Constr. Approx. (2001) 17: 227–247
DOI: 10.1007/s003650010033

CONSTRUCTIVE APPROXIMATION

Approximation in $L_p(\mathbf{R}^d)$ from a Space Spanned by the Scattered Shifts of a Radial Basis Function

J. Yoon

Abstract. A new multivariate approximation scheme on $\mathbf{R}^d$ using scattered translates of the "shifted" surface spline function is developed. The scheme is shown to provide spectral L_p-approximation orders with $1 \le p \le \infty$, i.e., approximation orders that depend on the smoothness of the approximands. In addition, it applies to noisy data as well as noiseless data. A numerical example is presented with a comparison between the new scheme and the surface spline interpolation method.

1. Introduction

1.1. *General*

Approximation schemes of the form

$$s(x) := \sum_{\xi \in \Xi} c_\xi \varphi(x - \xi), \qquad x \in \mathbf{R}^d, \tag{1.1}$$

with φ a "suitable" basis function are known to be effective for approximation to scattered data. The use of a radially symmetric basis function φ is particularly useful because this facilitates the evaluation of the approximant, while still leaving enough flexibility in the choice of φ. The set Ξ in $\mathbf{R}^d$ by which the radial basis function φ is shifted is usually referred to as a set of "centers." The common choices of φ include: $\varphi(x) = |x|^\lambda \log|x|$ with d and λ both even integers (surface spline), $\varphi(x) = (|x|^2 + c^2)^{\lambda/2}$ with λ and d both odd integers (multiquadric), and $\varphi(x) = \exp(-c|x|^2)$, $c > 0$ (Gaussian).

The initial approximation method using radial basis functions has been obtained by means of interpolation at finitely many scattered points Ξ in $\mathbf{R}^d$. However, while the interpolation method is certainly an important approach toward solving the scattered data problem, it has several drawbacks. For example, for a large class of basis functions (including multiquadrics and inverse multiquadrics), the existing theories guarantee the interpolant to approximate well for only a very small class of very smooth approximands (see [MN2]). Another drawback of the interpolation method is connected with the issue of numerical stability: as the number of centers increases, one needs to solve a large linear

Date received: March 22, 1999. Date revised: May 18, 1999; March 21, 2000; August 9, 2000. Date accepted: August 16, 2000. Communicated by Robert Schaback. On line publication: February 5, 2001.
AMS classification: 41A15, 41A25, 41A30, 41A63.
Key words and phrases: Radial basis function, Approximation order, "Shifted" surface, Spline, Multiquadrics.

system which is ill-conditioned. Last but not least, interpolation is never recommended when the data are known to be contaminated (= noisy data). All in all, there is an overwhelming need for approximation methods other than interpolation.

In view of the above discussion, we need a scheme with the following properties:

(i) it should approximate well a large class of functions;
(ii) it should be "local," namely, a coefficient in (1.1) should be determined by a few values of the data, even when many centers are involved in the scheme; and
(iii) the scheme should have a "smoothing" effect.

Thus, the main objective of this paper is, indeed, to construct an approximation scheme on nonuniformly distributed centers that satisfies all the above (and more).

It should be noted that noninterpolatory approximation schemes of the type (1.1) are also discussed intensively in the literature. However, most of the results in that direction deal with the case when the center set Ξ is infinite and uniform, i.e., a scale $\delta\mathbf{Z}^d$ of the integer lattice $\mathbf{Z}^d$. In fact, there are only a handful of treatments of noninterpolatory schemes for arbitrary center sets Ξ. Buhmann, Dyn, and Levin [BuDL] were among the first to construct a noninterpolatory approximation scheme for infinitely many scattered centers and to analyze its approximation power. Dyn and Ron [DR] showed that the scheme in [BuDL] can be understood as "an approximation to a uniform mesh approximation scheme." In both papers, quasi-interpolation schemes from radial basis function spaces with infinitely many centers Ξ were studied and both showed that the approximation orders obtained in the scattered case are identical to those that had been known on uniform grids. In particular, N. Dyn and A. Ron provided a general tool that allows us to convert *any* known approximation scheme on uniform grids to nonuniform grids, while preserving (to the extent that this is possible) the approximation orders known in the former case. The approach of [DR] can be described as follows: suppose that we are given an approximation scheme

$$f \mapsto \sum_{\alpha\in\mathbf{Z}^d} \lambda_\alpha(f)\varphi(\cdot-\alpha).$$

