L^P BERNSTEIN INEQUALITIES

AND RADIAL BASIS FUNCTION APPROXIMATION

A Dissertation

by

JOHN PAUL WARD

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

August 2010

Major Subject: Mathematics

L^P BERNSTEIN INEQUALITIES
AND RADIAL BASIS FUNCTION APPROXIMATION

A Dissertation

by

JOHN PAUL WARD

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Approved by:

Co-Chairs of Committee,	Francis J. Narcowich
	Joseph D. Ward
Committee Members,	P. Fred Dahm
	Joel Zinn
Head of Department,	Albert Boggess

August 2010

Major Subject: Mathematics

ABSTRACT

L^p Bernstein Inequalities

and Radial Basis Function Approximation. (August 2010)

John Paul Ward, B.S., University of Georgia

Co–Chairs of Advisory Committee: Dr. Francis J. Narcowich
Dr. Joseph D. Ward

In approximation theory, three classical types of results are direct theorems, Bernstein inequalities, and inverse theorems. In this paper, we include results about radial basis function (RBF) approximation from all three classes. Bernstein inequalities are a recent development in the theory of RBF approximation, and on $\mathbb{R}^d$, only L^2 results are known for RBFs with algebraically decaying Fourier transforms (e.g. the Sobolev splines and thin-plate splines). We will therefore extend what is known by establishing L^p Bernstein inequalities for RBF networks on $\mathbb{R}^d$. These inequalities involve bounding a Bessel-potential norm of an RBF network by its corresponding L^p norm in terms of the separation radius associated with the network. While Bernstein inequalities have a variety of applications in approximation theory, they are most commonly used to prove inverse theorems. Therefore, using the L^p Bernstein inequalities for RBF approximants, we will establish the corresponding inverse theorems. The direct theorems of this paper relate to approximation in $L^p(\mathbb{R}^d)$ by RBFs which are perturbations of Green's functions. Results of this type are known for certain compact domains, and results have recently been derived for approximation in $L^p(\mathbb{R}^d)$ by RBFs that are Green's functions. Therefore, we will prove that known results for approximation in $L^p(\mathbb{R}^d)$ hold for a larger class of RBFs. We will then show how this result can be used to derive rates for approximation by Wendland functions.

To my family

ACKNOWLEDGMENTS

The author would like to express his appreciation to his advisors, Dr. Narcowich and Dr. Ward, for their help with the work contained in this dissertation.

TABLE OF CONTENTS

1. INTRODUCTION

Radial basis function (RBF) approximation is primarily used for constructing approximants to functions that are only known at discrete sets of points. Some advantages of this theory are that RBF approximation methods work well in arbitrarily high dimensional spaces, where other methods break down, and ease of implementataion. For these reasons, RBF techniques are being used to solve a variety of applied problems, two examples being problems in statistical learning theory and numerical partial differential equations.

Given an RBF $\Phi : \mathbb{R}^d \to \mathbb{R}$ and a countable set of points $X \subset \mathbb{R}^d$, we can define an RBF approximation space $S_X(\Phi)$ to be a collection of linear combinations of functions from $\{\Phi(\cdot - \xi) : \xi \in X\}$. Approximation by such spaces has its origins in the work of Duchon [8, 9], on the thin-plate splines, and Hardy [12], on the multiquadrics, in the 1970s. Related work, including results about positive (semi-)definite functions, can be traced back even further to the 1920s, [14]. However, it was not until the late 1980s that a solid foundation began to be laid for RBF approximation. In the late 80's through the early 90's, researchers investigated the case where the sets X were required to be the scaled lattice points in $\mathbb{R}^d$. This was an ideal starting point where one has access to strong results like the Poisson summation formula and Wiener's lemma. With the work of researchers such as M.D. Buhmann and M.J.D. Powell, we now have a good understanding of approximation in this setting, and much of this work is summarized in the book of Buhmann, [5]. Unfortunately, in practice, one is not always given the choice to work with such nice data sets. Therefore, research in this field now focuses mainly on approximation spaces defined on scattered data

---

This dissertation follows the style of the Journal of Approximation Theory.

sites. F.J. Narcowich and J.D. Ward have made significant contributions to this field studying approximation on compact manifolds where the data sites are allowed to be quasi-uniform.

The initial approach to scattered data approximation on $\mathbb{R}^d$ concerned the stationary setting, with the RBFs being scaled to be proportional to the fill distance. In this paper, we will study the alternative nonstationary setting, which is relatively new and still contains many open problems. One notable contribution to this area has recently been produced by DeVore and Ron, [7]. In that paper, the authors derived rates for the approximation of L^p functions by RBF spaces $S_X(\Phi)$ where X consists of scattered points in $\mathbb{R}^d$ and Φ satisfies a Green's function-type condition. The Wendland functions, which have some nice properties and are hence popular for applications, do not fit the framework of [7]. Therefore, this paper will address a generalization of a result from [7] that could potentially be broad enough to capture the Wendland functions. Considering the fact that the Wendland functions, when restricted to the sphere $\mathbb{S}^n$, are perturbations of Green's functions, cf. [16, Proposition 3.1], we will extend a result of [7] to include similar perturbations.

The other main topic of this paper will be Bernstein inequalities for RBF approximants. In 1912, S.N. Bernstein proved the first inequality of this type for L^∞ norms of trigonometric polynomials, [4]. A generalization can be found in [6]; this result, which is credited to Zygmund, states that any trigonometric polynomial T of degree n satisfies

$$\left|\left|T^{(r)}\right|\right|_p \le n^r \left|\left|T\right|\right|_p$$

for $1 \le p \le \infty$. A Bernstein inequality of an RBF approximant will take the form of a L^p Sobolev norm being bounded by the corresponding L^p norm. The exact form and the relationship to the classical Bernstein inequality will be explained in Section

2. The first example of a Bernstein-type inequality for RBF approximants was proved in 2001 by Schaback and Wendland for approximants in bounded domains, [19]; the authors were able to show that a particular L^2 Sobolev norm of an approximant is bounded by its L^∞ norm. More recently, Narcowich, Ward, and Wendland proved a more standard type of Bernstein inequality, [18]. They proved L^2 Bernstein inequalities for approximants coming from an RBF approximation space $S_X(\Phi)$ on $\mathbb{R}^d$ where the Fourier transform of the RBF has algebraic decay. In the same year, Mhaskar proved L^p Bernstein inequalities for certain Gaussian networks on $\mathbb{R}^d$, [15]. Lastly, in [16], Mhaskar, Narcowich, Prestin, and Ward were able to prove Bernstein inequalities in L^p norms for a large class of spherical basis functions (SBFs). In this paper, we will prove results on $\mathbb{R}^d$ that are analagous to those in [16].

In the remainder of this section we will cover preliminary material that will be needed throughout the paper. We will begin by describing the approximation procedure that will be used. This will be followed by the definitions of the Fourier transform and several classes of smooth function spaces. Afterward, we will state some inequalities for L^p functions and close the section by recalling some results from measure theory.

Section 2 will cover the foremost topic of the paper; we will prove L^p Bernstein inequalities for RBFs that have algebraically decaying Fourier transforms. Proving these inequalities will require us to examine two properties of the RBF approximants. First, we will need to show that for $g = \sum_{\xi\in X} a_\xi \Phi(\cdot - \xi) \in S_X(\Phi)$, $||a||_p$ can be bounded by $||g||_p$, thus showing that a is stable with respect to the norm of g. The other property concerns the approximation of g by band-limited functions. After proving Bernstein inequalites for the RBF approximants, we will be able to use them to derive corresponding inverse theorems.

Direct theorems concerning approximation by RBFs will be the focus of Section

3. The abilty of Green's functions to invert differential operators was shown to be useful for determining approximation results in [7]; however, there are some common RBFs that do not fit into this framework. We will begin by proving analogous results for approximation by perturbations of Green's functions. This will then be followed by some examples. In particular, polynomial reproducing functionals will be used to show that the Wendland functions can be well approximated by local translates of themselves, and an error bound for approximation by some particular Wendland functions will be derived.

We conclude the paper in Section 4 with a summary of our results. We will additionally discuss some possibilities for future research based on both of the main topics contained in this paper. Due to the novelty of Bernstein inequalities and inverse theorems in RBF approximation, there are several directions in which to continue the research of Section 2. Future work relating to Section 3 will initially focus on proving approximation bounds for all of the Wendland functions.

1.1 RBF Approximation Spaces

In this paper we will be concerned with the approximation of functions in $L^p(\mathbb{R}^d)$ for $1 \leq p \leq \infty$. The approximants will be finite linear combinations of translates of an RBF Φ, and the translates will come from a countable set $X \subset \mathbb{R}^d$. The error of this approximation, which is measured in a Sobolev-type norm, depends on both the function Φ and the set X. Therefore, given an RBF Φ and a set X, we define the RBF approximation space $S_X(\Phi)$ by

$$S_X(\Phi) = \left\{ \sum_{\xi \in Y} a_\xi \Phi(\cdot - \xi) : Y \subset X, \#Y < \infty \right\} \cap L^1(\mathbb{R}^d).$$

By choosing Φ and X properly, one is able to prove results about rates of approximation as well as the stability of the approximation procedure.

Some RBFs that are commonly used in applications are the Gaussians, thin-plate splines, multiquadrics, Wendland functions, and Sobolev splines. In Section 2, we shall only consider RBFs that have algebraically decaying (generalized) Fourier transforms. Note that both the thin-plate splines and Sobolev splines fall into this category.

When analyzing an approximation procedure, one often wants to determine the error of the approximation and the stability of the approximation procedure. When considering an RBF approximation space $S_X(\Phi)$, these quantities are bounded in terms of certain measurements of the set X. The error of approximation will typically be given in terms of the fill distance

$$h_X = \sup_{x \in \mathbb{R}^d} \inf_{\xi \in X} ||x - \xi||_2 \,,$$

which measures how far a point in $\mathbb{R}^d$ can be from X, and the stability of the approximation will be determined by the separation radius

$$q_X = \frac{1}{2} \inf_{\substack{\xi, \xi' \in X \\ \xi \neq \xi'}} ||\xi - \xi'||_2 \,,$$

which measures how close two points in X may be. In order to balance the rate of approximation with the stability of the procedure, approximation will typically be restricted to sets X for which h_X is comparable to q_X, and sets for which the mesh ratio $\rho_X := h_X/q_X$ is bounded by a constant will be called quasi-uniform. In this paper, we will only consider approximation spaces $S_X(\Phi)$ where X is quasi-uniform.

1.2 The Fourier Transform

The Fourier transform is a fundamental tool for proving many results in approximation theory. We will use it to charcterize the RBFs under consideration and to prove several results. Throughout this paper, we will use the following convention. Given a function $f : \mathbb{R}^d \to \mathbb{R}$ in L^1, its Fourier transform $\hat{f}$ will be defined by

$$\hat{f}(\omega) = \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} f(x) e^{-ix\cdot\omega} dx.$$

1.3 Smoothness Classes

Let E be a subset of $\mathbb{R}^d$, then we denote by $C^k(E)$ the collection of real valued functions defined on E that have continuous partial derivatives up to order k. The set of functions in $C^k(\mathbb{R})$ that converge to 0 at infinity will be denoted by C_0^k, and we represent the compactly supported elements of $C^k(\mathbb{R}^d)$ by C_c^k.

We define the Schwartz class $\mathcal{S}$ on $\mathbb{R}^d$ as follows. A function $f : \mathbb{R}^d \to \mathbb{R}$ is said to be of Schwartz class if for all multi-indices α and β, there exists a constant $C_{\alpha,\beta} > 0$ such that

$$\left| x^\alpha D^\beta f(x) \right| \leq C_{\alpha,\beta}$$

for all $x \in \mathbb{R}^d$.

