1}^{1}(1-x)^{\alpha}(1-x^2)^{\lambda-1/2}P_n^{(\lambda)}(x)\\
&= \frac{(-1)^n(2\pi)^{1/2}2^{\alpha+1/2}\Gamma(\alpha+\lambda+1/2)\Gamma(2\lambda+n)\Gamma(\alpha+1)}{n!\Gamma(\alpha+2\lambda+n+1)\Gamma(\alpha-n+1)\Gamma(\lambda)}.
\end{aligned}$$

Proof: Using integration by parts

$$J_n^{\alpha,\lambda} = \int_{-1}^{1}\frac{(1-x)^{\alpha+1}}{\alpha+1}\frac{d}{dx}\left((1-x^2)^{\lambda-1/2}P_n^{(\lambda)}(x)\right)dx.$$

Now, using Rodriguez' formula (3.1.1),

$$(1-x^2)^{\lambda-1/2}P_n^{(\lambda)}(x) \quad = \quad \gamma_n^\lambda \frac{d^n}{dx^n}(1-x^2)^{n+\lambda-1/2}, \tag{3.1.1}$$

where

$$\gamma_n^\lambda \quad = \quad \frac{(-2)^n\Gamma(n+\lambda)\Gamma(n+2\lambda)}{n!\Gamma(\lambda)\Gamma(2n+2\lambda)}. \tag{3.1.2}$$

Thus differentiating Rodriguez's formula through once gives

$$\begin{aligned} \frac{d}{dx}(1-x^2)^{\lambda-1/2}P_n^{(\lambda)}(x) \quad &= \quad \gamma_n^\lambda \frac{d^{n+1}}{dx^{n+1}}(1-x^2)^{n+\lambda-1/2} \\ &= \quad \gamma_n^\lambda \frac{d^{n+1}}{dx^{n+1}}(1-x^2)^{n+1+(\lambda-1)-1/2}. \end{aligned}$$

Using (3.1.1) again we have upon substituting $n+1$ for n and $\lambda-1$ for λ

$$\frac{d^{n+1}}{dx^{n+1}}(1-x^2)^{n+\lambda-1/2} = \frac{(1-x^2)^{\lambda-3/2}P_{n+1}^{(\lambda-1)}}{\gamma_{n+1}^{\lambda-1}}.$$

Thus substituting this back into the preceeding equation:

$$\frac{d}{dx}(1-x^2)^{\lambda-1/2}P_n^{(\lambda)}(x) \quad = \quad \frac{\gamma_n^\lambda}{\gamma_{n+1}^{\lambda-1}}(1-x^2)^{\lambda-3/2}P_{n+1}^{(\lambda-1)}(x).$$

Putting this back into the integral gives

$$J_n^{\alpha,\lambda} = \frac{\gamma_n^\lambda}{\gamma_{n+1}^{\lambda-1}} \int_{-1}^{1} \frac{(1-x)^{\alpha+1}}{\alpha+1}(1-x^2)^{\lambda-3/2}P_{n+1}^{(\lambda-1)}(x)dx.$$

Therefore

$$J_n^{\alpha,\lambda} = \frac{\gamma_n^\lambda J_{n+1}^{\alpha+1,\lambda-1}}{\gamma_{n-1}^{\lambda-1}(\alpha+1)},$$

so that rearrangement gives

$$J_n^{\alpha,\lambda} = \frac{\alpha\gamma_n^\lambda J_{n-1}^{\alpha-1,\lambda+1}}{\gamma_{n-1}^{\lambda+1}}.$$

Now, from (3.1.2),

$$\gamma_{n-1}^{\lambda+1} = \frac{(-2)^{n-1}\Gamma(n+\lambda)\Gamma(n+2\lambda+1)}{(n-1)!\Gamma(\lambda+1)\Gamma(2n+2\lambda)}.$$

Thus

$$\frac{\gamma_n^\lambda}{\gamma_{n-1}^{\lambda+1}} = \frac{(-2)^n\Gamma(n+\lambda)\Gamma(n+2\lambda)}{n!\Gamma(\lambda)\Gamma(2n+2\lambda)} \times \frac{(n-1)!\Gamma(\lambda+1)\Gamma(2n+2\lambda)}{(-2)^{n-1}\Gamma(n+\lambda)\Gamma(n+2\lambda+1)},$$

giving, upon using $\Gamma(x+1) = x\Gamma(x)$,

$$\frac{\gamma_n^{\lambda}}{\gamma_{n-1}^{\lambda+1}} = \frac{-2\lambda\alpha}{n(n+2\lambda)},$$

with the result that

$$J_n^{\alpha,\lambda} = \frac{\alpha(-2\lambda)}{n(n+2\lambda)} J_{n-1}^{\alpha-1,\lambda+1}.$$

We can now proceed by induction. Assume the result is true for $n-1$, and all $\alpha > 0$ and $\lambda > -1/2$:

$$J_{n-1}^{\alpha-1,\lambda+1} = \frac{(-1)^{n-1} 2^{\alpha-1+2\lambda+2n} \Gamma(\alpha+\lambda+1/2)\Gamma(\lambda+n+1/2)}{(n-1)!\Gamma(2\lambda+2n)\Gamma(\alpha+2\lambda+n+1)} \times \frac{\Gamma(2\lambda+n+1)\Gamma(\alpha)\Gamma(\lambda+n-1)}{\Gamma(\alpha-n+1)\Gamma(\lambda-1)}.$$

Substituting this into the recurrence relation gives

$$\begin{aligned} J_n^{\alpha,\lambda} &= \frac{\alpha(-2\lambda)(-1)^{n-1} 2^{\alpha+2\lambda+2n-1}}{n(n+2\lambda)(n-1)!\Gamma(2\lambda+2n)} \\ &\quad \times \frac{\Gamma(\alpha+\lambda+1/2)\Gamma(\lambda+n+1/2)\Gamma(2\lambda+n+1)\Gamma(\alpha)\Gamma(\lambda+n)}{\Gamma(\alpha+2\lambda+n+1)\Gamma(\alpha-n+1)\Gamma(\lambda+1)} \\ &= \frac{(-1)^n 2^{\alpha+2\lambda+2n} \Gamma(\alpha+\lambda+1/2)\Gamma(\lambda+n+1/2)\Gamma(2\lambda+n)}{n!\Gamma(2\lambda+2n)\Gamma(\alpha+2\lambda+n+1)\Gamma(\alpha-n+1)\Gamma(\lambda)} \\ &\quad \times \Gamma(\alpha+1)\Gamma(\lambda+n). \end{aligned}$$

Using the duplication formula (see [**?** , Chapter 6])

$$\Gamma(2z) = (2\pi)^{-1/2} 2^{2z-1/2} \Gamma(z)\Gamma(z+1/2),$$

we can simplify the above equation (also putting $\lambda = (d-2)/2$) to get

$$J_n^{\alpha,\lambda} = \frac{(-1)^n (2\pi)^{1/2} 2^{\alpha+1/2} \Gamma(\alpha+d/2-1/2)\Gamma(d-2+n)\Gamma(\alpha+1)}{n!\Gamma(\alpha+d+n-1)\Gamma(\alpha-n+1)\Gamma(d/2-1)}.$$

We conclude the induction(putting λ back again) by noting from [**?** , Page 68], that for all $\alpha > 0$ and $\lambda > -1/2$:

$$J_0^{\alpha,\lambda} = \frac{2^{\alpha+2\lambda}\Gamma(\alpha+\lambda+1/2)\Gamma(\lambda+1/2)}{\Gamma(\alpha+2\lambda+1)}.$$