Then, we replace each $\varphi(\cdot-\alpha)$ by a suitable combination

$$\sum_{\xi\in\Xi} A(\alpha,\xi)\varphi(\cdot-\xi),$$

with Ξ the set of scattered centers we wish to use. This result, however, requires one to choose the density of a uniform grid that is associated with the given scattered set Ξ. As an alternative, we construct in this paper a new approximation scheme that, while based on the general idea of [DR], is not connected to uniform grid approximations. The approximation scheme that is developed and analyzed here is intrinsically "scattered": it employs directly the scattered shifts of the basis function φ. Furthermore, while the conversion tool in [DR] is applied there only to stationary schemes (see [BR], [DJLR]), we successfully apply our new scheme to the more general nonstationary case. This results in schemes that provide spectral approximation orders (i.e., approximation orders that depend only on the smoothness of the approximands f we approximate).

Before we advance our discussion further, we would like to comment on the notion of "radial basis function," a comment which, as a matter of fact, is valid for all studies in the area of scattered data approximation. Many basis functions that are not necessarily radially symmetric fit the theory developed here. The mere advantage of the radial symmetry concerns the ease of the actual calculations, and has no theoretical advantage.

Among the basis functions currently in use, we choose our basis function to be the "shifted" surface spline

$$\varphi_c(x) := \begin{cases} (|x|^2 + c^2)^{\lambda/2}, & \lambda \in \mathbf{Z}_+, \lambda, d \text{ odd}, \\ (|x|^2 + c^2)^{\lambda/2} \log(|x|^2 + c^2)^{1/2}, & \lambda \in \mathbf{Z}_+, \lambda, d \text{ even}, \end{cases} \tag{1.2}$$

whose properties are quite well understood, both theoretically and practically. One of the reasons for choosing this particular function (over the "shifted" surface spline) is the desire to use the parameter c as a "tension" parameter. Note that we stress this tension parameter by using the notation φ_c. When λ and d are odd integers, the function $\varphi_c(x) = (|x|^2 + c)^{\lambda/2}$ is usually referred to as a "multiquadric." In particular, if $c = 0$, the resulting function, $\varphi_0(x) = |x|^\lambda \log |x|$, is the so-called "surface spline."

The reader who is interested in knowing more about the state-of-the-art in the area of radial basis function methods may find it useful to consult with the surveys [Bu], [D], and [P]. Another important source is the work of Madych and Nelson [MN1,2], who developed a theory of interpolation based on reproducing kernel Hilbert spaces. The general conditions on φ that ensure the existence and uniqueness of a solution of the interpolation method have been given by Micchelli [M]. More recently, M. J. Johnson [J] established an upper bound on the approximation order when interpolating data that are defined in the unit ball of $\mathbf{R}^d$. His analysis covered the basis function φ of the form $\varphi = |\cdot|^{\lambda/2}$ for d, λ odd, and $\varphi = |\cdot|^{\lambda/2} \log |\cdot|$ for d, λ even.