We shall use the standard definition for the L^p spaces. A Lebesgue measurable function $f : \mathbb{R}^d \to \mathbb{R}$ is in $L^p(\mathbb{R}^d)$ for $1 \leq p < \infty$ if

$$||f||_p = \left(\int_{\mathbb{R}^d} |f(x)|^p \, dx \right)^{1/p} < \infty,$$

and a Lebesgue measurable function f is said to be in $L^\infty(\mathbb{R}^d)$ if

$$||f||_\infty = \operatorname*{ess\,sup}_{x\in\mathbb{R}^d} |f(x)| < \infty.$$

We will say a measurable function f is locally integrable, denoted by $L^1_{\text{loc}}(\mathbb{R}^d)$, if

$$\int_E |f(x)|\,dx < \infty$$

for each bounded, measurable set $E \subseteq \mathbb{R}^d$.

We will be concerned with approximating functions that lie in subspaces of L^p spaces. The spaces that we will be mainly interested in are the Bessel-potential spaces $L^{k,p}(\mathbb{R}^d)$, which coincide with the standard Sobolev spaces $W^{k,p}(\mathbb{R}^d)$ when k is a positive integer and $1 < p < \infty$, cf. [20, Section 5.3]. The Bessel potential spaces are defined by

$$L^{k,p} = \{f : \hat{f} = (1 + ||\cdot||_2^2)^{-k/2}\hat{g}, g \in L^p(\mathbb{R}^d)\}$$

for $1 \leq p \leq \infty$, and they are equipped with the norm

$$||f||_{L^{k,p}} = ||g||_p \, .$$

In Section 3, we will be working with smoothness spaces associated with linear operators. If $T : C_c^k(\mathbb{R}^d) \to C_c(\mathbb{R}^d)$ is linear, then we define a semi-norm and norm on $C_c^\infty(\mathbb{R}^d)$ by

$$\begin{aligned} |f|_{W(L^p(\mathbb{R}^d),T)} &= ||Tf||_p \\ ||f||_{W(L^p(\mathbb{R}^d),T)} &= ||f||_p + |f|_{W(L^p(\mathbb{R}^d),T)} \, . \end{aligned}$$

The completion of $C_c^\infty(\mathbb{R}^d)$, with respect to the above norm, will be denoted by $W(L^p(\mathbb{R}^d), T)$.

1.4 Some Results for L^p Spaces

We define the convolution of two measurable functions f and g by

$$f * g = \int_{\mathbb{R}^d} f(\cdot - t)g(t)dt$$

wherever the integral exists. The L^p norm of the convolution of two measurable functions can be bounded using Young's inequality. One version of this inequality is the following.

Theorem 1.1. ([11, Theorem 8.7]) *If $f \in L^1$ and $g \in L^p$ $(1 \leq p \leq \infty)$, then $f * g(x)$ exists for almost every x, $f * g \in L^p$, and*

$$\left|\left|f * g\right|\right|_p \leq \left|\left|f\right|\right|_1 \left|\left|g\right|\right|_p .$$

Another useful result for L^p norms is Hölders's inequality. Suppose $1 \leq p \leq \infty$ and $p^{-1} + q^{-1} = 1$. Given two measurable functions f and g, we have

$$\left|\left|fg\right|\right|_1 \leq \left|\left|f\right|\right|_p \left|\left|g\right|\right|_q ,$$

cf. [11, Chapter 6].

1.5 Measure Theory

We will now cover some results for $M(\mathbb{R}^d)$, the class of finite Borel measures on $\mathbb{R}^d$. This space is equipped with a norm defined by

$$\left|\left|\mu\right|\right| = \left|\mu\right| (\mathbb{R}^d),$$

where $|\mu|$ is the total variation of μ. We would first like to point out that the definition of the Fourier transform can be extended to $M(\mathbb{R}^d)$ as follows:

$$\hat{\mu}(\omega) = \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} e^{-ix\cdot\omega} d\mu(x).$$

Next we will show how to bound the norm of the convolution of a measure with a function. Note that the following proposition provides an analog of Young's inequality.

Proposition 1.2. ([11, Proposition 8.49]) *If $f \in L^p(\mathbb{R}^d)$ $(1 \leq p \leq \infty)$ and $\mu \in M(\mathbb{R}^d)$, then the integral $f * \mu(x) = \int_{\mathbb{R}^d} f(x-t)d\mu(t)$ exists for almost every x, $f * \mu \in L^p$, and $||f * \mu||_p \leq ||f||_p \, ||\mu||$.*

In Section 3, we will need to consider a decomposition of measures. Specifically, any measure $\mu \in M(\mathbb{R}^d)$ can be written as $\mu = \mu_a + \mu_s + \mu_d$, where μ_a is absolutely continuous with respect to Lebesgue measure, μ_d is a countable linear combination of Dirac measures, and $\mu_s = \mu - \mu_a - \mu_d$ is the singular continuous part of μ.

2. BERNSTEIN INEQUALITIES AND INVERSE THEOREMS

In this section, we will establish L^p Bernstein inequalities for certain RBF approximation spaces $S_X(\Phi)$, and these inequalities will take the form $||g||_{L^{k,p}} \leq Cq_X^{-k} ||g||_p$. To prove this, we will use band-limited approximation with the bandwidth proportional to $1/q_X$. Thus $1/q_X$ acts similarly to a Nyquist frequency, and viewing $1/q_X$ as a frequency, we can see the connection to the classical Bernstein inequalities for trigonometric polynomials. In particular, bandwidth is playing the role of the degree of the polynomial from the classical inequality.

The basic strategy that we will use is the following, which is the same as the one used in [16]. Given $g = \sum_{\xi\in X} a_\xi \Phi(\cdot - \xi) \in S_X(\Phi)$, we choose an appropriate band-limited approximant g_σ, and we have

$$||g||_{L^{k,p}} \leq ||g_\sigma||_{L^{k,p}} + ||g - g_\sigma||_{L^{k,p}}.$$

We then split the second term into two ratios.

$$||g||_{L^{k,p}} \leq ||g_\sigma||_{L^{k,p}} + \left(\frac{||a||_p}{||g||_p} \frac{||g - g_\sigma||_{L^{k,p}}}{||a||_p} \right) ||g||_p. \tag{2.1}$$

The term $||a||_p \,/\, ||g||_p$ will be bounded by a stabilty ratio $\mathcal{R}_{S,p}$ that is independent of the function g. We will then need to bound the error of approximating g by band-limited functions. Combining these results with a Bernstein inequality for band-limited functions, we will be able to prove the Bernstein inequality for all functions in $S_\Phi(X)$.

2.1 Radial Basis Functions

In order to prove the Bernstein inequality, we will need to require certain properties of the RBFs involved. As much of the work will be done in the Fourier domain, we state the constraints in terms of the RBFs' Fourier transforms. Given a radial function $\Phi : \mathbb{R}^d \to \mathbb{R}$ with (generalized) Fourier transform $\hat{\Phi}$, let $\phi : (0, \infty) \to \mathbb{R}$ be the function defined by $\hat{\Phi}(\omega) = \phi(||\omega||_2)$. We will say a function Φ is admissible of order β if there exist constants $C_1, C_2 > 0$ and $\beta > d$ such that for all $\sigma \geq 1$, $x \geq 1/2$, and $l \leq l_d := \lceil (d+3)/2 \rceil$, we have

(i) $C_1 \leq \phi(\sigma x)(1 + (\sigma x)^2)^{\beta/2} \leq C_2$

(ii) $\left| \left((1 + (\sigma x)^2)^{\beta/2} \phi(\sigma x) \right)^{(l)} \right| \leq C_2$

Two particular classes of admissible functions are the Sobolev splines and the thin-plate splines. The Sobolev spline Φ of order $\beta > d$ is given by

$$\Phi = C \, ||\cdot||_2^{(\beta - d)/2} \, K_{(d-\beta)/2}(||\cdot||_2),$$

where K is a modified Bessel function of the third kind. This function possesses the Fourier transform

$$\hat{\Phi} = (1 + ||\cdot||_2^2)^{-\beta/2}.$$

Since we will mainly be working with $\hat{\Phi}$, we will not discuss formulas and properties of K and Φ. Instead, we direct the reader to [1, 2, 20, 22, 23]. For a postive integer $m > d/2$, the thin-plate splines of order $2m$ take the form

$$\Phi = \begin{cases} ||\cdot||_2^{2m-d} & d \text{ odd} \\ ||\cdot||_2^{2m-d} \log ||\cdot||_2 & d \text{ even} \end{cases}$$

and possess the generalized Fourier transform

$$\hat{\Phi} = C \, ||\cdot||_2^{-2m} \, .$$

For further information on these functions see [23].

2.2 Stability

One of the essential results for proving the Bernstein inequalities is a bound of a stability ratio for $S_X(\Phi)$. We define the L^p stability ratio $\mathcal{R}_{S,p}$ associated with this collection by

$$\mathcal{R}_{S,p} = \sup_{S_X(\Phi)\ni g\neq 0} \frac{||a||_p}{||g||_p},$$

where $g = \sum_{\xi\in X} a_\xi \Phi(\cdot - \xi)$. The goal of this subsection is to bound the stabilty ratio by $Cq_X^{d/p'-\beta}$ for some C independent of a and X, where p' is the conjugate exponent to p. For this subsection, we will assume we are working with a fixed countable set $X \subset \mathbb{R}^d$ with $0 < q_X < 1$ and an admissible function Φ of order β.

To begin, fix $Y = \{\xi_j\}_{j=1}^N \subset X$ and $g = \sum_{j=1}^N a_j \phi(\cdot - \xi_j)$. We will derive a bound for $||a||_p / ||g||_p$ and show the bound is independent of Y and a. The strategy for proving this is as follows. Let K be a smooth function and define $\hat{K}_\sigma(\omega) = \hat{K}(\omega/\sigma)$. We will then consider the convolutions $K_\sigma * g(x) = \sum_{j=1}^N a_j K_\sigma * \Phi(x - \xi_j)$. For an appropriate choice of σ, the interpolation matrix $(A_\sigma)_{i,j} = K_\sigma * \Phi(\xi_i - \xi_j)$ will be invertible, and the norm of its inverse will be bounded. Then $a = A_\sigma^{-1}(K_\sigma * g)|_Y$ and $||a||_p \leq ||A_\sigma^{-1}||_p \, ||K_\sigma * g|_Y||_p$. We will then be left with bounding $||K_\sigma * g|_Y||_p$ in terms of q_X and $||g||_p$.

2.2.1 Convolution Kernels

We now define the class of smooth functions with which we will convolve g. Let $\mathcal{K}_1$ be the collection of Schwartz class functions $K : \mathbb{R}^d \to \mathbb{R}$ that satisfy:

(i) There is a $\kappa : [0, \infty) \to [0, \infty)$ such that $\hat{K}(\omega) = \kappa(||\omega||_2)$.

(ii) $\kappa(r) = 0$ for $r \in [0, 1]$ and κ is nonvanishing on an open set.

For each $K \in \mathcal{K}_1$, we define the family of functions $\{K_\sigma\}_{\sigma \geq 1}$ by $\hat{K_\sigma}(\omega) = \hat{K}(\omega/\sigma)$. Note that property (ii) requires each function K_σ to have a Fourier transform which is 0 in a neighborhood of the origin, and as σ increases, so does this neighborhood. The convolution $\Phi * K_\sigma$ will retain this property and allow us to obtain diagonal dominance in A_σ.