Putting $n = 0$ into our formula gives

$$J_0^{\alpha,\lambda} = \frac{2^{\alpha+2\lambda}\Gamma(\alpha+\lambda+1/2)\Gamma(\lambda+1/2)\Gamma(2\lambda)\Gamma(\alpha+1)\Gamma(\lambda)}{0!\Gamma(2\lambda)\Gamma(\alpha+2\lambda+1)\Gamma(\alpha+1)\Gamma(\lambda)}$$

and cancellation gives the result. ■

In Chapter 4.1 we will need to know the decay rates for these Gegenbauer coefficients. We can do this using the asymptotic formula for the Gamma function (see [**?** , Section 6.1])

$$\Gamma(az+b) \sim (2\pi)^{1/2} e^{-az} (az)^{az+b-1/2}.$$

Writing

$$C_{\alpha,\lambda} = (2\pi)^{1/2} 2^{\alpha+1/2} \frac{\Gamma(\alpha+\lambda+1/2)\Gamma(\alpha+1)}{\Gamma(\lambda)},$$

we have

$$\begin{aligned} J_n^{\alpha,\lambda} &= (-1)^n C_{\alpha,\lambda} \Gamma(2\lambda+n) n! \Gamma(\alpha+2\lambda+n+1)\Gamma(\alpha-n+1) \\ &\sim C_{\alpha,\lambda} \frac{(2\pi)^{1/2} e^{-n} (n)^{2\lambda+n-1/2}}{(2\pi)^{1/2} e^{-n} (n)^{n+1/2} (2\pi)^{1/2} e^{-n} (n)^{\alpha+2\lambda+n+1/2} (2\pi)^{1/2} e^{n} (-n)^{\alpha-n+1/2}} \\ &= \mathcal{O}(n^{-2\alpha-2}). \end{aligned} \tag{3.1.3}$$

Let

$$\Psi(x) = \frac{d}{dx} \log \Gamma(x) = \frac{\Gamma'(x)}{\Gamma(x)}$$

be the digamma function (see [**?** , Chapter 6]).

Theorem 3.3. *Let* $\alpha > 0$, $\lambda > -1/2$ *and*

$$K_n^{\alpha,\lambda} = \int_{-1}^{1} (1-x)^\alpha \log(1-x)(1-x^2)^{\lambda-1/2} P_n^{(\lambda)}(x)\, dx.$$

Then, if $\alpha \in \mathbf{N}$ *and* $n > \alpha$,

$$\begin{aligned} &K_n^{\alpha,\lambda} \\ &= \frac{(-1)^\alpha 2^{\alpha+2\lambda+2n}(n-\alpha-1)!\Gamma(\alpha+\lambda+1/2)\Gamma(\lambda+n+1/2)\Gamma(2\lambda+n)\Gamma(\alpha+1)\Gamma(\lambda+n)}{n!\Gamma(2\lambda+2n)\Gamma(\alpha+2\lambda+n+1)\Gamma(\lambda)}. \end{aligned}$$

Otherwise,

$$\begin{aligned} &K_n^{\alpha,\lambda} \\ &= J_n^{\alpha,\lambda}(\log 2 + \Psi(\alpha+\lambda+1/2) + \Psi(\alpha+1) - \Psi(\alpha+2\lambda+n) - \Psi(\alpha-n+1)). \end{aligned}$$

Proof: The key observation in this proof is that

$$\frac{d}{d\alpha}(1-x)^\alpha = (1-x)^\alpha \log(1-x).$$

Thus,

$$K_n^{\alpha,\lambda} = \frac{d}{d\alpha} J_n^{\alpha,\lambda}.$$

If we write

$$A_n^\lambda = \frac{(-1)^n 2^{2\lambda+2}\Gamma(\lambda+n+1/2)\Gamma(2\lambda+n)\Gamma(\lambda+n)}{n!\Gamma(2\lambda+2n)\Gamma(\lambda)}$$

then

$$J_n^{\alpha,\lambda} = A_n^\lambda \frac{2^\alpha\Gamma(\alpha+\lambda+1/2)\Gamma(\alpha+1)}{\Gamma(\alpha+2\lambda+n)\Gamma(\alpha-n+1)}.$$

Thus, using the facts that $\Gamma'(x) = \Gamma(x)\Psi(x)$ and that

$$\frac{d}{dx}\frac{1}{\Gamma(x)} = -\frac{\Psi(x)}{\Gamma(x)},$$

we see that

$$\begin{aligned}
K_n^{\alpha,\lambda} &= A_n^\lambda \frac{\partial}{\partial\alpha}\frac{2^\alpha\Gamma(\alpha+\lambda+1/2)\Gamma(\alpha+1)}{\Gamma(\alpha+2\lambda+n)\Gamma(\alpha-n+1)} \\
&= A_n^\lambda \frac{2^\alpha\Gamma(\alpha+\lambda+1/2)\Gamma(\alpha+1)}{\Gamma(\alpha+2\lambda+n)\Gamma(\alpha-n+1)} \\
&\quad (\log 2 + \Psi(\alpha+\lambda+1/2) + \Psi(\alpha+1) - \Psi(\alpha+2\lambda+n) - \Psi(\alpha-n+1))),
\end{aligned}$$

and the result follows. We justify the last line above (a simple conseqence of a compound product rule) in what follows. Let

$$u = 2^{\alpha+2\lambda+2n}\Gamma(\alpha+\lambda+1/2)\Gamma(\alpha+1)$$

and let

$$v = \Gamma(\alpha+2\lambda+n)\Gamma(\alpha-n+1).$$

Then let

$$u_1 = 2^{\alpha+2\lambda+2n}\Gamma(\alpha+\lambda+1/2)$$

and

$$u_2 = \Gamma(\alpha+1).$$

Then

$$\frac{\partial u_1}{\partial\alpha} = \Gamma(\alpha+\lambda+1/2)\log 2.2^{\alpha+2\lambda+2n} + \Gamma(\alpha+\lambda+1/2)2^{\alpha+2\lambda+2n}\Psi(\alpha+\lambda+1/2)$$

and

$$\frac{\partial u_2}{\partial\alpha} = \Gamma(\alpha+1)\Psi(\alpha+1).$$

So since

$$\frac{\partial u}{\partial \alpha} = \frac{\partial u_1}{\partial \alpha} u_2 + \frac{\partial u_2}{\partial \alpha} u_1$$

$$= 2^{\alpha+2\lambda+2n}\Gamma(\alpha+1)\Gamma(\alpha+\lambda+1/2)(\log 2 + \Psi(\alpha+1) + \Psi(\alpha+\lambda+1/2).$$

$$\frac{\partial v}{\partial \alpha} = \Gamma(\alpha+2\lambda+n)\Gamma(\alpha-n+1)(\Psi(\alpha+2\lambda+n) + \Psi(\alpha-n+1).$$

Now

$$\partial\left(\frac{u}{v}\right) = \frac{v\frac{\partial u}{\partial \alpha} - u\frac{\partial v}{\partial \alpha}}{v^2}$$

$$= \frac{\Gamma(\alpha+2\lambda+n)\Gamma(\alpha-n+1)\Gamma(\alpha+1)\Gamma(\alpha+\lambda+1/2).2^{\alpha+2\lambda+2n}}{(\Gamma(\alpha+2\lambda+n))^2(\Gamma(\alpha-n+1))^2}$$

$$\times(\log 2 + \Psi(\alpha+\lambda+1/2) + \Psi(\alpha+1))$$

$$\frac{-2^{\alpha+2\lambda+2n}\Gamma(\alpha+\lambda+1/2)\Gamma(\alpha+1)\Gamma(\alpha+2\lambda+n)\Gamma(\alpha-n+1)}{(\Gamma(\alpha+2\lambda+n))^2(\Gamma(\alpha-n+1))^2}$$

$$\times(\Psi(\alpha+2\lambda+n) + \Psi(\alpha-n+1))$$

and upon cancellation the result follows.