The following notations are used throughout this paper. For the given function φ_c and a discrete $\Xi \subset \mathbf{R}^d$, we define

$$S_\Xi(\varphi_c) := \text{closure } S_0(\varphi_c),$$

under the topology of uniform convergence on compact sets, with

$$S_0(\varphi_c) := \text{span}\{\varphi_c(\cdot - \xi) \colon \xi \in \Xi\},$$

the finite span of $\{\varphi_c(\cdot - \xi) : \xi \in \Xi\}$. The Fourier transform of $f \in L_1(\mathbf{R}^d)$ is defined as

$$\hat{f}(\theta) := \int_{\mathbf{R}^d} f(t) e_{-\theta}(t)\, dt, \qquad e_\theta \colon x \mapsto e^{i\theta \cdot x}.$$

Also, for a function $f \in L_1(\mathbf{R}^d)$, we use the notation $f^\vee$ for the inverse Fourier transform. We assume that the reader is familiar with the usual properties of the Fourier transform. In particular, the Fourier transform can be uniquely extended to the space of tempered distributions on $\mathbf{R}^d$. Several different function norms are used. In particular, the L_p-norm, $1 \le p \le \infty$, is denoted as

$$\|f\|_p := \|f\|_{L_p(\mathbf{R}^d)}.$$

For $x = (x_1, \ldots, x_d)$ in $\mathbf{R}^d$:

$$|x| := (x_1^2 + x_2^2 + \cdots + x_d^2)^{1/2}$$

stands for its Euclidean norm and, for $\alpha \in \mathbf{Z}_+^d := \{\beta \in \mathbf{Z}^d : \beta \geq 0\}$, we set $\alpha! := \alpha_1! \cdots \alpha_d!$ and $|\alpha|_1 := \sum_{k=1}^d \alpha_k$. Finally, Π_k stands for the space of all polynomials of degree $\leq k$ in d variables.

Given a function f in $C(\mathbf{R}^d \backslash 0)$, we say that f **has a singularity of order** k **at the origin** if there exist some constants $c_1, c_2 > 0$ such that $c_1 \leq |\cdot|^k |f| \leq c_2$ in some punctured neighborhood of the origin.

Asymptotic approximation properties are usually quantified by the notion of approximation order. In order to make this notion precise, we define the density of Ξ by

$$h := h(\Xi) := \sup_{x \in \mathbf{R}^d} \inf_{\xi \in \Xi} |x - \xi|. \tag{1.3}$$

Then, given a sequence $(L_h)_h$ of approximation schemes, we say that $(L_h)_h$ provides an L_p-approximation order $k > 0$ if, for every sufficiently smooth $f \in L_p(\mathbf{R}^d)$,

$$\| f - L_h f \|_p = O(h^k), \qquad 1 \leq p \leq \infty,$$

as h tends to 0. In our study, the range of L_h is the space $S_\Xi(\varphi_c)$ for certain Ξ and c. Note that we have to index the operator L_h by the density $h := h(\Xi)$ of the center set Ξ. While this is convenient in the context of discussing approximation orders, the reader should keep in mind that the actual scheme L_h depends on the choice of φ_c as well as on the given set Ξ because L_h maps to $S_\Xi(\varphi_c)$.

We used above the loose term of "sufficiently smooth f." More precisely, our approximands are chosen from the Sobolev space

$$W_p^k(\mathbf{R}^d), \qquad 1 \leq p \leq \infty, \quad k \in \mathbf{Z}_+,$$

of all functions whose derivatives of orders $\leq k$ are in $L_p(\mathbf{R}^d)$. By $|\cdot|_{k,p}$, we shall denote the homogeneous kth order L^p-Sobolev semi-norm, i.e.,

$$|f|_{k,p} := \sum_{|\alpha|_1 = k} \|D^\alpha f\|_p.$$

1.2. *An Outline of our Approach*

Let $f \colon \mathbf{R}^d \to \mathbf{R}^d$ be a smooth function, and let φ_c be the "shifted" surface spline as in (1.2). The basis function φ_c should not be used directly in any approximation scheme since it increases polynomially fast around ∞. However, a suitable bell-shape

$$\psi_c(x) = \sum_{\alpha \in \mathbf{Z}^{d^d}} \mu(\alpha) \varphi_c(x - \alpha) \tag{1.4}$$

is obtained by applying a difference operator. To simplify the discussion at this introductory stage, we assume that the localization sequence μ is of finite support.