Before moving on, we will need to determine certain bounds on the functions in $\mathcal{K}_1$. First we need an L^∞ bound.

$$\begin{aligned} |K_\sigma(x)| &\leq C_d \int_{\mathbb{R}^d} \hat{K}_\sigma(\omega) d\omega \\ &\leq C_d \sigma^d \int_1^\infty \kappa(t) t^{d-1} dt, \end{aligned}$$

so

$$|K_\sigma(x)| \leq C_{K,d} \sigma^d. \tag{2.2}$$

Next we will need a bound on K_σ for $r = ||x||_2 > 0$. Writing K_σ as a Fourier integral, we see

$$|K_\sigma(x)| = r^{-(d-2)/2} \left| \int_\sigma^\infty \kappa(t/\sigma) t^{d/2} J_{(d-2)/2}(rt) dt \right|,$$

and by a change of variables, we have

$$|K_\sigma(x)| = \sigma^{d/2+1} r^{-(d-2)/2} \left| \int_1^\infty \kappa(t) t^{d/2} J_{(d-2)/2}(\sigma r t) dt \right|.$$

Therefore by Proposition A.2

$$|K_\sigma(x)| \leq C_{K,d} \frac{\sigma^{d/2+1-l_d}}{r^{(d-2)/2+l_d}}. \tag{2.3}$$

Using (2.2) and (2.3), we will prove a bound of L^q norms of linear combinations of translates of K_σ.

Proposition 2.1. *Let* $T : \mathbb{R}^N \to L^1(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$ *be the linear operator defined by*

$$T(h) = \sum_{j=1}^{N} h_j K_\sigma(x - \xi_j).$$

Then

$$||T(h)||_q \leq C_{\kappa,d}\sigma^{d/q'} \left(1 + \frac{1}{(\sigma q)^{(d-2)/2+l_d}}\right)^{1/q'} ||h||_q$$

Proof. After proving the bound in the cases $q = 1$ and $q = \infty$, the result will then follow by the Riesz-Thorin theorem (cf. [21, Chapter 5]). First, by Proposition B.1 we have,

$$\begin{aligned}\sum_{j=1}^{N} |K_\sigma(x - \xi_j)| &\leq ||K_\sigma||_\infty + \sum_{||x-\xi_j|| \geq q} |K_\sigma(x - \xi_j)| \\ &\leq C_{K,d}\left(\sigma^d + \frac{\sigma^{d/2+1-l_d}}{q^{(d-2)/2+l_d}}\right).\end{aligned}$$

Simplifying the previous expression, we obtain

$$\sum_{j=1}^{N} |K_\sigma(x - \xi_j)| \leq C_{K,d}\sigma^d \left(1 + \frac{1}{(\sigma q)^{(d-2)/2+l_d}}\right). \tag{2.4}$$

Therefore

$$\begin{aligned}||T(h)||_\infty &= \left|\left|\sum_{j=1}^{N} h_j K_\sigma(x - \xi_j)\right|\right|_\infty \\ &\leq \left|\left|\sum_{j=1}^{N} |K_\sigma(x - \xi_j)|\right|\right|_\infty ||h||_\infty \\ &\leq C_{\kappa,d}\sigma^d \left(1 + \frac{1}{(\sigma q)^{(d-2)/2+l_d}}\right) ||h||_\infty\end{aligned}$$

Now in the case $q = 1$,

$$\begin{aligned} ||T(h)||_1 &= \left|\left|\sum_{j=1}^{N} h_j K_\sigma(x-\xi_j)\right|\right|_1 \\ &\leq \sum_{j=1}^{N} |h_j| \, ||K_\sigma||_1 \\ &\leq C_K \, ||h||_1 \end{aligned}$$

□

2.2.2 Interpolation Matrices

The interpolation matrices $(A_\sigma)_{i,j} = K_\sigma * \Phi(\xi_i - \xi_j)$ will be shown to be invertible by the following lemma. In addition, we will at the same time find a bound for the ℓ^p norm of A_σ^{-1}. Given an $n \times n$ matrix A with diagonal part D and $F = A - D$ we have the following.

Lemma 2.2. ([16, Lemma 5.2]) *If D is invertible and $||D^{-1}F||_1 < 1$, then A is invertible and $||A^{-1}||_1 < ||D^{-1}||_1 \, (1 - ||D^{-1}F||_1)^{-1}$.*

The diagonal entries of A_σ are equal to $K_\sigma * \Phi(0)$, and the off diagonal absolute column sums are of the form $\sum_{i \neq j} |K_\sigma * \Phi(\xi_i - \xi_j)|$. In order to apply the lemma, we must bound the former from below and the latter from above. First we have

$$\begin{aligned} K_\sigma * \Phi(0) &= C_d \int_\sigma^\infty \kappa(t/\sigma)\phi(t)t^{d-1}dt \\ &= C_d \sigma^{d-\beta} \int_1^\infty \frac{\kappa(t)}{t^{\beta-d+1}}(\sigma t)^\beta \phi(\sigma t)dt \\ &\geq C_{\Phi,d}\sigma^{d-\beta} \int_1^\infty \frac{\kappa(t)}{t^{\beta-d+1}}dt \end{aligned}$$

,

and hence

$$K_\sigma * \Phi(0) \geq C_{\Phi,K,d}\sigma^{d-\beta}. \tag{2.5}$$

Next, we need a bound on $|K_\sigma * \Phi(x)|$ for $x \neq 0$. Since $K_\sigma * \Phi$ has a radial Fourier transform in $L_1(\mathbb{R}^d)$, we can write it as a one dimensional integral. Note that in the following integral $r = ||x||_2$.

$$\begin{aligned} |K_\sigma * \Phi(x)| &= C_d r^{-(d-2)/2} \left| \int_\sigma^\infty \kappa(t/\sigma)\phi(t)t^{d/2} J_{(d-2)/2}(rt)dt \right| \\ &= C_d \frac{\sigma^{(d+2)/2-\beta}}{r^{(d-2)/2}} \left| \int_1^\infty \frac{\kappa(t)}{t^\beta}(\sigma t)^\beta \phi(\sigma t)t^{d/2} J_{(d-2)/2}(\sigma rt)dt \right|, \end{aligned}$$

and by Proposition A.2

$$|K_\sigma * \Phi(x)| \leq C_{\Phi,K,d} \frac{\sigma^{(d+2)/2-\beta}}{r^{(d-2)/2}(\sigma r)^{l_d}}$$

With this estimate we can bound the off diagonal absolute column sums of A_σ. Using Proposition B.1 we have

$$\sum_{i \neq j} |K_\sigma * \Phi(\xi_i - \xi_j)| \leq C_{\Phi,K,d} \frac{\sigma^{d-\beta}}{(\sigma q_X)^{(d-2)/2+l_d}}. \tag{2.6}$$

We are now ready to apply the lemma. Define

$$M = \max\left\{1, \left(\frac{2C^2_{\Phi,K,d}}{C^1_{\Phi,K,d}}\right)^{1/((d-2)/2+l_d)}\right\},$$

where the constants $C^1_{\Phi,K,d}$ and $C^2_{\Phi,K,d}$ are from (2.5) and (2.6) respectively. We then define $\sigma_0 = M/q_X$, so

$$(K_{\sigma_0} * \Phi(0))^{-1} \sum_{i \neq j} |K_{\sigma_0} * \Phi(\xi_i - \xi_j)| \leq \frac{1}{2}.$$

Therefore

$$\left|\left|A^{-1}_{\sigma_0}\right|\right|_1 \leq C_{\Phi,K,d}\sigma_0^{\beta-d},$$

and in terms of q_X,

$$\left|\left|A_{\sigma_0}^{-1}\right|\right|_1 \leq C_{\Phi,K,d} q_X^{d-\beta}. \tag{2.7}$$

As A_{σ_0} is self-adjoint the same bound holds for $\left|\left|A_{\sigma_0}^{-1}\right|\right|_\infty$. The Riesz-Thorin interpolation theorem can then be applied to get

$$\left|\left|A_{\sigma_0}^{-1}\right|\right|_p \leq C_{\Phi,K,d} q_X^{d-\beta} \tag{2.8}$$

for $1 \leq p \leq \infty$.

2.2.3 A Marcinkiewicz-Zygmund Type Inequality

To finish the bound of the stability ratio we require a bound of a discrete norm by a continuous one. To accomplish this, we can use an argument similar to the proof of [17, Theorem 1]. Let μ denote Lebesgue measure on $\mathbb{R}^d$ and let $\nu = \sum_{j=1}^N \delta_{\xi_j}$.

Proposition 2.3. *If $f \in L_\mu^1 \cap L_\mu^\infty$ and $K \in \mathcal{K}_1$, then*

$$||K_{\sigma_0} * f|_Y||_p \leq C_{K,d} q_X^{-d/p} ||f||_{p,\mu}$$

Proof. First, define $T(\tau, f) = \int_{\mathbb{R}^d} K_{\sigma_0}(\cdot - y) f(y) d\tau_y$. Then choose a compactly supported simple function H such that $||H||_{p',\nu} = 1$ and

$$||K_{\sigma_0} * f|_Y||_p = \sum_{j=1}^N H(\xi_j) K_{\sigma_0} * f(\xi_j).$$

Appealing to Fubini's theorem, we get

$$\begin{aligned}
||T(\mu,f)||_{1,\nu} &= \left|\int_{\mathbb{R}^d} T(\mu,f)(x)H(x)d\nu_x\right| \\
&= \left|\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} K_{\sigma_0}(x-y)f(y)H(x)d\mu_y d\nu_x\right| \\
&= \left|\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} K_{\sigma_0}(x-y)f(y)H(x)d\nu_x d\mu_y\right| \\
&= \left|\int_{\mathbb{R}^d} T(\nu,H)(y)f(y)d\mu_y\right|.
\end{aligned}$$

Now we may apply Hölders's inequality and Proposition 2.1 as follows:

$$\begin{aligned}
||T(\mu,f)||_{1,\nu} &\le ||T(\nu,H)||_{p',\mu}\,||f||_{p,\mu} \\
&= \left|\left|\sum_{j=1}^{N} K_{\sigma_0}(\cdot-\xi_j)H(\xi_j)\right|\right|_{p',\mu} ||f||_{p,\mu} \\
&\le C_{K,d}q_X^{-d/p}\,||f||_{p,\mu}.
\end{aligned}$$

□

2.2.4 Stability Ratio Bound

We are now in a position to prove the bound on the stabilty ratio for $p \in [1,\infty]$. Recall X is a countable subset of $\mathbb{R}^d$ with $0 < q_X < 1$, and Φ is an admissible function of order β.

Theorem 2.4. *Let $\mathcal{R}_{S,p}$ be the stability ratio associated with $S_X(\Phi)$. Then*

$$\mathcal{R}_{S,p} = \sup_{S_X(\Phi)\ni g\neq 0} \frac{||a||_p}{||g||_p} \le C_{\Phi,d}q_X^{d/p'-\beta}$$

Proof. It has been shown that the interpolation matrix $(A_{\sigma_0})_{i,j} = K_{\sigma_0} * \Phi(\xi_i - \xi_j)$ is invertible. Therefore $||a||_p \le \left|\left|A_{\sigma_0}^{-1}\right|\right|_p ||K_{\sigma_0} * g|_Y||_p$. Using (2.8), we get

$$||a||_p \le C_{\Phi,K,d}q_X^{d-\beta}\,||K_{\sigma_0} * g|_Y||_p\,.$$

Finally, applying the M-Z inequality and taking the infimum over $K \in \mathcal{K}_1$ gives the result. □

2.3 Band-Limited Approximation

As in the previous subsection, X will be a fixed countable set with $0 < q_X < 1$, and Φ will be an admissible function of order β. At this point, we are left with bounding the two remaining terms of (2.1). This will require choosing band-limited functions that approximate the elements of $S_X(\Phi)$ and satisfy the Bernstein inequality as well. In particular, given $g \in S_X(\Phi)$ we need to find a band-limited function g_σ so that

$$\frac{||g - g_\sigma||_{L^{k,p}}}{||a||_p} \leq C q_X^{\beta-k-d/p'}$$

for $1 \leq p \leq \infty$. Since most of the work will be done in the Fourier domain, we will require $k < \beta - d$.