Now, when $\alpha \in \mathbf{N}$ and $n > \alpha$, $\Gamma(\alpha-n+1)$ is undefined, as is $\Psi(\alpha-n+1)$. Thus in this case we need to compute the limit

$$\lim_{x\to -k} \frac{\Psi(x)}{\Gamma(x)}$$

for $k \in \mathbf{N}$. Since Γ possesses a simple pole at $-k$ with residue $(-1)^k/k!$, (see [? , Page 255]), Ψ also possesses a simple pole with residue -1. Thus, for $k \in \mathbf{N}$,

$$\lim_{x\to -k} \frac{\Psi(x)}{\Gamma(x)} = (-1)^{k+1}/k!.$$

Therefore, for $\alpha \in \mathbf{N}$ and $n \geq \alpha$,

$$K_n^{\alpha,\lambda} = (-1)^{n-\alpha} A_n^{\lambda} \frac{2^{\alpha}(n-\alpha-1)!\Gamma(\alpha+\lambda+1/2)\Gamma(\alpha+1)}{\Gamma(\alpha+2\lambda+n)} \qquad \blacksquare$$

Using the duplication formula again and putting $\lambda = (d-2)/2$ gives the following simplification

$$\frac{(-1)^{\alpha}(2\pi)^{1/2}(2)^{\alpha+1/2}\Gamma(\alpha+d/2-1/2)\Gamma(d+n-2)\Gamma(\alpha+1)}{n!\Gamma(\alpha+d+n-1)(n-\alpha-1)!\Gamma(d/2-1)}$$

We calculated the decay rate for this coefficient, using the result (3.1). For $\alpha \in \mathbf{N}$ and $n > \alpha$, which is most important for our applications we again get

$$K_n^{\alpha,\lambda} = \mathcal{O}(n^{-2\alpha-2}). \tag{3.1.4}$$

3.2 Concluding remarks

In this chapter we have seen that the Gegenbauer coefficients for the standard radial basis functions on the spheres can be computed and confirm the degree of positive definiteness of these kernels on the sphere. It was conjectured by Levesley and Hubbert [**?**] that the decay rate of the Gegenbauer coefficients could be inferred directly from the rate of decay of the Fourier transform of the associated RBF in the ambient Euclidean space. This was confirmed independently at a similar time by Narcowich et. al [**?**] and zu Castell [**?**]. We would expect to be able to compute Jacobi expansions for the same RBFs and hence compute the appropriate native spaces for approximation by RBFs on other two point homogeneous spaces.

CHAPTER 4

Interpolation on Compact Homogeneous Manifolds

In this chapter we will be concerned with the approximation on a compact homogeneous manifold M, at a finite point set $Y \in M$ using a G-invariant kernel. As discussed in the introduction, such kernels can be seen to be generalisations of radial basis functions and zonal functions on the sphere. We will follow the standard variational approach for interpolation; the contents of this chapter can be found in more condensed form in [**?**]. In order to prove our result we will use the notion of a norming set, which was introduced in [**?**], and which we discuss in Section 4.3. To use these results we will need to use a Bernstein-type inequality, and we will discuss such inequalities in Section 4.2.

4.1 Variational theory

In Chapter 2 we introduced native spaces for our spaces M. We will assume that we have a strictly positive definite kernel κ with spectral decomposition

$$\kappa(x,y) = \sum_{n=0}^{\infty} \sum_{k=1}^{\nu_n} a_{nk}(\kappa) p_{nk}(x,y),$$

where p_{nk} is the reproducing kernel for H_{nk}, as decsribed in Section 2.2. We are interested in approximating a function $f \in \mathcal{N}_\kappa$.

We form an interpolant s_Y in the standard manner

$$s_Y(x) = \sum_{y \in Y} c_y \kappa(y, x)$$

where the coefficients $c_y, y \in Y$, are determined by the interpolation conditions:

$$s_Y(y) = f(y), \; y \in X.$$

If the kernel κ is strictly positive definite then the matrix system is invertible regardless of the configuration of the point set Y.

The error $|f(x) - s_Y(x)|$ is examined as the set Y becomes dense in spherical subsets $S_\eta = \{y \in M | d(y, p) \leq \eta\}$ (which may in the case of global error analysis be M itself). The measure of density chosen is the mesh norm.

Definition 4.1. *The mesh norm is defined to be*

$$h(Y) = \max_{y \in S_\eta} \min_{x \in Y} d(x, y).$$

We are interested in estimating decay rates for the error as $h(Y) \to 0$. When $\eta < \max_{x,y \in M} d(x, y)$ these become local error estimates.

Using the reproducing kernel property of Proposition 2.2 we can measure the interpolation error for $x \in M$ by

$$\begin{aligned} |f(x) - s_Y(x)| &\leq \langle f - s_Y, \kappa(\cdot, x) \rangle_\kappa \\ &= \langle f - s_Y, \kappa(\cdot, x) + \sum_{y \in Y} \alpha_y \kappa(\cdot, y) \rangle_\kappa \\ &\leq \|f - s_Y\|_\kappa \|\kappa(\cdot, x) + \sum_{y \in Y} \alpha_y \kappa(\cdot, y)\|_\kappa, \end{aligned} \tag{4.1.1}$$

where $\alpha_y, y \in Y$ are arbitrary, since, because Y is a set of interpolation points,

$$\begin{aligned} |f(y) - s_Y(y)| &= \langle f - s_Y, \kappa(\cdot, y) \rangle_\kappa \\ &= 0, \quad y \in Y. \end{aligned}$$

To simplify (4.1.1) we prove two auxiliary lemmas. The first is a consequence of the well-known norm minimisation property for variational splines. The second is used to provide a more convenient expression for the second quantity on the right hand side.

Lemma 4.1.

$$\|f\|_\kappa^2 = \|f - s_Y\|_\kappa^2 + \|s_Y\|_\kappa^2$$

Proof: Consider

$$\langle f - s_Y, s_Y\rangle_\kappa.$$

Now,

$$s_Y(x) = \sum_{y\in Y} \alpha_y \kappa(x, y),$$

so that

$$\begin{aligned} \langle f - s_Y, s_Y\rangle_\kappa &= \langle f - s_Y, \sum_{y\in Y} \alpha_y \kappa(x, y)\rangle_\kappa \\ &= 0, \end{aligned}$$

since we are evaluating $f - s_Y$ at the interpolation points, where f and s_Y are equal. Using Pythagoras' Theorem we get the stated result. ∎

Corollary 4.1. *For all $f \in \mathcal{N}_\kappa$,*

$$\|f - s_Y\|_\kappa \leq \|f\|_\kappa.$$

Lemma 4.2. *Let $Z \subset M$ be finite, and $\lambda \in Z^*$. Then*

$$\|\lambda_w \kappa(\cdot, w)\|_\kappa \leq \sup_{\|g\|_\kappa=1} |\lambda g|.$$

Proof: For any $g \in N_\kappa$, with $\|g\|_\kappa = 1$,

$$\begin{aligned} |\lambda g| &= \langle g, \lambda_w \kappa(\cdot, w)\rangle_\kappa \\ &\leq \|g\|_\kappa \|\lambda_w \kappa(\cdot, w)\|_\kappa \\ &= \|\lambda_w \kappa(\cdot, w)\|_\kappa. \end{aligned}$$