In our study, we obtain an approximation scheme into the space spanned by the scattered shifts of φ_c by employing the following 2-step method. In the first step we mollify f: we approximate f by

$$\sigma_\omega^\vee * f$$

with σ_ω a cut-off function, where ω is a parameter depending on h, i.e.,

$$\omega := \omega(h),$$

(we discuss this point later in further detail). In particular, we choose the parameter c in φ_c so that the ratio

$$\rho := c/\omega$$

is held fixed. This will lead to the construction of numerically stable schemes. Next, we look for an approximant from the space $S_\Xi(\varphi_c)$ which approximates $\sigma_\omega^\vee * f$ in some sense. To this end we realize that we know how to approximate well $\sigma_\omega^\vee * f$ by convolution:

$$\sigma_\omega^\vee * f \approx [\psi_\rho * \Lambda f(\omega\cdot)](\cdot/\omega).$$

Here, the function ψ_ρ is a localization of φ_ρ as in (1.4) so that, after some calculations, one can see that

$$\text{(1.5)} \qquad \psi_\rho(\cdot/\omega - t) = \psi_{c/\omega}(\cdot/\omega - t) = \sum_{\alpha\in\mathbf{Z}^d} \mu(\alpha)\varphi_c(\cdot - (t+\alpha)\omega)/\omega^\lambda.$$

Also, Λ is an operator depending on σ_ω and it is of the form

$$\text{(1.6)} \qquad \Lambda f = m_{c,\omega} * f$$

with $m_{c,\omega}*$ a mollifier that "inverts" (in a suitable sense) the convolution $\psi_\rho*$. However, the above approximant then lies in the space spanned by *all* the translates of φ_c, contrary to our desire to have the approximants in $S_\Xi(\varphi_c)$. Therefore, we approximate the convolution kernel $(x, t) \mapsto \psi_\rho(x/\omega - t)$ by a kernel

$$(x, t) \mapsto K(x, t).$$

The basic properties of K are as follows:

(a) for a fixed $t \in \mathbf{R}^d$, $K(\cdot, t) \in S_\Xi(\varphi_c)$; and
(b) for a fixed $x \in \mathbf{R}^d$, $K(x, \cdot) \in L_1(\mathbf{R}^d)$.

Our quasi-interpolant is then initially written in the integral form

$$\text{(1.7)} \qquad R_\Xi f := \int_{\mathbf{R}^d} K(\cdot, t)(\Lambda f)(\omega t)\, dt \approx \sigma_\omega^\vee * f$$

with the same operator Λ in (1.6).

The construction of the kernel $K(\cdot, t)$ is done as follows: we first approximate each $\varphi_c(\cdot - t)$ by a linear combination

$$\text{(1.8)} \qquad \varphi_c(\cdot, t) := \sum_{\xi\in\Xi} A(t, \xi)\varphi_c(\cdot - \xi) \in S_\Xi(\varphi_c).$$

Here the sequence $(A(t, \cdot))$ is assumed to be finitely supported for each $t \in \mathbf{R}^d$. (Note that the above sequence is defined on Ξ; however, we do not necessarily assume Ξ to be finite.) Further properties of the map $(t, \xi) \mapsto A(t, \xi)$ that are essential for the success of (1.7) will be analyzed later. Next, by replacing each shift $\varphi_c(\cdot - t)$ in (1.5) by $\varphi_c(\cdot, t)$, we define $K(\cdot, t)$ by

$$K(\cdot, t) := \sum_{\alpha \in N} \mu(\alpha)\varphi_c(\cdot, (t+\alpha)\omega)/\omega^\lambda \approx \psi_\rho(\cdot/\omega - t). \tag{1.9}$$

It is obvious from this form that $K(\cdot, t) \in S_\Xi(\varphi_c)$, for each $t \in \mathbf{R}$. Note also that the kernel K depends on the locations of the centers Ξ, but not on the approximand f.