2.3.1 Band-Limited Approximants

We begin by defining a class of band-limited functions. A function $g \in S_X(\Phi)$ will be convolved with one of these functions in order to define its band-limited approximant. Let $\mathcal{K}_2$ be the collection of Schwartz class functions $K : \mathbb{R}^d \to \mathbb{R}$ that satisfy the following properties:

(i) There is a nonincreasing $\kappa : [0, \infty) \to [0, \infty)$ such that $\hat{K}(\omega) = \kappa(||\omega||_2)$

(ii) $\kappa(\omega) = 1$ for $\omega \leq \frac{1}{2}$, and $\kappa(\omega) = 0$ for $\omega \geq 1$

For each $K \in \mathcal{K}_2$, we define the family of functions $\{K_\sigma\}_{\sigma \geq 1}$ by $\hat{K}_\sigma(\omega) = \hat{K}(\omega/\sigma)$. Band-limited approximants to $g \in S_X(\Phi)$ are then defined by $g_\sigma = K_\sigma * g$. The first thing we must check is that g_σ satisfies the Bernstein inequality. The following lemma addresses this issue.

Lemma 2.5. *Let* $f \in L^1(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$, *and let* $K \in \mathcal{K}_2$. *Then*

$$||f * K_\sigma||_{L^{m,p}} \leq C_{K,d}\sigma\, ||f * K_\sigma||_{L^{m-1,p}}$$

for $1 \leq p \leq \infty$ *and any positive integer* m.

Proof. We begin with an application of Young's inequality.

$$\begin{aligned} ||f * K_\sigma||_{L^{m,p}} &= \left|\left|\left[(1+||\cdot||_2^2)^{m/2}\hat{K}_\sigma \hat{f}\right]^\vee\right|\right|_p \\ &= \left|\left|\left[(1+||\cdot||_2^2)^{1/2}\hat{K}_{2\sigma}(1+||\cdot||_2^2)^{(m-1)/2}\hat{K}_\sigma \hat{f}\right]^\vee\right|\right|_p \\ &\leq \left|\left|\left[(1+||\cdot||_2^2)^{1/2}\hat{K}_{2\sigma}\right]^\vee\right|\right|_1 \left|\left|\left[(1+||\cdot||_2^2)^{(m-1)/2}\hat{K}_\sigma \hat{f}\right]^\vee\right|\right|_p \\ &= \left|\left|\left[(1+||\cdot||_2^2)^{1/2}\hat{K}_{2\sigma}\right]^\vee\right|\right|_1 ||f * K_\sigma||_{L^{m-1,p}}. \end{aligned}$$

Now it is known that there exist finite measures ν and λ such that

$$(1+||x||_2^2)^{1/2} = \hat{\nu}(x) + 2\pi\, ||x||_2\, \hat{\lambda}(x),$$

cf. [20, Chapter 5]. We therefore have

$$\begin{aligned} \left|\left|\left[(1+||\cdot||_2^2)^{1/2}\hat{K}_{2\sigma}\right]^\vee\right|\right|_1 &= \left|\left|\left[\hat{\nu}\hat{K}_{2\sigma} + 2\pi\, ||\cdot||_2\, \hat{\lambda}\hat{K}_{2\sigma}\right]^\vee\right|\right|_1 \\ &\leq ||\nu * K_{2\sigma}||_1 + 2\pi \left|\left|\lambda * \left[||\cdot||_2\, \hat{K}_{2\sigma}\right]^\vee\right|\right|_1 \\ &\leq C\left(1 + \left|\left|\left[||\cdot||_2\, \hat{K}_{2\sigma}\right]^\vee\right|\right|_1\right), \end{aligned}$$

and it remains to prove $\left\| \left[\|\cdot\|_2 \hat{K}_{2\sigma} \right]^\vee \right\|_1 \le C\sigma$. First

$$\begin{aligned} \left\| \left[\|\cdot\|_2 \hat{K}_{2\sigma} \right]^\vee \right\|_\infty &\le C_d \int_{\mathbb{R}^d} \|\omega\|_2 \hat{K}_{2\sigma}(\omega) d\omega \\ &= C_d \int_0^{2\sigma} t\kappa\left(\frac{t}{2\sigma}\right) t^{d-1} dt \\ &= C_d \sigma^{1+d} \int_0^2 \kappa\left(\frac{t}{2}\right) t^d dt, \end{aligned}$$

and hence

$$\left\| \left[\|\cdot\|_2 \hat{K}_{2\sigma} \right]^\vee \right\|_\infty \le C_{K,d} \sigma^{1+d}. \tag{2.9}$$

Now, for $\|x\|_2 = r > 0$

$$\begin{aligned} \left| \left[\|\cdot\|_2 \hat{K}_{2\sigma} \right]^\vee (x) \right| &= r^{-(d-2)/2} \left| \int_0^{2\sigma} t\kappa\left(\frac{t}{2\sigma}\right) t^{d/2} J_{(d-2)/2}(rt) dt \right| \\ &= r^{-(d-2)/2} \sigma^{d/2+2} \left| \int_0^2 t\kappa\left(\frac{t}{2}\right) t^{d/2} J_{(d-2)/2}(\sigma r t) dt \right|, \end{aligned}$$

so by Proposition A.2

$$\left| \left[\|\cdot\|_2 \hat{K}_{2\sigma} \right]^\vee (r) \right| \le C_{K,d} \frac{r^{-(d-2)/2} \sigma^{d/2+2}}{(\sigma r)^{l_d}}. \tag{2.10}$$

Utilizing inequalities 2.9 and 2.10, we get

$$\begin{aligned} \left\| \left[\|\cdot\|_2 \hat{K}_{2\sigma} \right]^\vee \right\|_1 &\le C_d \sigma^{-d} \left\| \left[\|\cdot\|_2 \hat{K}_{2\sigma} \right]^\vee \right\|_\infty + \int_{\|x\|_2 \ge \frac{1}{\sigma}} \left| \left[\|\cdot\|_2 \hat{K}_{2\sigma} \right]^\vee (x) \right| dx \\ &\le C_{K,d}\sigma + C_{K,d} \int_{\frac{1}{\sigma}}^\infty \frac{r^{-(d-2)/2} \sigma^{d/2+2}}{(\sigma r)^{l_d}} r^{d-1} dr \\ &\le C_{K,d}\sigma \end{aligned}$$

□

Now if we define $\sigma_1 := 1/q_X$, then the next corollary follows easily.

Corollary 2.6. *For all $g \in S_X(\Phi)$,*

$$||g * K_{\sigma_1}||_{L^{k,p}} \leq C_{\Phi,K,d} q_X^{-k} ||g||_p$$

2.3.2 Approximation Analysis

Now that we know the band-limited approximants to the elements of $S_X(\Phi)$ satisfy the Bernstein inequality, we must bound the error of approximation in $L^{k,p}$. We will begin by bounding the approximation error in $L^{k,1}$ and $L^{k,\infty}$ and then use interpolation to obtain the result for all other values of p. In both extremal cases, this reduces to bounding the error of approximating the RBF by band-limited functions. For $p = 1$ this is straightforward; however, the case $p = \infty$ is more involved.

In order to simplify some expressions, we define the functions

$$\begin{aligned} E_{\Phi,K,k} &:= \left| \left((1 + ||\cdot||_2^2)^{k/2} (\Phi - \Phi * K_{\sigma_1})^\wedge \right)^\vee \right|, \\ h(t) &:= \phi(t)(1+t^2)^{\beta/2}. \end{aligned}$$

If we are to bound the error of approximating Φ by band-limited functions, we will certainly need a point-wise bound of $E_{\Phi,K,k}$. Let us begin with an L^∞ bound.

$$\begin{aligned} E_{\Phi,K,k}(x) &\leq \int_{\mathbb{R}^d} (1 - \hat{K}_{\sigma_1}(\omega)) \hat{\Phi}(\omega) (1 + ||\omega||_2^2)^{k/2} d\omega \\ &\leq C_d {\sigma_1}^{d-\beta+k} \int_{1/2}^\infty \frac{(1-\kappa(t)) h(\sigma_1 t) t^{d-1}}{(1/{\sigma_1}^2 + t^2)^{(\beta-k)/2}} dt. \end{aligned}$$

Therefore

$$E_{\Phi,K,k}(x) \leq C_{\beta,K,d} {\sigma_1}^{d-\beta+k}. \tag{2.11}$$

Next, for $||x||_2 = r > 0$

$$\begin{aligned} E_{\Phi,K,k}(x) &= r^{-(d-2)/2}\left|\int_{\sigma_1/2}^{\infty} \frac{(1-\kappa(t/\sigma_1))h(t)}{(1+t^2)^{(\beta-k)/2}} t^{d/2} J_{(d-2)/2}(rt)dt\right| \\ &= \frac{{\sigma_1}^{d/2+1-\beta+k}}{r^{(d-2)/2}}\left|\int_{1/2}^{\infty} \frac{(1-\kappa(t))h(\sigma_1 t)}{(1/{\sigma_1}^2+t^2)^{(\beta-k)/2}} t^{d/2} J_{(d-2)/2}(\sigma_1 rt)dt\right|. \end{aligned}$$

Therefore by Proposition A.2

$$E_{\Phi,K,k}(x) \leq C_{\Phi,K,d}\frac{r^{-(d-2)/2}{\sigma_1}^{d/2+1-\beta+k}}{(\sigma_1 r)^{l_d}} \tag{2.12}$$

With these results we are now able to bound the error in approximating $S_X(\Phi)$. Let $Y = \{\xi_i\}_{i=1}^N$ be a finite subset of X, and let $K \in \mathcal{K}_2$

Theorem 2.7. *Given $g = \sum_{j=1}^N a_j\Phi(\cdot - \xi_j) \in S_X(\Phi)$, we have*

$$||g - g * K_{\sigma_1}||_{L^{k,p}} \leq C_{\Phi,d} q_X^{\beta-k-d/p'} ||a||_p$$

for $1 \leq p \leq \infty$.

Proof. We will show this holds when $p = 1$ and $p = \infty$, and the result will follow from the Riesz-Thorin interpolation theorem. Letting $g_{\sigma_1} = g * K_{\sigma_1}$, we have

$$\begin{aligned} ||g - g_{\sigma_1}||_{L^{k,\infty}} &= \left|\left|\sum_{j=1}^N a_j\Phi(\cdot - \xi_j) - \sum_{j=1}^N a_j\Phi * K_{\sigma_1}(\cdot - \xi_j)\right|\right|_{L^{k,\infty}} \\ &= \left|\left|\sum_{j=1}^N a_j \left((1+||\cdot||_2^2)^{k/2}(\Phi - \Phi * K_{\sigma_1})^\wedge\right)^\vee (\cdot - \xi_j)\right|\right|_\infty, \end{aligned}$$

and using the notation for $E_{\Phi,K,k}$ above

$$||g - g_{\sigma_1}||_{L^{k,\infty}} \leq ||a||_\infty \left|\left|\sum_{j=1}^N E_{\Phi,K,k}(\cdot - \xi_j)\right|\right|_\infty$$

Now by (2.11), (2.12), and Proposition B.1 we have

$$\begin{aligned}\sum_{j=1}^{N} E_{\Phi,K,k}(x-\xi_j) &\le ||E_{\Phi,K,k}||_\infty + \sum_{||x-\xi_j||_2 \ge q_Y} E_{\Phi,K,k}(x-\xi_j)\\ &\le C_{\Phi,K,d}{\sigma_1}^{d-\beta+k}\left(1+\frac{1}{(\sigma_1 q_X)^{(d-2)/2+l_d}}\right).\end{aligned}$$

For $p=1$, we have $||g-g_{\sigma_1}||_{L^{k,1}} \le ||a||_1 \, ||E_{\Phi,K,k}||_1$, and

$$\begin{aligned}||E_{\Phi,K,k}||_1 &\le C_d q_X^d \, ||E_{\Phi,K,k}||_\infty + \int_{||x||_2\ge q_X} E_{\Phi,K,k}(x)dx\\ &\le C_{\Phi,K,d}{\sigma_1}^{k-\beta} + C_{\Phi,K,d}\int_q^\infty \frac{{\sigma_1}^{d/2+1-\beta+k-l_d}}{r^{(d-2)/2+l_d-d+1}}dr\\ &\le C_{\Phi,K,d}{\sigma_1}^{k-\beta}\end{aligned}$$

We finish the proof by taking the infimum over $K \in \mathcal{K}_2$. □

2.4 Bernstein Inequalities and Inverse Theorems

In approximation theory, there are a variety of applications for Bernstein inequalities. While they are most commonly associated with the derivation of inverse theorems, they can also be useful in proving direct theorems. For example, a Bernstein inequality for multivariate polynomials is used in certain RBF approximation error estimates, cf. [23, Chapter 11]. However, in this paper, we will only address the Bernstein inequalities themselves and their matching inverse theorems.