If we set

$$g = \frac{\lambda_w \kappa(\cdot, w)}{\|\lambda_w \kappa(\cdot, w)\|_\kappa},$$

we get

$$|\lambda g| = \|\lambda_w \kappa(\cdot, w)\|_\kappa.$$

Hence

$$\|\lambda_w \kappa(\cdot, w)\|_\kappa \leq \sup_{\|g\|_\kappa = 1} |\lambda g|. \qquad \blacksquare$$

Using the previous two lemmas and (4.1.1) we get

$$\begin{aligned}
|f(x) - s_Y(x)| &\leq \langle f - s_Y, \kappa(\cdot, x)\rangle_\kappa \\
&= \langle f - s_Y, \kappa(\cdot, x) + \sum_{y \in Y} \alpha_y \kappa(\cdot, y)\rangle_\kappa \\
&\leq \|f - s_Y\|_\kappa \inf_{\{\alpha_y : y \in Y\}} \sup_{\|g\|_\kappa = 1} \left| g(x) - \sum_{y \in Y} \alpha_y g(y) \right| \qquad (4.1.2) \\
&\leq \|f\|_\kappa \inf_{\{\alpha_y : y \in Y\}} \sup_{\|g\|_\kappa = 1} \left| g(x) - \sum_{y \in Y} \alpha_y g(y) \right|. \qquad (4.1.3)
\end{aligned}$$

In order to progress we need to be able to produce a functional which annihilates polynomials of some degree, which has controllable size. This is the objective of the next section.

4.2 Bernstein and Markov inequalities

On the circle the famous standard Bernstein inequality is

$$\|t'_N\|_\infty \leq N \|t_N\|_\infty,$$

for any trigonometric polynomial of degree N. Similarly, for an algebraic polynomial on an open interval we have the Markov inequality

$$\|t'_N\|_{\infty,[a,b]} \leq \frac{2}{(b-a)} N^2 \|t_n\|_{\infty,[a,b]}.$$

Bos et. al [**?**] give a Bernstein inequality on algebraic manifolds. Kroo [**?**] has proved a Bernstein and a Markov-type inequality for non-symmetric convex bodies (where a convex body in $\mathbf{R}^n$ is a convex, compact set with non-empty interior).

To define directional derivatives on our homogeneous spaces we need to use the exponential of an element of the Lie algebra $L(G)$.

Definition 4.2. *Let $L(G)$ be the Lie algebra of a group G acting on a manifold M. $L(G)$ consists of those $l \times l$ skew-symmetric matrices D such that* $\exp tD \in G$ *for all $t \in \mathbf{R}$. Each $D \in L(G)$ acts as a differential operator on $C^1(M)$ by*

$$Df(m) = (d/dt)f(\exp tDm)|_{t=0}, \quad f \in C^1(M).$$

4.2.1 Two point homogeneous spaces

For a compact two point homogeneous space each of the geodesics is closed (see Helgason [**?**]) so that the restriction any polynomial of degree n to a (closed) minimal geodesic joining two nearby points will be a trigonometric polynomial of degree n. Hence, for such spaces it is clear that for differentiation along any geodesic (each geodesic through a point defines a direction) we have

Lemma 4.3. *Let M be a compact 2-point homogeneous space. Then for any polynomial p_N of degree N and any element $D \in L(G)$ of unit length, we have the Bernstein inequality*

$$\max_{t\in\mathbf{R}} \left| \frac{d}{dt} p_N(\exp(tD)) \right| \le N \max_{t\in\mathbf{R}} |p_N(\exp(tD))|.$$

For a restricted interval of length less that the diameter of the manifold, we have the Markov inequality

$$\max_{t\in[0,\tau]} \left| \frac{d}{dt} p_N(\exp(tD)) \right| \le \frac{2}{\tau} N^2 \max_{t\in[0,\tau]} |p_N(\exp(tD))|.$$

4.2.2 General compact homogeneous spaces

Here we need to see the following result due to Ragozin [**?**] to give

Proposition 4.1. *Let M be a compact homogeneous space. Then for any polynomial p_N of degree N and any element $D \in L(G)$ of unit length, we have the Bernstein inequality*

$$\max_{t\in\mathbf{R}} \left| \frac{d}{dt} p_N(\exp(tD)) \right| \le R_1 N \max_{t\in\mathbf{R}} |p_N(\exp(tD))|.$$

, for some fixed positive constant R_1. For a restricted interval of length less that the diameter of M, we have the Markov ineqaulity

$$\max_{t\in[0,\tau]}\left|\frac{d}{dt}p_N(\exp(tD))\right| \le R_2N^2 \max_{t\in[0,\tau]}|p_N(\exp(tD))|,$$

for some fixed positive constant R_2.

With a slight abuse of notation we will call

$$Dp_N(t) = \frac{d}{dt}p_N(\exp(tD)).$$

Remark 4.1. *We remark that there are also Bernstein inequalities resulting from generalisations of multiplier operators such as in Ditzian [**?**] for two point homogeneous spaces, and Kushpel and Levesley [**?**] for the complex spheres, which are examples of rank 2 spaces.*

4.3 Norming sets

We are interested in exploiting the notion of norming sets as used in Jetter et al. [**?**], to obtain both global and local error estimates.

Definition 4.3. *A norming set U for a subset $V \subset C(M)$ is a point set such that, for each $v \in V$ there exists $u \in U$ such that $|v(u)| \ge c\|v\|_\infty$ for some $0 < c < 1$.*

Intuitively, the norming set U is rich enough to tell us how big the elements of V are just by measurement on U.

Theorem 4.1. *Any knot set $Y \subset M$ with mesh norm $h(Y) \le (2R_1N)^{-1}$ gives rise to a norming set of P_N with constant $c = 1/2$.*

Proof: We first assume that Y is a knot set with mesh norm $h(Y) \le (2R_1N)^{-1}$. Let $q \in P_N$ be such that $\|q\|_\infty = 1$ (as P_N is a normed linear space there is no loss of generality). We know that M is a compact manifold and q is a continuous function, so it attains its least upper bound on M. Thus $|q(z)| = 1$ for some $z \in M$. From our assumptions and the definition of mesh norm we have:

$$\sup_{m\in M}\inf_{y\in Y} d(m,y) \le (2R_1N)^{-1}.$$

Thus

$$\inf_{y \in Y} d(m, y) \leq (2R_1 N)^{-1},$$

for all $m \in M$. Hence we can find $x \in Y$ such that

$$d(z, x) \leq \frac{1+\epsilon}{2R_1 N}$$

with ϵ taken to be arbitrarily small. Now we use Proposition 4.1 to see that

$$\|Dq\|_\infty \leq R_1 N \|q\|_\infty.$$

If we apply the mean value theorem we see that

$$|q(z) - q(x)| \leq R_1 N d(z, x) \|q\|_\infty < \frac{1+\epsilon}{2}.$$

Thus

$$|q(z) - q(x)| < \frac{1+\epsilon}{2}.$$

However,

$$|q(z) - q(x)| \geq ||q(z)| - |q(x)||$$

and

$$|q(z)| \geq |q(x)|$$

by the definition of $q(z)$. Thus

$$|q(z)| - |q(x)| < \frac{1+\epsilon}{2},$$

so that

$$|q(x)| > |q(z)| - \frac{1+\epsilon}{2}.$$

In other words

$$|q(x)| > 1 - \left(\frac{1+\epsilon}{2}\right) = \frac{1-\epsilon}{2}.$$

Letting $\epsilon \to 0$ gives $|q(x)| \geq 1/2$. Thus we have found an $x \in Y$ for which

$$q(x) > \frac{1}{2}\|q\|_\infty,$$

so that Y is a norming set for P_N. ■

We are also interested in local error estimates and rates on spherical caps of the manifold, $S_\eta = \{y \in M : d(y, p) \leq \eta\}$. Following the same proof and using the Markov inequality in Proposition 4.1 instead of the Bernstein inequality, we have the following corollary.