With the approximation scheme R_Ξ above at hand, our goal of this study is to prove error bounds of the following form (Theorem 1.1). The theorem we state now can be regarded as a prototype for the main results. We remind the reader that the parameter λ appears in the definition of φ_c (and should not be confused with the convolution operator Λ) while d is the spatial dimension.

Theorem 1.1. *Let the basis function φ_c and the approximation scheme R_Ξ be given as above. Let the coefficients $(A(t,\xi))_{\xi\in\Xi}$ for $\varphi_c(\cdot, t)$ in* (1.8) *satisfy the relation*

$$\sum_{\xi\in\Xi} A(t,\xi)q(\xi) = q(t) \qquad \text{for any} \quad q \in \Pi_n$$

with n a "sufficiently large" integer. Assume that $f \in W_p^k(\mathbf{R}^d) \cap W_1^m(\mathbf{R}^d)$ with $m := \min(k, \lambda + d)$, where $1 \le p \le \infty$. Then, if $k < \lambda + d$, we have an estimate

$$\| f - R_\Xi f \|_p = o(h^k).$$

Also, if $k \ge \lambda + d$, we have

$$\| f - R_\Xi f \|_p = o(h^{rk}),$$

which is valid for every $0 < r < 1$.

Note that the theorem indicates that we obtain approximation orders that depend on the smoothness of f only. We emphasize that the number n, which represents "accuracy" (as well as the complexity) of the scheme, depends on $r \in (0, 1)$ and the smoothness parameter k: the higher approximation order we desire, the more complex our scheme is. We also emphasize that the "tension parameter" c that appears in the definition of φ_c may vary with Ξ.

The layout of this paper is as follows: Section 2 is devoted to the development of the approximation scheme R_Ξ. In Section 2.1, some basic properties of φ_c are discussed and then the scheme R_Ξ is constructed. Also discussed there is the issue of the smoothing effects of R_Ξ. In Section 2.2, we discuss properties of the matrix $A(\cdot, \cdot)$ (which appears in (1.8)) that are essential for the success of (1.7). Section 3 is devoted to error analysis. Specifically, the approximation order of the scheme R_Ξ is discussed in Section 3.2. As was said, we will show that R_Ξ provides *spectral approximation orders*: the convergence rate of the scheme will be determined by the decay at infinity of $\hat{f}$ (i.e., by the smoothness of the approximand f). Finally, in Section 4, a specific numerical example which illustrates the accuracy of approximation by using the scheme R_Ξ is provided. The initial numerical tests reveal that the new approximation scheme gives better results than the surface spline interpolation method, which is one of the best well-known methods in this area.

2. The Approximation Scheme R_Ξ and the Function $\varphi_c(\cdot, t)$

2.1. *The Approximation Scheme*

As mentioned in the Introduction, we choose to focus on "basis" functions that are obtained from the fundamental solution of the iterated Laplacian by the shifting $|x| \mapsto (|x|^2 + c^2)^{1/2}$ with $c > 0$:

$$\varphi_c(x) = \begin{cases} (|x|^2 + c^2)^{\lambda/2}, & \lambda \in \mathbf{Z}_+, \lambda, d \text{ odd}, \\ (|x|^2 + c^2)^{\lambda/2} \log(|x|^2 + c^2)^{1/2}, & \lambda \in \mathbf{Z}_+, \lambda, d \text{ even}. \end{cases}$$

The generalized Fourier transforms of these functions satisfy that

$$\hat{\varphi}_c(\theta) = c(\lambda, d)|\theta|^{-\lambda-d}\tilde{K}_{(d+\lambda)/2}(|c\theta|), \tag{2.1}$$