With the bound of the stability ratio and the band-limited approximation estimate in hand, we are in a position to prove the Bernstein inequalities. Let X be a countable set with $0 < q_X < 1$, and let Φ be an admissible function of order β.

Theorem 2.8. *If $k < \beta - d$, $1 \leq p \leq \infty$, and $g \in S_X(\Phi)$, then*

$$||g||_{L^{k,p}} \leq C_{\Phi,d} q_X^{-k} ||g||_p$$

Proof. Let g_{σ_1} be the previously defined approximant of g. Then

$$\begin{aligned} ||g||_{L^{k,p}} &\leq ||g_{\sigma_1}||_{L^{k,p}} + ||g - g_{\sigma_1}||_{L^{k,p}} \\ &= ||g_{\sigma_1}||_{L^{k,p}} + \left(\frac{||a||_p}{||g||_p} \frac{||g - g_{\sigma_1}||_{L^{k,p}}}{||a||_p} \right) ||g||_p \end{aligned}$$

Applying Theorem 2.4, Corollary 2.6, and Theorem 2.7, we get

$$\begin{aligned} ||g||_{L^{k,p}} &\leq ||g_{\sigma_1}||_{L^{k,p}} + \left(\frac{||a||_p}{||g||_p} \frac{||g - g_{\sigma_1}||_{L^{k,p}}}{||a||_p} \right) ||g||_p \\ &\leq C_{\Phi,d} q_X^{-k} ||g||_p + \left(C_{\Phi,d} q_X^{\frac{d}{p'} - \beta} \right) \left(C_{\Phi,d} q_X^{\beta - k - \frac{d}{p'}} \right) ||g||_p \\ &\leq C_{\Phi,d} q_X^{-k} ||g||_p . \end{aligned}$$

□

Having established Bernstein inequalities for $S_X(\Phi)$, we can now prove the corresponding inverse theorems. In what follows, X_n will denote a sequence of countable sets in $\mathbb{R}^d$ such that $X_n \subset X_{n+1}$, $\rho_{X_n} = \rho_n \leq C$, and $0 < h_n, q_n < 2^{-n}$, and Φ will be an admissible RBF of order β. We will additionally use the notation $S_n = S_{X_n}(\Phi)$. In this situation, we define the error of approximation by

$$\mathcal{E}(f, S_n)_{L_p} = \inf_{g \in S_n} ||f - g||_{L_p(\mathbb{R}^d)} .$$

Using the standard technique for proving inverse theorems from Bernstein inequalities, we will show that if a function is well approximated by S_n then it must lie in some Bessel-potential space.

Theorem 2.9. *Let $1 \leq p \leq \infty$, and let $f \in L^p(\mathbb{R}^d)$. If there is a constant $c_f > 0$, independent of n, and a positive integer l such that*

$$\mathcal{E}(f, S_n)_{L_p} \leq c_f h_n^l,$$

then for every $0 \leq k < \min\{\beta - d, l\}$, $f \in L^{k,p}$.

Proof. Let $f_n \in S_n$ be a sequence of functions satisfying $||f - f_n||_p \leq 2c_f h_n^l$. Note that $f_n \in S_m$ for $m > n$. We now have

$$\begin{aligned} ||f_{n+1} - f_n||_{L^{k,p}} &\leq C_{\Phi,d} \left(\frac{h_{n+1}}{q_{n+1}} \right)^k h_{n+1}^{-k} ||f_{n+1} - f_n||_p \\ &\leq C_{\Phi,d} h_{n+1}^{-k} (||f_{n+1} - f||_p + ||f - f_n||_p) \\ &\leq C_{\Phi,d,f} h_{n+1}^{-k} (h_{n+1}^l + h_n^l), \end{aligned}$$

and since $h_n < 2^{-n}$, it follows that

$$||f_{n+1} - f_n||_{L^{k,p}} \leq C_{\Phi,d,f} 2^{-(l-k)n}.$$

This shows f_n is a Cauchy sequence in $L^{k,p}$. Since $L^{k,p}$ is complete, f_n converges to some function $h \in L^{k,p}$. Since f_n converges to both f and h in L^p, $f = h$ a.e., and therefore $f \in L^{k,p}$. □

3. APPROXIMATION ERROR ESTIMATES

Now that we have established Bernstein inequalities and inverse theorems for RBF approximants, we turn to the issue of direct theorems. Our approach will be to extend a recent result of DeVore and Ron. In the paper [7], the authors made use of a Green's function-type condition in order to determine rates of approximation. They were concerned with approximating functions in $L^p(\mathbb{R}^d)$ by approximation spaces $S_X(\Phi)$, where X has no accumulation points and h_X is finite. The essence of their argument is as follows.

Property 1. *Suppose* $T : C_c^k(\mathbb{R}^d) \to C_c(\mathbb{R}^d)$ *is a linear operator,* $\Phi \in L^1_{loc}(\mathbb{R}^d)$*, and for any* $f \in C_c^k(\mathbb{R}^d)$ *we have*

$$f = \int_{\mathbb{R}^d} Tf(t)\Phi(\cdot - t)dt. \tag{3.1}$$

Now, given an $f \in C_c^k(\mathbb{R}^d)$, we can form an approximant by replacing Φ with a suitable kernel in (3.1). Since we are interested in approximating f by $S_X(\Phi)$, the kernel should have the form

$$K(\cdot, t) = \sum_{\xi \in X(t)} A(t, \xi)\Phi(\cdot - \xi). \tag{3.2}$$

The collection of possible kernels is restricted by requiring $X(t) \subset B(t, Ch_X) \cap X$ for a constant $C > 0$ and additionally requiring $A(t, \xi)$ to be in L^1 for all ξ. An essential ingredient for deriving approximation rates is showing that Φ can be well approximated by K. We therefore define the error kernel

$$E(x, t) := \Phi(x - t) - K(x, t). \tag{3.3}$$

and assume the following property.

Property 2. *Given Φ, there exists a kernel K of the form (3.2) and constants $l > d$, $\kappa > 0$, and $C > 0$ such that*

$$|E(x,t)| \leq Ch_X^{\kappa}(1 + ||x - t||_2)^{-l}.$$

While this approach can be used to provide estimates for some popular RBFs, e.g. the thin-plate splines, not all RBFs are Green's functions. Therefore our goal will be to extend the class of applicable RBFs. Note that in [7], the authors were concerned with proving error bounds that account for the local density of the data sites. As our goal is to extend the class of applicable RBFs, we will not discuss this more technical approach; instead we will assume X is quasi-uniform, and our bound will be written in terms of the global density parameter h_X.

3.1 General Error Estimates

In this subsection, we will show that it is possible to replace Property 1 by a condition that only requires Φ to be "close" to a function that satisfies this property. Specifically, we will consider RBFs Φ that satisfy the following.

Property 1'. *Let $T : C_c^k(\mathbb{R}^d) \to C_c(\mathbb{R}^d)$ and $G \in L^1_{loc}(\mathbb{R}^d)$ be a pair satisfying Property 1, and suppose $G = \Phi * \mu_n + \nu_n$ where $\Phi \in L^1_{loc}(\mathbb{R}^d)$, μ_n is a sequence of compactly supported, finite Borel measures with $||\mu_n||$ bounded by a constant, and ν_n is a sequence of finite Borel measures with $||\nu_n||$ converging to 0.*

The next theorem provides an error bound for approximation by $S_X(\Phi)$ where X is quasi-uniform. For $1 \leq p \leq \infty$, we define the error of approximation by

$$\mathcal{E}(f, S_X(\Phi))_p = \inf_{S \in S_X(\Phi)} ||f - S||_{L^p(\mathbb{R}^d)}$$

Theorem 3.1. *Suppose* Φ *is an RBF satisfying Property 1' and Property 2, and let* $f \in C_c^k(\mathbb{R}^d)$, *then*

$$\mathcal{E}(f, S_X(\Phi))_p \leq Ch_X^{\kappa} \, |f|_{W(L^p(\mathbb{R}^d),T)}$$

Proof. First, we define a sequence of approximants to f by

$$\begin{aligned} F_n &:= \int_{\mathbb{R}^d} Tf * \mu_n(t) K(\cdot, t) dt \\ &= \sum_{\xi \in X} \Phi(\cdot - \xi) \int_{\mathbb{R}^d} Tf * \mu_n(t) A(t, \xi) dt. \end{aligned}$$

Note that this sum is finite due to the compact support of $Tf * \mu_n$ and the conditions imposed on $A(\cdot, \cdot)$. Now

$$\begin{aligned} |f - F_n| &= \left| \int_{\mathbb{R}^d} Tf(t) G(\cdot - t) dt - \int_{\mathbb{R}^d} Tf * \mu_n(t) K(\cdot, t) dt \right| \\ &= \left| \int_{\mathbb{R}^d} Tf * \mu_n(t) \left(\Phi(\cdot - t) - K(\cdot, t) \right) dt + Tf * \nu_n \right| \\ &\leq \int_{\mathbb{R}^d} |Tf * \mu_n(t)| \, |\Phi(\cdot - t) - K(\cdot, t)| \, dt + |Tf * \nu_n| \,, \end{aligned}$$

and by Property 2,

$$|f - F_n| \leq Ch_X^k \int_{\mathbb{R}^d} |Tf * \mu_n(t)| \, (1 + ||\cdot - t||_2)^{-l} dt + |Tf * \nu_n| \,.$$

Therefore

$$\begin{aligned} ||f - F_n||_p &\leq Ch_X^k \left| \left| |Tf * \mu_n| * (1 + ||\cdot||_2)^{-l} \right| \right|_p + ||Tf * \nu_n||_p \\ &\leq Ch_X^k \, ||Tf||_p \, ||\mu_n|| + ||Tf||_p \, ||\nu_n|| \end{aligned}$$

by Proposition 1.2. Hence, for some N sufficiently large, we will have

$$||f - F_N||_p \leq Ch_X^k \, ||Tf||_p \,,$$

where C depends only on Φ, G, and the constant from Property 2. □

Now that we have established this result for C_c^k, we would like to extend it to all of $W(L^p(\mathbb{R}^d),T)$. This is accomplished in the following corollary using a density argument.