Corollary 4.2. *Any knot set $Y \subset S_\eta$ with mesh norm $h(Y) \leq (2R_2N^2)^{-1}$ gives rise to a norming set of P_N with constant $c = 1/2$.*

Let n_Y be the cardinality of the set Y. Then we define $l_\infty(Y) = \{v \in \mathbf{R}^{n_Y} : \sup_{y\in Y} |v_y| < \infty\}$. In other words $l_\infty(Y)$ is the set of bounded sequences whose elements are indexed by the members of Y.

Lemma 4.4. *The sampling operator $T : C(M) \to l_\infty(Y)$, where Y is a finite knot set with mesh norm $h(X) \leq (2R_1N)^{-1}$, $Y \subset M$, restricted to U, a finite-dimensional subspace of $C(M)$, is an isomorphism (denoted by T_U).*

Proof: To show that T_U is 1-1 (injective) we suppose that $\ker(T_U) \neq 0$ and obtain a contradiction. So suppose there exists $w \in U$ with $w \neq 0$ and $T_U(w) = 0$. Since Y is a norming set

$$\sup_{y\in Y} |\delta_y(u)| \geq 1/2\|u\|_\infty.$$

However, since $T_U(w) = 0$,

$$\delta_x(w) = w(x) = 0.$$

This implies

$$0 \geq 1/2\|w\|_\infty,$$

so that

$$\|w\|_\infty = 0.$$

However $w \neq 0$, which is the desired contradiction. Thus $\ker T_U = 0$ and T_U is injective. Since T_U is onto by definition

$$T_U(U) = T(U) \subset l_\infty(Y),$$

we obtain the required isomorphism. ■

To prove the main result of this section we will require the Hahn-Banach theorem which one can find in any functional analysis book, in particular Rudin [**?**].

Proposition 4.2. *Let λ be a bounded linear functional on a subspace Z of a normed space W. Then there exists a bounded linear functional $\tilde{\lambda}$ on W which is an extension of λ to W and has the same norm*

$$\|\tilde{\lambda}\|_W = \|\lambda\|_Z.$$

Definition 4.4. *Let X and Y be normed linear spaces and let $L : X \to Y$ be a linear operator. Then the norm of L is defined to be:*

$$\sup_{x \in X} \frac{\|L(x)\|_Y}{\|x\|_X} = \sup_{\|x\|_X = 1} \|L(x)\|_Y.$$

Definition 4.5. *Let X be a normed linear space of functions. Then the dual of X, X^*, is defined to be the linear space of bounded linear functionals on X.*

Definition 4.6. *Let X and Y be a Hilbert spaces and A an operator, $A : X \to Y$. The adjoint of A, A^*, is defined by the following equation:*

$$\langle Ax, y \rangle = \langle x, A^* y \rangle,$$

for all $x \in X$ and $y \in Y$.

Before we start the proof it is useful to draw a small diagram which demonstrates the actions of various operators and their duals, which we have in Figure 4.1.

Theorem 4.2. *Let U be a finite-dimensional subspace of $C(M)$ and $Y^* = \operatorname{span}\{\delta_y : y \in Y\}$, where $Y \subset M$ is a finite knot set. Assume that Y is a norming set of W with norming constant $c \geq 1/2$. Then U^* can be identified with the space Y^*. Moreover, any $w^* \in U^*$ with $\|w^*\|$=1, can be identified with some $\sum_{y \in Y} a_y \delta_y$ where $\sum_{y \in Y} |a_y| \leq 2$.*

Proof: As before, let $T : C(M) \to l_\infty(Y)$ be the sampling operator, and T_U be its restriction to U:

$$T_U = T|_U : C(M) \to T(U) \subset l_\infty(Y).$$

By the Lemma 4.4 T_U is an isomorphism and so possesses an inverse. We first wish to show that $\|T^{-1}U\| \leq 2$. In order to do this we use $1 = \|I\| \leq$

$$
\begin{tikzcd}
\mathrm{U} \arrow[r, "\mathrm{T}"] \arrow[d] & \mathrm{T(U)} \arrow[d] \\
\mathrm{U}^* & (\mathrm{T(U)})^* \arrow[l, "\mathrm{T}^*"]
\end{tikzcd}
$$

Figure 4.1: Operators and their duals

$\|T_U\|\|T_U^{-1}\|$ and prove $\|T_U\| \geq 1/2$. By definition,

$$
\begin{aligned}
\|T_U\| &= \sup_{\|w\|=1} \|T_U(w)\| \\
&= \sup_{\|w\|=1} \sup_{y \in Y} |w(y)| \\
&= \sup_{\|w\|=1} \sup_{y \in Y} |\delta_y(w)| \\
&\geq \sup_{\|w\|=1} 1/2\|w\| = 1/2.
\end{aligned}
$$

From this we deduce that

$$\|T_U^{-1}\| \leq 2.$$

By standard results in functional analysis $[T_U^{-1}]^*$ is an isomorphism and $\|[T_U^{-1}]^*\| = \|T_U^{-1}\| \leq 2$. As T_U^* is an isomorphism, and in particular is onto U^*, then given any $w^* \in U^*$ such that $\|w^*\| = 1$ there exists a $t^* \in (T(U))^*$ with $T_U^*(t^*) = w^*$. Now $t^* = [T_U^*]^{-1}(w^*)$ and since $\|[T_U^*]^{-1}\| \leq 2$ it follows that $\|t^*\| \leq 2$ by the definition of an operator norm.

Since $T(U)$ is a finite-dimensional subspace of $l_\infty(Y)$ we may apply Proposition 4.2 to get a norm preserving extension $l^* \in [l_\infty(Y)]^*$ to t^*, with the representation

$$l^* = \sum_{y \in Y} a_y \delta_y.$$

The norm of this extension

$$\begin{aligned} \|l^*\| &= \sup_{\|w\|=1} |l^*(w)| \\ &= \sup_{\|w\|=1} |\sum_{y \in Y} a_y w(y)| \\ &\leq \sum_{y \in Y} |a_y|, \end{aligned}$$

with equality if $w(y) = a_y/|a_y|$. Hence $\sum_{y \in Y} |a_y| \leq 2$. ■

From Theorems 4.1 and 4.2 we immediately get

Corollary 4.3. *(i) Let $Y \subset M$ be any finite set of interpolation points with density $h(Y) \leq (2R_1N)^{-1}$. Then for any linear functional λ, supported on M, on P_N with $\sup_{p \in P_N} |\lambda p|/\|p\|_\infty = 1$, there exist a set of real numbers $\{a_y : y \in Y\}$ with $\sum_{y \in y} |a_y| \leq 2$, such that*

$$\lambda(P_N) = \sum_{y \in y} a_y p_N(y), p_N \in P_N.$$

(ii) Let $Y \subset S_\eta$ be any finite set of interpolation points with density $h(Y) \leq (2R_2N)^{-2}$. Then for any linear functional λ, supported on M, on P_N with $\sup_{p \in P_N} |\lambda p|/\|p\|_\infty = 1$, there exist a set of real numbers $\{a_y : y \in Y\}$ with $\sum_{y \in y} |a_y| \leq 2$, such that

$$\lambda(P_N) = \sum_{y \in y} a_y p_N(y), p_N \in P_N.$$

4.4 Convergence rates

In this section we use the results of the previous sections to obtain a convergence rate for interpolation of functions in $\mathcal{N}_\kappa$ by κ–splines. We begin with the error estimate of (4.1.2). Choosing $\{a_y : y \in X\}$ to be the coefficients described in Corollary 4.3, we have