[GS], where $c(\lambda, d)$ is a constant depending on λ and d, and $\tilde{K}_\nu(|t|) := |t|^\nu K_\nu(|t|)$ with $K_\nu(|t|)$ the modified Bessel function of order ν. (Note that despite the similarity in the notations, there is no direct connection between the above $\tilde{K}$ and the kernel K.) The following properties of $\tilde{K}_\nu$ are related to our analysis:

$$\begin{aligned} &\tilde{K}_\nu(|t|) \approx \sqrt{\frac{\pi}{2}} t^{\nu-1/2} e^{-|t|} \qquad (t \to \infty), \\ &\tilde{K}_\nu(|t|) \in C^{2\nu-1}(\mathbf{R}^d) \cap C^\infty(\mathbf{R}^d \backslash 0), \end{aligned} \tag{2.2}$$

see [AS]. One property of φ_c which is important in our development is the following:

Lemma 2.1. *Let φ_c be as above. For any $\nu \in \mathbf{Z}_+^d$ with $|\nu|_1 =: n + 1 > \lambda + d + 1$,*

$$D^\nu \varphi_c \in L_p(\mathbf{R}^d), \qquad 1 \le p \le \infty.$$

Proof. It is sufficient to prove that the Fourier transform of $(\ )^\alpha D^\nu \varphi_c : x \mapsto x^\alpha D^\nu \varphi_c(x)$ with $|\alpha|_1 = d + 1$ is in $L_1(\mathbf{R}^d)$. Note that

$$[(\ )^\alpha D^\nu \varphi_c]^\wedge = (-1)^{d+1} i^{n-d} D^\alpha((\ )^\nu \hat{\varphi}_c).$$

Hence by using Leibnitz' rule, we will show that

$$D^\gamma(\ )^\nu D^{\alpha-\gamma}\hat{\varphi}_c \in L_1(\mathbf{R}^d). \tag{2.3}$$

Since $D^{\alpha-\gamma}\hat{\varphi}_c$ decays fast around ∞, the function in (2.3) is in $L_1(N_\infty)$ for some neighborhood N_∞ of ∞. Next, from (2.1), we see that the distribution $D^{\alpha-\gamma}\hat{\varphi}_c$ has a singularity of order $\lambda + 2d + 1 - |\gamma|_1$ at the origin and $D^\gamma(\ )^\nu$ has a zero of order $n + 1 - |\gamma|_1$ at the origin. Thus, we find that the function in (2.3) has a singularity of order $\lambda + 2d - n$, and $\lambda + 2d - n < d$ by assumption. It implies that $D^\gamma(\ )^\nu D^{\alpha-\gamma}\hat{\varphi}_c$ is in $L_1(N_0)$ with N_0 a neighborhood at the origin. ∎

Now, let us turn to the construction of the approximation scheme R_Ξ in (1.7). Suppose that we look for an approximant in the span of $\varphi_c(\cdot - t)$, $t \in \mathbf{R}^d$, for a smooth function $f\colon \mathbf{R}^d \to \mathbf{R}$. Then we may try to find the exact solution f^* of the convolution equation

$$\varphi_c * f^* = f.$$

However, this equation is not always solvable. In order to make sure that there is a solution for this equation, the function f needs to be very smooth in a way that depends on the basis function φ_c. Thus, rather than solving the equation exactly, we approximate first the function f by

$$\sigma_\omega^\vee * f$$

where ω is a parameter depending on h, i.e.,

$$\omega := \omega(h),$$

$\sigma_\omega : x \mapsto \sigma(\omega x)$, and $\sigma : \mathbf{R}^d \to [0,1]$ is a nonnegative C^∞-cutoff function whose support σ lies in the ball $B_\eta := \{x \in \mathbf{R}^d\colon |x| < \eta\} \subset [-2\pi, 2\pi]^d$ with $\sigma = 1$ on $B_{\eta/2}$ and $\|\sigma\|_\infty = 1$. Then we look for an approximant from the space $S_\Xi(\varphi_c)$ which approximates $\sigma_\omega^\vee * f$.