Corollary 3.2. *Suppose $f \in W(L^p(\mathbb{R}^d),T)$, then*

$$\mathcal{E}(f,S_X(\Phi))_p \leq Ch_X^\kappa \, |f|_{W(L^p(\mathbb{R}^d),T)}$$

Proof. Let $f_n \in C_c^\infty$ be a sequence converging to f in $W(L^p(\mathbb{R}^d),T)$, and let $F_n \in S_X(\Phi)$ be a sequence of approximants satisfying

$$||f_n - F_n||_p \leq Ch_X^\kappa \, |f_n|_{W(L^p(\mathbb{R}^d),T)} .$$

Then

$$\begin{aligned} ||f - F_n||_p &\leq ||f - f_n||_p + ||f_n - F_n||_p \\ &\leq ||f - f_n||_p + Ch_X^\kappa \, |f_n|_{W(L^p(\mathbb{R}^d),T)} . \end{aligned}$$

Now since f_n converges to f in $W(L^p(\mathbb{R}^d),T)$, the result follows. □

3.2 Special Cases

We will now give some examples showing how to verify the properties listed above. One goal of this theory is to derive approximation results for the Wendland functions, which are a class of compactly supported RBFs. We will denote by $\Phi_{d,k}$, the Wendland function in $C^{2k}(\mathbb{R}^d)$ of minimal degree. In this section, we will always assume that X is a quasi-uniform subset of $\mathbb{R}^d$.

First we will show that Property 2 holds for the Wendland functions when $k \geq 1$. Following the proof for the thin-plate splines presented in [7], we will make use of local polynomial reproduction. To that end, let P denote the space of polynomials of degree at most $2k-1$ on $\mathbb{R}^d$. Now for a finite set $Y \subset \mathbb{R}^d$, let Λ_Y be the set of extensions to $C(\mathbb{R}^d)$ of linear combinations of the point evalution functionals $\delta_y : P|_Y \to \mathbb{R}$. We will assume the following:

(i) There is a constant C_1 such that $X(t) \subset B(t, C_1 h_X) \cap X$ for all $t \in \mathbb{R}^d$.

(ii) For all t, there exists $\lambda_t \in \Lambda_{X(t)}$ such that λ_t agrees with δ_t on P.

(iii) $||\lambda_t|| \leq C_2$ for some constant C_2 independent of t.

Based on these assumptions, λ_t takes the form $\sum_{\xi \in X(t)} A(\xi, t)\delta_\xi$, and we can define the kernel approximant to $\Phi_{d,k}$ by

$$\begin{aligned} K(x,t) &= \lambda_t(\Phi_{d,k}(x-\cdot)) \\ &= \sum_{\xi \in X(t)} A(t,\xi)\Phi(x-\xi). \end{aligned}$$

One way to verify the validity of our assumptions is the following. For each $x \in h_X\mathbb{Z}^d$, we denote by Q_x the cube of side length h_X centered at x. This provides a partition of $\mathbb{R}^d$ into cubes. We then associate to each cube Q_x the ball $B(x, C_3 h_X)$ for some constant C_3, and for each t in a fixed Q_{x_0}, we define $X(t) = X \cap B(x_0, C_3 h_X)$. By choosing C_3 appropriately and bounding the possible values of h_X, one can show that there exists λ_t satisfying the above properties, with $C_2 = 2$, for all $t \in Q_{x_0}$, cf. [23, Chapter 3]. As x_0 was arbitrary, the result holds for all $t \in \mathbb{R}^d$. Note that by using the same set $X(t)$ for all t in a given cube Q_{x_0}, we are able to choose the coefficients $A(\xi, t)$ so that they are continuous with respect to t in the interior of the cube. To see this, let $\{p_i\}_{i=1}^m$ be a basis for P and let $B(x_0, C_3 h_X) \cap X = \{\xi_j\}_{j=1}^N$. We now define the matrix $M_{i,j} := p_i(\xi_j)$ and the t dependent vector $\beta(t)_i := p_i(t)$. Since λ_t reproduces polynomials, there exists a vector $\alpha(t)$ such that $M\alpha(t) = \beta(t)$ for all $t \in Q_{x_0}$. We could therefore choose a specific $\alpha(t)$ by means of a pseudo-inverse, i.e.

$\alpha(t) := M^T(MM^T)^{-1}\beta(t)$. In this form it is clear that α is a continuous function of t and $\lambda_t := \sum_{i=1}^{N} \alpha_i(t)\delta_{\xi_i}$ satisfies the required properties.

We will now use the polynomial reproducing functionals λ_t to verify Property 2. Note that $\Phi_{d,k}(x-t)$ and $\lambda_t(\Phi_{d,k}(x-\cdot))$ are both zero for $||x-t||_2 > 1 + C_1 h_X$. Therefore in order to verify Property 2, it suffices to show that for $||x-t||_2 < 1+C_1 h_X$ we have $|\Phi_{d,k}(x-t) - \lambda_0(\Phi_{d,k}(x-\cdot))| \le Ch_X^{2k}$. This inequality will only be verified for $t = 0$, as all other cases work similarly. To begin, fix x, and let R be the $2k-1$ degree Taylor polynomial of $\Phi_{d,k}$ at x. Then

$$\begin{aligned}
|\Phi_{d,k}(x) - \lambda_0(\Phi_{d,k}(x-\cdot))| &= |\Phi_{d,k}(x) - R(x) - \lambda_0(\Phi_{d,k}(x-\cdot) - R(x-\cdot))| \\
&\le ||\lambda_0||\, ||\Phi_{d,k}(x-\cdot) - R(x-\cdot)||_{L^\infty(B(0,C_1 h_X))} \\
&= ||\lambda_0||\, ||\Phi_{d,k} - R||_{L^\infty(B(x,C_1 h_X))} \\
&\le C_{\Phi_{d,k}}\, ||\lambda_0||\, ||\Phi_{d,k}||_{W^{2k,\infty}}\, h_X^{2k}
\end{aligned}$$

We have therefore shown that each $\Phi_{d,k}$ satisfies Property 2 with $\kappa = 2k$.

In order to prove Property 1', we can work in the Fourier domain, where the convolution becomes a standard product. For example, to show that $G = \Phi * \mu$ for some $\mu \in M(\mathbb{R}^d)$, we can verify that $\hat{G}/\hat{\Phi}$ is the Fourier transform of some $\mu \in M(\mathbb{R}^d)$. The difficulty lies in characterizing the space of Fourier transforms of $M(\mathbb{R}^d)$. This is known to be a very difficult problem, cf. [3]. However, in certain situations, we are able to make this determination. Suppose that we know $\hat{\Phi}/\hat{G}$ is the Fourier transform of some $\mu \in M(\mathbb{R})$. Then the following result will give conditions for $\hat{\mu}^{-1} = \hat{G}/\hat{\Phi}$ being the Fourier transform of an element of $M(\mathbb{R})$. Recall the decomposition of measures from Section 1; a Borel measure μ can be written as $\mu_a + \mu_s + \mu_d$. Using this notation, we state the following theorem of Benedetto. Note that the author proved this theorem in a more general setting with $\mathbb{R}$ replaced by an arbitrary locally compact abelian group.

Theorem 3.3. ([3, Theorem 2.4.4]) *Let $\mu \in M(\mathbb{R})$ such that $|\hat{\mu}|$ never vanishes and*

$$||\mu_s|| < \inf_{x\in\mathbb{R}} |\hat{\mu}_d(x)| .$$

Then $\hat{\mu}^{-1}$ is the Fourier transform of an element of $M(\mathbb{R})$

With this theorem, we will be able to verify Property 1' for the Wendland functions on $\mathbb{R}$. At this point we are only able to handle the 1-dimensional case; however, we would like to point out that the authors of [16] were able to show that when restricted to $\mathbb{S}^n$, all of the Wendland functions are perturbations of Green's functions.

We now fix $k \geq 1$, and we must select a Green's function G and corresponding differential operator T so that $\hat{\Phi}_{1,k}/\hat{G}$ satisfies the necessary conditions. Considering the decay of $\hat{\Phi}_{1,k}$, we choose $\hat{G} = (1 + |\cdot|^{2k+2})^{-1}$ and $T = (1 + (-1)^{k+1}\Delta^{k+1})$. Note that for $1 \leq p < \infty$ we have $W(L^p(\mathbb{R}), T)$ is equivalent to the spaces $W^{2k+2,p}(\mathbb{R})$ and $L^{2k+2,p}(\mathbb{R})$. We now verify the hypotheses of Theorem 3.3. By Proposition C.2, we know that for some $B_k > 0$ and $C \in \mathbb{R}$ we have

$$\frac{\hat{\Phi}_{1,k}}{\hat{G}}(x) = \hat{\Phi}_{1,k} + B_k \left(\frac{1}{k!} + \frac{(-1)^{k+1}}{k!2^k} \cos(x) - C\frac{\sin(x)}{x} + \hat{h}(x) \right)$$

where $\hat{h}$ is the Fourier transform of an L^1 function, h. The fact that $\hat{\Phi}_{1,k}/\hat{G}$ is positive follows from the positivity of $\hat{\Phi}_{1,k}$, cf. [23, Chapter 10]. Now let μ be the measure defined by

$$\mu = \Phi_{1,k} + \sqrt{2\pi}B_k \left(\frac{1}{k!}\delta_0 + \frac{(-1)^{k+1}}{k!2^{k+1}}(\delta_{-1} + \delta_1) - \frac{C}{2}\chi_{[-1,1]} + \frac{1}{\sqrt{2\pi}}h \right).$$

Then μ is a finite Borel measure with $\hat{\mu} = \hat{\Phi}_{1,k}/\hat{G}$. Additionally, μ has no singular continuous part, and

$$|\hat{\mu}_d(\omega)| = B_k \left(\frac{1}{k!} + \frac{(-1)^{k+1}}{k!2^k} \cos(\omega) \right) \geq \frac{B_k}{2k!}$$

Therefore Theorem 3.3 implies that $1/\hat{\mu} = \hat{G}/\hat{\Phi}_{1,k}$ is the Fourier transform of some finite Borel measure $\tilde{\mu}$. By letting $\tilde{\mu}_n$ be the restriction of $\tilde{\mu}$ to $B(0,n)$, we can see that

$$\begin{aligned} G &= \Phi_{1,k} * \tilde{\mu} \\ &= \Phi_{1,k} * \tilde{\mu}_n + \Phi_{1,k} * (\tilde{\mu} - \tilde{\mu}_n), \end{aligned}$$

and hence corollary 3.2 applies.

Theorem 3.4. *Let $\Phi_{1,k}$ be as above, and let X be a quasi-uniform subset of $\mathbb{R}$. Then for $f \in W^{2k+2,p}(\mathbb{R})$ and $1 \leq p < \infty$, we have*

$$\mathcal{E}(f, S_X(\Phi_{1,k}))_p \leq C h_X^{2k} \, |f|_{W^{2k+2,p}(\mathbb{R})}$$

By examining this example, we can determine what the analogous result would be for the remaining Wendland functions. If $k \geq 1$, then it is known that $\hat{\Phi}_{d,k} > 0$ and

$$c_1(1 + ||\omega||_2^2)^{-(d+2k+1)/2} \leq \hat{\Phi}_{d,k}(\omega) \leq c_2(1 + ||\omega||_2^2)^{-(d+2k+1)/2}$$

for some positive constants c_1 and c_2, cf. [23, Theorem 10.35]. Therefore, we could choose G to be the Sobolev spline of order $d + 2k + 1$ (i.e. $\hat{G}(\omega) = (1 + ||\omega||_2^2)^{-(d+2k+1)/2}$), which is the Green's function for the pseudo-differential operator $T = (1 - \Delta)^{(d+2k+1)/2}$, so that $\hat{G}$ and $\hat{\Phi}_{d,k}$ have similar decay. In this situation, $\hat{G}/\hat{\Phi}_{d,k}$ is a continous function that is bounded above and bounded away from 0. If we assume $\hat{G}/\hat{\Phi}_{d,k}$ is the Fourier transform of an element of $M(\mathbb{R}^d)$, then we could apply corollary 3.2 to obtain the following.