$$\begin{aligned} &\inf_{\alpha_y : y \in Y} \sup_{\|g\|_\kappa = 1} \left| g(x) - \sum_{y \in Y} \alpha_y g(y) \right| \\ &\leq \sup_{\|f\|_\kappa = 1} |(g(x) - T_N g(x)) - \sum_{y \in Y} a_y (g(y) - T_N g(y))|, \end{aligned}$$

where T_N is the orthogonal projector onto P_N. Thus,

$$
\begin{aligned}
&\sup_{\|g\|_\kappa=1}\left|g(x)-\sum_{y\in Y} c_y g(y)\right| \\
&\le (1+\sum_{y\in Y}|a_y|)\sup_{\|g\|_\kappa=1}|g(x)-T_N g(x)| \\
&\le 3\max_{x\in M}\sup_{\|g\|_\kappa=1}\left|\sum_{n>N}\sum_{k=1}^{\nu_n} T_{nk}g(x)\right| \\
&\le 3\sup_{\|g\|_\kappa=1}\sum_{n>N}\sum_{k=1}^{\nu_n}\|T_{nk}g\|_\infty \\
&\le 3\sup_{\|g\|_\kappa=1}\left(\sum_{n>N}\sum_{k=1}^{\nu_n} a_{nk}^{-1}\|T_{nk}g\|_2^2\right)^{1/2}\left(\sum_{n>N}\sum_{k=1}^{\nu_n} a_{nk}\frac{\|T_{nk}g\|_\infty^2}{\|T_{nk}g\|_2^2}\right)^{1/2} \\
&\le 3\left(\sum_{n>N}\sum_{k=1}^{\nu_n} a_{nk}d_{nk}\right)^{1/2},
\end{aligned}
$$

where we have used the Cauchy–Schwarz inequality, the definition of the $\|f\|_\kappa$ norm and Corollary 2.1.

Thus we arrive at our main result:

Theorem 4.3. *Let s_Y be the κ–spline interpolant to $f \in \mathcal{N}_\kappa$ on the finite point set $Y \subset M$, where $h(Y) \le 1/(2R_1N)$. Then, for $x \in M$*

$$|f(x)-s_Y(x)| \le 3\left(\sum_{n>N}\sum_{k=1}^{\nu_n} a_{nk}d_{nk}\right)^{1/2}\|f\|_\kappa.$$

Using the same argument we prove

Corollary 4.4. *Let s_Y be the κ–spline interpolant to $f \in \mathcal{N}_\kappa$ on the finite point set $Y \subset S_\eta$ where $h(Y) \le 1/(2R_2N^2)$. Then, for $x \in S_\eta$*

$$|f(x)-s_k(x)| \le 3\left(\sum_{n>N}\sum_{k=1}^{\nu_n} a_{nk}d_{nk}\right)^{1/2}\|f\|_\kappa.$$

We will use the following proposition to make some remarks on the rate of convergence of the interpolation process with respect to the number of interpolation points.

Proposition 4.3. *A compact d-dimensional manifold can be covered with $m = Ch^{-d}$ balls of radius h.*

Consider a covering, $\{U_\alpha : \alpha \in I\}$ of a compact d-dimensional manifold, M. Then this covering has a finite subcover $\{U_{\alpha_i} : 1 \leq i \leq n\}$. Consider the subset of $\mathbf{R}^d$ consisting of the union of the image of the coordinate maps corresponding to U_{α_i}. Cover this set with hypercubes of length h.This can be done with h^{-d} hypercubes. Cover each cube with a hypersphere of radius h. Consider the pre-image of the union of these spheres under the coordinate maps again. As these maps are diffeomorphisms only a certain degree of distortion may occur. Thus the manifold is covered by Ch^{-d} balls. ∎

Remark 4.2. *Suppose the coefficients $a_{nk} = \mathcal{O}(n^{-\beta})$ as $n \to \infty$, and let us assume that the dimensions of the polynomial space H_n is $\mathcal{O}(n^{d-1})$ $(\sum_{k=1}^{\nu_n} d_{nk} = \mathcal{O}(n^{d-1}))$. Then, if $\beta > d$*

$$\begin{aligned}\left(\sum_{n>N}\sum_{k=1}^{\nu_n} a_{nk}d_{nk}\right)^{1/2} &= \mathcal{O}\left(\sum_{n>N} n^{-\beta}n^{d-1}\right)^{1/2} \\ &= \mathcal{O}(N^{(d-\beta)/2}).\end{aligned}$$

Now, for the total manifold M we have $N = \mathcal{O}(h^{-1})$, so we achieve a result of the form

$$|f(x) - s_Y(x)| \leq Ch^{(\beta-d)/2}.$$

From the previous proposition we have that

$$|f(x) - s_Y(x)| \leq Cm^{(d-\beta)/(2d)},$$

where m is the number of interpolation points.

We also investigated whether additional smoothness of the function to be interpolated leads to an improved error estimate (Schaback's 'doubling trick') [**?**], used by Morton and Neamtu for spheres [**?**]. Consider

$$\begin{aligned}\|f - s_Y\|_\kappa^2 &= \langle f - s_Y, f\rangle \\ &= \sum_n\sum_k \frac{\langle T_{nk}(f-s_Y), T_{nk}f\rangle}{a_{nk}} \\ &\leq \left(\sum_n\sum_k T_{nk}(f-s_Y)\right)^{1/2}\left(\sum_n\sum_k \frac{T_{nk}f}{a_{nk}^2}\right)^{1/2} \\ &= \|f\|_{\kappa*\kappa}\|f-s_Y\|_{L_2}. \end{aligned} \tag{4.4.1}$$

Here, as before, $T_{nk}f$ denotes the orthogonal projection onto H_{nk} and

$$\begin{aligned}\kappa * \kappa(x,z) &= \int_M \int_M \kappa(d(x,y))\kappa(d(y,z))d\mu(y) \\ &= \sum_{n=0}^{\infty} \sum_{k=1}^{\nu_n} a_{nk}^2 p_{nk}(x,z).\end{aligned}$$

From 4.1.3, integrating and combining with 4.4.1 we obtain:

$$\|f - s_Y\|_{L_2} \leq 3 \left(\sum_{n>N} \sum_{k=1}^{\nu_n} a_{nk} d_{nk} \right) \|f - s_Y\|_\kappa,$$

since the μ is the normalised measure on M. Combining this inequality with 4.4.1 we obtain:

$$\|f - s_Y\|_\kappa^2 \leq 3\|f\|_{\kappa*\kappa} \left(\sum_{n>N} \sum_{k=1}^{\nu_n} a_{nk} d_{nk} \right) \|f - s_Y\|_\kappa.$$

Dividing through by $\|f - s_Y\|_\kappa$ gives the following result:

Theorem 4.4. *If f is in the native space of $\kappa * \kappa$ then*

$$\|f - s_Y\|_\kappa \leq 3\|f\|_{\kappa*\kappa} \left(\sum_{n>N} \sum_{k=1}^{\nu_n} a_{nk} d_{nk} \right).$$

In the next section we will present the results of some numerical experiments, investigating the effect the relative smoothness of the target function has on convergence rates when interpolating on the complex 2-sphere, for which $d = 3$. We also predicted what the convergence rates should be from the theory. We used a result of Schaback's on the general decay rate of the Fourier transform for Wendland functions (see [?]) to obtain the decay rate of the Fourier transform for our basis functions:

Theorem 4.5. *For $\alpha \in N/2$, the d-variate Fourier transform $F_d(\Psi_{\mu,\alpha})$ of $\Psi_{\mu,\alpha}$(the general Wendland function) with*

$$\mu = [d/2 + \alpha] + 1 \geq 3$$

satisfies

$$F_d(\Psi_{\mu,\alpha})(r) = O(r^{-(d+2\alpha+1)}) \text{ for } r \to \infty.$$

Here μ and α are parameters which allow us to specify a particular Wendland function, given the general function.