Conjecture 3.5. *Let $\Phi_{d,k}$ be as above with $k \geq 1$, and let X be a quasi-uniform subset of $\mathbb{R}^d$. Then for $f \in L^{d+2k+1,p}(\mathbb{R}^d)$ and $1 \leq p < \infty$, we have*

$$\mathcal{E}(f, S_X(\Phi_{d,k}))_p \leq C h_X^{2k} \left| f \right|_{L^{d+2k+1,p}(\mathbb{R}^d)} .$$

4. CONCLUSIONS AND FUTURE RESEARCH

In this paper, we have proved results concerning the approximation of functions in $L^p(\mathbb{R}^d)$ by RBF approximation spaces. Using the stability of the approximants and band-limited approximation, we were able to show that functions in $S_X(\Phi)$ satisfy a Bernstein inequality if Φ has an algebraically decaying (generalized) Fourier transform. Making use of these inequalities, we were able to prove corresponding inverse theorems. Specifically, we showed that if a function can be approximated well by $S_{X_n}(\Phi)$, where h_{X_n} converges to 0, then that function must lie in a Bessel-potential space.

While the Bernstein inequalities contained in this paper used norms defined on $\mathbb{R}^d$, there are many instances for which one is interested in approximation on a bounded domain. In [19], Schaback and Wendland proved a Bernstein-type inequality for bounded domians; however, it was restricted to a particular L^2 Sobolev norm of an approximant being bounded by its L^∞ norm. It would be interesting to know if one could prove a Bernstein or Markov type inequality for RBF approximants on a bounded domain $\Omega \subset \mathbb{R}^d$, one in which an L^p Sobolev norm is bounded by the corresponding L^p norm.

Another avenue for future research in this area concerns the choice of RBF. Our method of proving the Bernstein inequalities relied on the algebraic decay of the Fourier transforms of the RBFs involved. Is it possible to modify this proof so that we can include RBFs with exponentially decaying Fourier transforms, e.g. the Gaussians and inverse multiquadrics? In the case of the Gaussians, Mhaskar has shown that a Bernstein inequality does hold for RBF approximants if one bounds the number of centers in X, [15].

In Section 3, we proved a direct theorem concerning approximation by perturbations of Green's functions. To accomlish this, we used the fact that Green's functions provide a way to invert pseudo-differential operators. While we were able to verify the hypotheses for some specific Wendland functions, we could not show that the result holds in general. One direction for future research is to show that all of the Wendland functions satisfy the necessary conditions.

Prior to the work [7] of Devore and Ron, there were results for L^p approximation by Radial Basis functions on $\mathbb{R}^d$, [13, 24]. One improvement of the Devore-Ron paper is that the error bounds account for the local density of the data. Previous work had only dealt with bounding the error in terms of a global density parameter, as we did in Section 3. Since we have adapted the method of [7], it seems reasonable that we should be able to prove a result about approximating by perturbations of Green's functions where we account for the local density of the data sites.

REFERENCES

[1] N. Aronszajn, K.T. Smith, Theory of Bessel potentials. I, Ann. Inst. Fourier (Grenoble) 11 (1961) 385-475.

[2] N. Aronszajn, P. Szeptycki, On spaces of potentials connected with L^p classes, Ann. Inst. Fourier (Grenoble) 13 (1963) 211-306.

[3] J.J. Benedetto, Spectral Synthesis, Academic Press, Inc., London, 1975.

[4] S.N. Bernstein, Sur l'ordre de la meilleure approximation des fonctions continues par les polyômes de degré donné, Mem. Cl. Sci. Acad. Roy. Belg. 4 (1912) 1-103.

[5] M.D. Buhmann, Radial Basis Functions: Theory and Implementations, Cambridge University Press, New York, 2003.

[6] R. DeVore, G.G. Lorentz, Constructive Approximation, Springer-Verlag, Berlin, 1993.

[7] R. DeVore, A. Ron, Approximation using scattered shifts of a multivariate function, Trans. Amer. Math. Soc., in press.

[8] J. Duchon, Interpolation des fonctions de deux variables suivant le principe de la flexion des plaques minces, RAIRO Anal. Numer. 10 (R-3) (1976) 5-12.

[9] J. Duchon, Sur l'erreur d'interpolation des fonctions de plusieurs variables par les D^m-splines. RAIRO Anal. Numer. 12 (4) (1976) 325-334.

[10] G.B. Folland, Fourier Analysis and Its Applications, Brooks/Cole Publishing Company, Pacific Grove, CA, 1992.

[11] G.B. Folland, Real Analysis, John Wiley & Sons, New York, 1999.

[12] Hardy, Multiquadric equations of topography and other irregular surfaces, J. Geophys. Res. 76 (8) (1971) 1905-1915.

[13] M.J. Johnson, Approximation in $L^p(\mathbb{R}^d)$ from spaces spanned by the perturbed integer translates of a radial function, J. Approx. Theory 107 (2) (2000) 163-203.

[14] M. Mathias, Über positive Fourier-integrale, Math Z. 16 (1) (1923) 103-124.

[15] H.N. Mhaskar, A Markov-Bernstein inequality for Gaussian networks, in: Trends and Applications in Constructive Approximation, vol. 151 of Internat. Ser. Numer. Math., Birkhäuser, Basel, 2005, pp. 165-180.

[16] H.N. Mhaskar, F.J. Narcowich, J. Prestin, J.D. Ward, L^p Bernstein estimates and approximation by spherical basis functions, Math. Comp. 79 (2010) 1647-1679.

[17] H.N. Mhaskar, J. Prestin, On Marcinkiewicz-Zygmund-type inequalities, in: Approximation Theory: In Memory of A. K. Varma, Marcel Dekker, New York, 1998, pp. 389-404.

[18] F.J. Narcowich, J.D. Ward, H. Wendland, Sobolev error estimates and a Bernstein inequality for scattered data interpolation via radial basis functions, Constr. Approx. 24 (2) (2006) 175-186.

[19] R. Schaback, H. Wendland, Inverse and saturation theorems for radial basis function interpolation, Math. Comp. 71 (238) (2002) 669-681.

[20] E.M. Stein, Singular Integrals and Differentiability Properties of Functions, Princeton University Press, Princeton, NJ, 1970.

[21] E.M. Stein, G. Weiss, Introduction to Fourier Analysis on Euclidean Spaces, Princeton University Press, Princeton, NJ, 1971.

[22] G.N. Watson, A Treatise on the Theory of Bessel Functions, Cambridge University Press, Cambridge, UK, 1944.

[23] H. Wendland, Scattered Data Approximation, Cambridge University Press, Cambridge, UK, 2005.

[24] J. Yoon, Approximation in $L^p(\mathbb{R}^d)$ from a space spanned by the scattered shifts of a radial basis function, Constr. Approx. 17 (2) (2001) 227-247.

APPENDIX A

BESSEL FUNCTIONS AND FOURIER INTEGRALS

A d-dimensional Fourier integral of a radial function reduces to a 1-dimensional integral involving a Bessel function of the first kind. In what follows, we will list some of the properties of these Bessel functions and prove two propositions which will be useful for bounding Fourier integrals.

Proposition A.1. ([23, Proposition 5.4 & Proposition 5.6])

(1) $\frac{d}{dz}\{z^\nu J_\nu(z)\} = z^\nu J_{\nu-1}(z)$

(2) $J_{1/2}(z) = \sqrt{\frac{2}{\pi z}}\sin(z),\ J_{-1/2}(z) = \sqrt{\frac{2}{\pi z}}\cos(z)$

(3) $J_\nu(r) = \sqrt{\frac{2}{\pi r}}\cos(r - \frac{\nu\pi}{2} - \frac{\pi}{4}) + \mathcal{O}(r^{-3/2})$ *for* $r \to \infty$ *and* $\nu \in \mathbb{R}$

(4) $J^2_{l/2}(r) \leq \frac{2^{l+2}}{\pi r}$ *for* $r > 0$ *and* $l \in \mathbb{N}$

(5) $\lim_{r\to 0} r^{-l} J^2_{l/2}(r) = \frac{1}{2^l\Gamma^2(l/2+1)}$ *for* $l \in \mathbb{N}$

The next proposition makes use of integration by parts in order to bound the Fourier integral of a function whose support lies outside of a neighborhood of the origin.

Proposition A.2. *Let* $\alpha \geq 1$*, and let* $f \in C^n([0,\infty))$ *for some natural number* $n > 1$*. Also, assume there are constants* $C, \epsilon > 0$ *such that* $f = 0$ *on* $[0, \frac{1}{2}]$ *and* $\left|f^{(j)}(t)\right| \leq Ct^{-d-\epsilon}$ *for* $j \leq n$*. Then there is a constant* C_ϵ *such that*
$\left|\int_{1/2}^\infty f(t)t^{d/2}J_{(d-2)/2}(\alpha t)dt\right| \leq C_\epsilon \alpha^{-n}$

Proof. We first define a sequence of functions arising when integrating by parts. Let $f_0 = f$, $f_1 = f'$, and $f_j = \left(\frac{f_{j-1}}{t}\right)'$ for $j \geq 2$. Note that when $j \geq 2$, there are constants $c_{j,l}$ such that $f_j(t) = \sum_{l=1}^{j} c_{j,l} f^{(l)}(t) t^{-2j+l+1}$, and therefore $|f_j(t)| \leq Ct^{-d-j+1-\epsilon}$.

Applying the Dominated Convergence Theorem, we have

$$\left|\int_{1/2}^{\infty} f_0(t)t^{d/2}J_{(d-2)/2}(\alpha t)dt\right| = \lim_{b\to\infty}\left|\int_{1/2}^{b} f_0(t)t^{d/2}J_{(d-2)/2}(\alpha t)dt\right|.$$

After ingetrating by parts and taking the limit we get

$$\left|\int_{1/2}^{\infty} f_0(t)t^{d/2}J_{(d-2)/2}(\alpha t)dt\right| = \frac{1}{\alpha}\left|\int_{1/2}^{\infty} \frac{f_1(t)}{t}t^{d/2+1}J_{d/2}(\alpha t)dt\right|.$$

Integrating by parts j times, we have

$$\begin{aligned}\left|\int_{1/2}^{\infty} f_0(t)t^{d/2}J_{(d-2)/2}(\alpha t)dt\right| &= \frac{1}{\alpha^j}\lim_{b\to\infty}\left|\int_{1/2}^{b} \frac{f_j(t)}{t}t^{d/2+j}J_{d/2+j-1}(\alpha t)dt\right| \\ &= \frac{1}{\alpha^{j+1}}\left|\int_{1/2}^{\infty} \frac{f_{j+1}(t)}{t}t^{d/2+j+1}J_{d/2+j}(\alpha t)dt\right|,\end{aligned}$$

and therefore

$$\left|\int_{1/2}^{\infty} f_0(t)t^{d/2}J_{(d-2)/2}(\alpha t)dt\right| \leq \frac{1}{\alpha^n}\int_{1/2}^{\infty}\left|\frac{f_n(t)}{t}t^{d/2+n}J_{d/2+n-1}(\alpha t)\right|dt.$$

□

We will also need a bound on Fourier integrals of functions that are identically one in a neighborhood of the origin and have compact support. As in the previous case, the proof relies on integration by parts.