We have $\alpha = 1$ and $\alpha = 2$ respectively for the Wendland functions $W_1(r) = (1-r)_+^4(1+4r)$ and $W_2(r) = (1-r)_+^6(35r^2+18r+3)$.

We also employ the following result of zu Castell and Filbir (see [**?**]):

Theorem 4.6. *If, for some $0 < \gamma < d$, the generalised Fourier transform of a radial function ϕ which is conditionally positive definite of order $k \in N_0$ satisfies*

$$\phi(t) = \mathcal{O}(t^{-2k-\gamma}), \quad as \quad t \to 0,$$

then the coefficients a_n in the zonal series expansion

$$\Phi(\zeta,\eta) = \phi((2-2\zeta^t\eta)^{1/2}) = \sum_{n=0}^{\infty} a_n \sum_{k=1}^{c_n,k} S_{n,k}(\zeta)S_{n,k}(\eta), \quad \zeta,\eta \in S^{d-1},$$

satisfy

$$a_n = \mathcal{O}(n^{-2k-\gamma+1}), \ as \ t \to \infty.$$

After using our Remark 4.2, specifically that the convergence rate is $O(n^{(d-\beta)/2})$, this led to the predicted convergence rates, which we summarise in Table **??**.

Specifically, as $d = 3$, and $\alpha = 1$ for W_1, from Theorem 4.5 the decay rate for the Fourier transform is $\mathcal{O}(r^{-6})$, $r \to \infty$, giving $\mathcal{O}(n^{-5})$ for the decay of the Fourier coefficients on using Theorem 4.6. Finally on applying Remark 4.2 with $d = 3$ we obtain rate of convergence $1/3 = (5-3)/6$ for the convergence rate. For $\alpha = 2$ for W_2 we obtain similarly the rate of convergence $2/3 = (7-3)/6$.

We also ran our experiments with $r^2 \log r$ as basis function. We appreciate that this is not an entirely appropriate procedure for our theory, as this basis function is conditionally positive definite, not positive definite. However, we felt it would still be interesting as it would allow us to use the decay rate for the Gegenbauer coefficients, associated integrals for which we calculated in Chapter 3, combining it with Remark 4.2 to obtain a predicted convergence rate which we could then compare with the results of our experiment. The decay rate of the coefficients is $\mathcal{O}(n^{-2-2\alpha})$ from (3.1.3) .

For the thin plates spline $\alpha = 1$, and for the complex 2-sphere case we have with $\lambda = 1$

$$K_n = \int_{-1}^{1} (1-t)\log(1-t)(1-t^2)^{1/2} P_n^{(1)}(t)dt = \mathcal{O}(n^{-4}).$$

However, for the application of the machinery above we need the coefficients we are referring to in Remark 4.2 above.

It is straightforward to show that, if

$$(1-t)\log(1-t) = \sum_{n=0}^{\infty} b_n P_n^{(1)}(t),$$

then

$$b_n = \frac{K_n}{h_n}, \quad n \in \mathbf{N},$$

where

$$h_n = \int_{-1}^{1} (1-t^2)^{1/2} |P_n^{(1)}(t)|^2 dt.$$

Also, if

$$b_n P_n^{(1)}(xy) = a_n \sum_{j=1}^{d_n} Y_{nj}(x) Y_{nj}(y), \quad x, y \in M,$$

then, putting $x = y$ and integrating over M gives

$$b_n P_n^{(1)}(1) = a_n d_n,$$

so that

$$b_n = \frac{a_n d_n}{P_n^{(1)}(1)}.$$

Thus, the coefficient a_n is related to our computed coefficient K_n via

$$a_n = \frac{K_n P_n^{(\lambda)}(1)}{h_n d_n}.$$

From [**?**], we have $h_n = \mathcal{O}(1)$, and $P_n^{(\lambda)}(1) = \mathcal{O}(n)$. For the complex 2-sphere the dimension of the harmonic polynomial spaces is $d_n = \mathcal{O}(n^2)$, just as for the real 3-sphere. Hence,

$$a_n = \mathcal{O}(n^{-5}).$$

This is the same as for W_1 above, and we similarly get a convergence rate of 1/3.

Table 4.1: Predicted convergence rates

W_1	W_2	$r^2 \log r$
1/3	2/3	1/3

4.5 Numerical Experiments

This section presents some numerical experiments, looking at interpolation on a particular manifold. Our aim was to see how the smoothness of the basis functions and the target functions affected experimental convergence rates. The manifold that we investigated was the complex 2-sphere.

Definition 4.7. *The complex 2-sphere is defined by the following:*

$$\{(r_1 e^{i\theta_1}, r_2 e^{i\theta_2}) | (r_1)^2 + (r_2)^2 = 1, 0 \leq \theta_1, \theta_2 \leq 2\pi\}.$$

In the first set of experiments we used the radial basis functions W_1 and W_2 to form our interpolants. We set up a grid of n^3 data points, where n was varied, on the complex sphere, to interpolate functions of increasing smoothness $f_1(x) = \|x - \eta\| + \|x - \xi\|$, $f_2(x) = \|x - \eta\|^3 + \|x - \xi\|^3$ and $f_3(x) = \|x - \eta\|^5 + \|x - \xi\|^5$, where $\eta = (0, \mathbf{i})$, and $\xi = (0.1 + 0.1\mathbf{i}, (0.97)^{1/2} + 0.1\mathbf{i})$.

Let s_n denote the error of interpolation for n^3 data points, and $e_n = \max_{z \in Z} |(f(z) - s_n(z))|$, where Z is the set of test points on a 15^3 grid. The decay rate for the approximate infinity error is given by:

$$\frac{\log(e_n) - \log(e_{n-1})}{3(\log(n) - \log(n-1))}.$$

We then took the geometric mean of the convergence rates for $n = 2$ to $n = 13$ with the results given in Tables **??** and **??**.

Table 4.2: Convergence rates for with W_1 as basis

f_1	f_2	f_3
0.59	1.44	1.62

In both cases the greater smoothness of f_2 and f_3 over f_1 ensured a faster rate of convergence, and we somewhat a doubling of the rate of convergence, even

Table 4.3: Convergence rates for with W_2 as basis

f_1	f_2	f_3
0.67	1.43	2.26

Table 4.4: Convergence rates for with $r^2 \log r$ as basis

f_1	f_2	f_3
0.64	1.28	1.52

thought we are not seeing the predicted rates (but twice these). We should also note that we see little improvement in convergence from f_2 to f_3 suggesting that we have reached saturation with the smoothness.

The additional smoothness of f_3 over f_2 particularly showed up when interpolating with W_2, suggesting we are have nor reached saturation with f_2 this time. Also, we recover similar rates of convergence for f_1 in both cases. This result is what is predicted in the Euclidean case by the results of Light and Brownlee [**?**] and Narcowch, Ward and Wendland [**?**], who explore convergence when the target function is not in the native space of the associated basis function.

As mentioned above we also ran an experiment with $r^2 \log r$ as basis function. These results are in Table **??**.

As we expected the convergence rates are similar to those for W_1. Again the convergence rates are larger than expected, particularly with the smoother functions. We believe, in the case of the smoother functions, that error doubling once again is responsible for this.

4.6 Concluding Remarks

In this chapter we have presented a calculation of an error bound for compact homogeneous spaces on interpolating with a G-invariant kernel using the norming set approach. Another method, the smoothness approach, has provided a tighter error bound. It remains to be seen whether it is possible to obtain a similar result using norming sets.