Proposition A.3. *Let $\alpha \geq 1$, and let f be a function in $C^n([0,\infty))$ for some natural number $n > 1$. Also, assume $f = 1$ in a neighborhood of 0 and $f = 0$ for $x > 2$. Then there is a constant C such that*
$\left|\int_0^2 tf(t)t^{d/2}J_{(d-2)/2}(\alpha t)dt\right| \leq C\alpha^{-n}$

Proof. We first define a sequence of functions arising when integrating by parts. Let $f_0(t) = tf(t)$, $f_1 = f_0'$, and $f_j = \left(\frac{f_{j-1}}{t}\right)'$ for $j \geq 2$. Note that when $j \geq 2$, $f_j(t) =$

$O(t^{-2j+2})$ as $t \to 0$. After integrating by parts n times, we have

$$\left|\int_0^2 f_0(t)t^{d/2}J_{(d-2)/2}(\alpha t)dt\right| = \frac{1}{\alpha^n}\left|\int_0^2 \frac{f_n(t)}{t}t^{d/2+n}J_{d/2+n-1}(\alpha t)dt\right|.$$

□

APPENDIX B

SUMS OF FUNCTION VALUES

Proposition B.1. *Let $X \subset \mathbb{R}^d$ be a countable set with $q_X > 0$, and let $Y = \{y_j\}_{j=1}^N$ be a subset of X such that $||y_j||_2 \geq q_X$ for $1 \leq j \leq N$. If $f : \mathbb{R}^d \to \mathbb{R}$ is a function with $|f(x)| \leq C\,||x||_2^{-d-\epsilon}$ for some $C, \epsilon > 0$, then*

$$\sum_{j=1}^{N} |f(y_j)| \leq 3^d(1 + 1/\epsilon)(Cq_X^{-d-\epsilon}).$$

Proof. We can bound the sum using the volume argument found in the proof of [23, Theorem 12.3]. Following the same procedure, we define

$$E_m = \{x \in \mathbb{R}^d : mq_Y \leq ||x||_2 < (m+1)q_Y\}.$$

for each positive integer m. Now by comparing the volume of E_m to the volume of a ball of radius q_Y, one finds that $\#Y \cap E_m \leq 3^d m^{d-1}$. Therefore,

$$\begin{aligned}
\sum_{j=1}^{N} |f(y_j)| &\leq \sum_{j=1}^{N} \frac{C}{||y_j||_2^{d+\epsilon}} \\
&\leq \sum_{m=1}^{\infty} (\#Y \cap E_m) \max_{y \in E_m} \left\{ \frac{C}{||y||_2^{d+\epsilon}} \right\} \\
&\leq \sum_{m=1}^{\infty} 3^d m^{d-1} \frac{C}{(mq_Y)^{d+\epsilon}} \\
&\leq 3^d(1 + 1/\epsilon)Cq_X^{-d-\epsilon}
\end{aligned}$$

□

APPENDIX C

WENDLAND FUNCTIONS

The Wendland functions are a class of compactly supported RBFs that are radially defined as piecewise polynomials, and some examples are provided in the table below. Wendland's book [23] (particularly chapter 9) provides a detailed analysis of these functions and some their approximation properties. In this appendix, our focus will be on computing the Fourier transforms of these functions. The Wendland functions are determined by a dimension parameter d and a smoothness parameter k, and $\Phi_{d,k} \in C^{2k}(\mathbb{R}^d)$. In what follows, r will denote $||x||_2$, and $\doteq$ will be used to indicate equality up to some positive constant factor.

Table 1

Examples of Wendland functions

Function	Smoothness
$\Phi_{1,0}(x) = (1-r)_+$	C^0
$\Phi_{1,1}(x) \doteq (1-r)_+^3(3r+1)$	C^2
$\Phi_{1,2}(x) \doteq (1-r)_+^5(8r^2+5r+1)$	C^4
$\Phi_{3,0}(x) = (1-r)_+^2$	C^0
$\Phi_{3,1}(x) \doteq (1-r)_+^4(4r+1)$	C^2
$\Phi_{3,2}(x) \doteq (1-r)_+^6(35r^2+18r+3)$	C^4
$\Phi_{3,3}(x) \doteq (1-r)_+^8(32r^3+25r^2+8r+1)$	C^6

We will now derive an explicit form of the Fourier transform of $\Phi_{d,k}$ in the case d is odd. Using the notation of [23, Section 10.5], let $d = 2n+1$ and $m = n+k$. Then by [23, Lemma 6.19] and the definition of $\Phi_{d,k}$, we have

$$\hat{\Phi}_{d,k}(x) = B_m f_m(r) r^{-3m-2} \tag{C.1}$$

where B_m is a positive constant and the Laplace transform of f_m satisfies

$$\mathcal{L} f_m(r) = \frac{1}{r^{m+1}(1+r^2)^{m+1}}.$$

In order to find the inverse Laplace transform of the above expression, we will make use of partial fractions. First, note that there exist constants α_j, β_j, and γ_j such that

$$\frac{1}{s^{m+1}(1+s^2)^{m+1}} = \sum_{j=0}^{m} \frac{\alpha_j}{s^{j+1}} + \sum_{j=0}^{m} \frac{\beta_j}{(s+i)^{j+1}} + \sum_{j=0}^{m} \frac{\gamma_j}{(s-i)^{j+1}}, \tag{C.2}$$

and this decomposition is unique. Now for any real s, the expression on the left is real. Therefore, taking the complex conjugate of both sides, we get

$$\frac{1}{s^{m+1}(1+s^2)^{m+1}} = \sum_{j=0}^{m} \frac{\bar{\alpha}_j}{s^{j+1}} + \sum_{j=0}^{m} \frac{\bar{\beta}_j}{(s-i)^{j+1}} + \sum_{j=0}^{m} \frac{\bar{\gamma}_j}{(s+i)^{j+1}}.$$

Uniqueness of the decomposition then implies that for each j:

(i) α_j is real

(ii) $\beta_j = \bar{\gamma}_j$.

To further characterize the coefficients, we replace s by $-s$ in (C.2). First, we have

$$\frac{(-1)^{m+1}}{s^{m+1}(1+s^2)^{m+1}} = \sum_{j=0}^{m} \frac{(-1)^{j+1}\alpha_j}{s^{j+1}} + \sum_{j=0}^{m} \frac{(-1)^{j+1}\beta_j}{(s-i)^{j+1}} + \sum_{j=0}^{m} \frac{(-1)^{j+1}\bar{\beta}_j}{(s+i)^{j+1}},$$

and therefore

$$\frac{1}{s^{m+1}(1+s^2)^{m+1}} = \sum_{j=0}^{m} \frac{(-1)^{j+m}\alpha_j}{s^{j+1}} + \sum_{j=0}^{m} \frac{(-1)^{j+m}\beta_j}{(s-i)^{j+1}} + \sum_{j=0}^{m} \frac{(-1)^{j+m}\bar{\beta}_j}{(s+i)^{j+1}}. \tag{C.3}$$

Again using the uniqueness of the partial fraction decomposition, it follows that

(i) $(-1)^{j+m}\alpha_j = \alpha_j$

(ii) $(-1)^{j+m}\beta_j = \bar{\beta}_j$

for each j. The first property implies that $\alpha_j = 0$ for either all odd j or all even j. The second property tell us that β_j is real when $j+m$ is even, and it is imaginary when $j+m$ is odd.

We now compute f_m as the inverse Laplace transform of the sum in (C.2).

$$f_m(r) = \sum_{j=0}^{m} \frac{\alpha_j}{j!} r^j + \sum_{j=0}^{m} \frac{\beta_j}{j!} r^j e^{-ir} + \sum_{j=0}^{m} \frac{\bar{\beta}_j}{j!} r^j e^{ir}$$

Now if m is odd, we have $m = 2l+1$ and

$$f_m(r) = \sum_{j=0}^{l} \frac{\alpha_{2j+1}}{(2j+1)!} r^{2j+1} + \sum_{j=0}^{l} \frac{\beta_{2j+1}}{(2j+1)!} r^{2j+1}(e^{-ir}+e^{ir}) - \sum_{j=0}^{l} \frac{\beta_{2j}}{(2j)!} r^{2j}(-e^{-ir}+e^{ir}),$$

which can be simplified to

$$f_{2l+1}(r) = \sum_{j=0}^{l} \frac{\alpha_{2j+1}}{(2j+1)!} r^{2j+1} + \sum_{j=0}^{l} \frac{2\beta_{2j+1}}{(2j+1)!} r^{2j+1}\cos(r) - \sum_{j=0}^{l} \frac{2i\beta_{2j}}{(2j)!} r^{2j}\sin(r).$$

Similarly, when $m = 2l$ we have

$$f_{2l}(r) = \sum_{j=0}^{l} \frac{\alpha_{2j}}{(2j)!} r^{2j} - \sum_{j=0}^{l-1} \frac{2i\beta_{2j+1}}{(2j+1)!} r^{2j+1}\sin(r) + \sum_{j=0}^{l} \frac{2\beta_{2j}}{(2j)!} r^{2j}\cos(r).$$

Lemma C.1. *Let $\Phi_{d,k}$ be a Wendland function with d odd, and define n and m by $d = 2n+1$ and $m = n+k$. Then the exact form of the Fourier transform of $\Phi_{d,k}$ is found by substituting the above representations of f_m into $\hat{\Phi}_{d,k}(x) = B_m f_m(r) r^{-3m-2}$.*

We will now take a closer look at the 1-dimensional case.

Proposition C.2. *If $k \in \mathbb{N}$, then there exists $C \in \mathbb{R}$ and $h \in L^1(\mathbb{R})$ such that*

$$x^{2k+2}\hat{\Phi}_{1,k}(x) = B_k \left(\frac{1}{k!} + \frac{(-1)^{k+1}}{k! 2^k} \cos(x) + C\frac{\sin(x)}{x} + \hat{h}(x) \right), \tag{C.4}$$

where $\hat{h}$ is the Fourier transform of h.

Proof. Since the case where k is odd is similar to the case where k is even, we will only prove the former. Recall that $\hat{\Phi}_{1,k}(x) = B_k f_k(r) r^{-3k-2}$, so let us begin by examining the function f_k. For $k = 2l+1$, we have

$$r^{-1} f_k(r) = \sum_{j=0}^{l} a_j r^{2j} + \sum_{j=0}^{l} b_j r^{2j} \cos(r) - \sum_{j=0}^{l} c_j r^{2j} \frac{\sin(r)}{r} \tag{C.5}$$

for some constants a_j, b_j, and c_j. We can then define an analytic function $\tilde{f}_k : \mathbb{R} \to \mathbb{R}$ by

$$\tilde{f}_k(x) = |x|^{-1} f_k(|x|),$$

and we will have $x^{2k+2}\hat{\Phi}_{1,k}(x) = B_k x^{-k+1} \tilde{f}_k(x)$. Now since $\Phi_{1,k} \in L^1$ and $\hat{\Phi}_{1,k}(x) = B_k f_k(r) r^{-3k-2}$, $\tilde{f}_k(x)$ must have a zero of order $3k+1$ at 0, and therefore $\tilde{f}_k(x)$ has a power series of the form

$$\tilde{f}_k(x) = \sum_{j=3k+1}^{\infty} d_j x^j. \tag{C.6}$$

In order to verify (C.4), we first need to determine some of the coefficients in (C.5). From our previous work, we know that $c_l \in \mathbb{R}$, and we can find a_l and b_l as follows. In the partial fraction decomposition (C.2), multiply both sides by $s^{m+1}(1+s^2)^{m+1}$. By substituting the values $s = 0$ and $s = -i$, we find that $\alpha_m = 1$

and $\beta_m = (-1)^m/2^{m+1}$, and therefore $a_l = 1/k!$ and $b_l = (-1)^k/(k!2^{k+1})$.

We can now finish the proof by showing that

$$\tilde{h}_k(x) := x^{-k+1}\tilde{f}_k(x) - \left(\frac{1}{k!} + \frac{(-1)^{k+1}}{k!2^k}\cos(x) - c_l\frac{\sin(x)}{x}\right)$$

is the Fourier transform of an L^1 function. Since $\tilde{h}_k(x)$ is identically 0 for $k = 1$, we need only consider $k \geq 3$. This can be verified by determining that $\tilde{h}_k(x)$ has two continuous derivatives in L^1, cf. [10, p. 219]. Considering the representation (C.6), it is clear that $\tilde{h}_k(x)$ has two continuous derivatives, and the decay of these functions can be bounded using (C.5). □

VITA

Name:	John Paul Ward
Address:	Department of Mathematics
	Texas A&M University
	College Station, TX 77843-3368

Education:	B.S., Mathematics, University of Georgia, 2005