This thesis has dealt with interpolation on a particular class of differentiable manifold, the compact homogeneous space. We have proved some results for a specific set of compact homogeneous spaces, the spheres, and also proved a result of much greater generality. We have proved the formulae for the Fourier-Gegenbauer coefficients for two radial basis functions using elementary methods, formulae which enable the determination of the native spaces for these functions. Further work here could include a similarly elementary calculation of the Fourier-Jacobi coefficients for two-point spaces other than the spheres. Our experience with spheres was used to generalise a result known for spheres to compact homogeneous spaces in general. This result was the calculation of an error bound for such spaces on interpolation with G-invariant kernels. Further work here could include an adaptation of the method so that the degradation of the error bound in comparison with the smoothness approach is overcome. Also the work could be applied to a compact homogeneous space in particular. Additionally a number of numerical experiments were performed and we discovered that the smoothness of the basis function and that of the target function can very much affect experimental convergence rates.

References

[1] M. Abramowitz and I. Stegun, *Handbook of Mathematical Functions*, Dover Publications, 1972.

[2] A. Baker, *Matrix Groups. An Introduction to Lie Group Theory.*, Springer, 2002.

[3] B. Baxter and S. Hubbert, Radial basis functions for the sphere, in *Recent Advances in Approximation Theory* (ISNM Vol. 137), (W.a Haussmann, K. Jetter and M. Reimer eds), Birkhauser, 2001, 33–47.

[4] W. M. Boothby, *An Introduction to Differentiable Manifolds and Riemannian Geometry.*, Academic Press, 2003.

[5] L. P. Bos and P. D. Milman, Sobolev-Gagliardo-Nirenberg and Markov type inequalities on subanalytic domains, *Geometric and Functional Analysis* **5** (1996), 854–923.

[6] R. Brownlee, and W. A. Light, Approximation orders for interpolation by surface splines to rough functions, *IMA J. Numer. Anal.* **24** (2004) 179–192.

[7] W. Cheney and W. Light, *A Course in Approximation Theory*, Brooks/Cole Series in Advanced Mathematics, Brooks/Cole Publishing Company, 2000.

[8] P. Davis, *Interpolation and Approximation* , Dover Publications, Inc., 1975.

[9] Z. Ditzian, Fractional derivatives and best approximation, *Acta Mathematica Hungarica* **81** (1998), 323–348.

[10] J. Duchon, Splines minimizing rotation invariant seminorms in Sobolev spaces, *Constructive Theory of Functions of Several Variables*, **1** (1976), 85-100.

[11] N. Dyn, F. J. Narcowich and J. D. Ward, Variational principles and Sobolev–type estiates for generalised interpolation on a Riemannian manifold, *Contr. Approx.* **15** (1999), 175–208.

[12] F. Filbir and W. zu Castell, Radial basis functions and corresponding zonal series expansions on the sphere, *Journal of Approximation Theory* **134** (2005), 65–79.

[13] M. Von Golitschek, and W. A. Light, Interpolation by polynomials and radial basis functions on spheres, *Constr. Approx.* **17** (2001), 1–18.

[14] S. Helgason, *Differential Geometry and Symmetric Spaces*, Pure and Applied Mathematics **12**, Academic Press, 1962.

[15] E. Hewitt, and K. A. Ross, *Abstract Harmonic Analysis I and II*, Springer-Verlag, 1963 and 1969.

[16] S. Hubbert, On the accuracy of surface spline interpolation on the unit sphere (private communication)

[17] S. Hubbert and J. Levesley, Open problem concerning fourier transforms of radial functions in Euclidean space and on sphere, in *Recent Advances in Approximation Theory* (ISNM Vol. 137), (W. Haussmann, K. Jetter and M. Reimer eds), Birkhauser, 2001, 225–226.

[18] K. Jetter, J. Stöckler, and J. D. Ward, Error estimates for scattered data interpolation on spheres, *Math. Comp.* **68** (1999), 733–747.

[19] A. Kroo and S. Revesz, On Bernstein and Markov-Type Inequalities for Multivariate Polynomials on Convex Bodies, *Journal of Approximation Theory* **99** (1999), 134–152.

[20] A. K. Kushpel and J. Levesley, A multiplier version of the Bernstein inequality on the complex sphere, submitted for publication, 2012.

[21] J. Levesley and D. L. Ragozin, Radial basis interpolation on homogeneous manifolds: Convergence rates, *Advances in Computational Mathematics* **27** (2007), 237–246.

[22] J. Levesley and D. L. Ragozin, The density of translates of zonal kernels on compact homogeneous spaces, *Journal of Approximation Theory* **103** (2000), 252–268.

[23] J. Levesley and C. Odell, Evaluation of some integrals arising in approximation on the sphere, *Numerical Functional Analysis and Optimization* **23** (2002), 359–365.

[24] J. Levesley, C. Odell, and David L. Ragozin, Scattered Data on Homogeneous Manifolds: the Norming Set Approach, in *Curve and Surface Fitting: Saint-Malo 2002*, (A. Cohen, J.-L. Merrien, and L. L. Schumaker eds), Nashboro Press, 2003, 69–278.

[25] W. Light, J. Levesley, D.L. Ragozin and X. Sun, A simple approach to the variational theory for interpolation on spheres", in *International Series of Numerical Mathematics* **132**, Birkhauser, 1999.

[26] J. C. Mairhuber, On Haar's theorem concerning Chebyshev approximation problems having unique solutions, *Proc. Am. Math. Soc.*, **7** (1956), 609-615.

[27] T. M. Morton and M. Neamtu, Error bounds for solving pseudodifferential equations on spheres by collocation with zonal kernels, *Journal of Approximation Theory* **114** (2002), 242–268.

[28] C. Muller, *Spherical Harmonics*, Springer, 1966.

[29] F. J. Narcowich, J. D. Ward, and H. Wendland, Refined error estimates for radial basis function interpolation, *Constr. Approx.*, **19** (2003), 541–564.

[30] F. J. Narcowich, X. Sun, and J. D. Ward, Approximation Power of RBFs and Their Associated SBFs: A Connection, *Advances in Computational Mathematics* **27** (2007), 107–124.

[31] A. Ron, and X. Sun, Strictly positive definite functions on spheres in Euclidean spaces, *Math. Comp.* **65** (1996), 1513–1530.

[32] W. Rudin, *Functional Analysis*, International Series in Pure & Applied Mathematics, McGraw-Hill Higher Education, 1990.

[33] D. L. Ragozin, Polynomial approximation on compact manifolds and homogeneous spaces, *Transactions of the American Mathematical Society* **150** (1970), 41–53.

[34] R. Schaback, Improved error bounds for scattered data interpolation by radial basis functions, *Mathematics of Computation* **68** (1999), 201–216.

[35] R. Schaback, The missing Wendland functions (preprint).

[36] I. J. Schoenberg, Positive definite functions on spheres, *Duke Math. J.* **9** (1942), 96–108.

[37] G. Szegö, *Orthogonal Polynomials*, Amer. Math. Soc., Providence, 1991.

[38] H. Wendland, Piecewise polynomials, positive definite and compactly supported radial functions of minimal degree,*Adv. Comput. Math.* **4** (1995), 389–396.

[39] Y. Xu, and E. Cheney, Strictly positive definite functions on spheres, *Proc. Amer. Math. Soc.* **116** (1992), 977–981.

[40] W. zu Castell and Frank Filbir, Radial basis functions and corresponding zonal series expansions on the sphere, *Journal of Approximation Theory*, Vol.134, Issue 1 (2005) 65–79.