Since the function $\sigma_\omega^\vee * f$ is band-limited, after substituting $\sigma_\omega^\vee * f$ for f in the above convolution equation, we can find a solution

$$f^* = (\sigma_\omega \hat{f} / \hat{\varphi}_c)^\vee$$

of the following equation

$$\varphi_c * f^* = \sigma_\omega^\vee * f. \tag{2.4}$$

Since our real intent is to approximate the function $\sigma_\omega^\vee * f$ from the space $S_\Xi(\varphi_c)$, a natural way to construct an approximant from the space $S_\Xi(\varphi_c)$ can be to replace the kernel $\varphi(\cdot - t)$ in the convolution equation $\varphi_c * f^*$ in (2.4) by the approximation $\varphi_c(\cdot, t)$ in (1.8). However, a close look at the left-hand side of the above expression (2.4) shows that this attempt encounters several obstacles. First, the basis function φ_c grows at some polynomial degree away from zero, and there are inherent "cancellations;" hence loss of significance in the integration. Furthermore, in order for the above integration to make sense, we need to impose some extra conditions on f. (Namely, the function f is required to satisfy the condition $\hat{f} \in C^k(\mathbf{R}^d)$, $k > d + \lambda$.) The standard way to circumvent those difficulties is via a "localization process." Our strategy is to localize the kernel $\varphi_c(\cdot - t)$ in the above convolution equation (2.4) by applying a difference operator to φ_c, which constructs a new bell-shaped kernel

$$\psi_c = \sum_{\alpha \in \mathbf{Z}^d} \mu(\alpha) \varphi_c(\cdot - \alpha)$$

where the localization sequence $\mu\colon \mathbf{Z}^d \to \mathbf{C}$ decays fast at ∞ and the above sum converges uniformly on compact sets. In fact, μ is chosen to have finite support in this paper and the function ψ_c is assumed to satisfy the condition

$$\sup_{x \in \mathbf{R}^d} (1 + |x|)^{m_{\psi_c}} |\psi_c(x)| < \infty$$

for some $m_{\psi_c} > d$. The ability to perform this localization process is due to the structure of the Fourier transform of φ_c. The crucial fact is that the Fourier transform $\hat{\varphi}_c$ of φ_c is very smooth off the origin (see [DJLR]). This means that in order to localize φ_c we only need to make sure that the Fourier transform $\hat{\psi}_c$ of ψ_c is smooth at the origin. To ensure numerical stability, we need to insist that

$$\hat{\psi}_c(0) \neq 0.$$

In other words, considering the localization condition on ψ_c, the function $\hat{\psi}_c$ is continuous everywhere, especially at the origin. Hence, the function

$$\tau := \sum_{\alpha \in \mathbf{Z}^d} \mu(\alpha) e_{-i\alpha}$$

with $\hat{\psi}_c = \tau \hat{\varphi}_c$ has a high-order zero at the origin. We note here that τ is a $2\pi \mathbf{Z}^d$-periodic function, and since the only singularity of $\hat{\varphi}_c$ is at the origin and $\hat{\varphi}_c \neq 0$ on $\mathbf{R}^d \backslash 0$, we can assume that τ does not vanish on some punctured neighborhood $\Omega \backslash 0$ of the origin. This ensures that $\hat{\psi}_c$ does not vanish on Ω. Expressing the inverse Fourier transform of τ as

$$\tau^{\vee} = \sum_{\alpha \in \mathbf{Z}^d} \mu(\alpha) \delta_\alpha, \tag{2.5}$$

the above convolution equation (2.4) implies the relation

$$\sigma_\omega^{\vee} * f = \varphi_c * \tau(\omega \cdot)^{\vee} * \left(\frac{\sigma_\omega \hat{f}}{\tau(\omega \cdot) \hat{\varphi}_c} \right)^{\vee} \tag{2.6}$$

in which the property $(gh)^{\vee} = g^{\vee} * h^{\vee}$ is used. Invoking the relation in (2.1), we obtain

$$\tau(\omega \cdot) \hat{\varphi}_c = \omega^{\lambda + d} \hat{\psi}_{c/\omega}(\omega \cdot). \tag{2.7}$$

Denoting

$$\rho := c/\omega,$$

the expression (2.7) leads to

$$\tau(\omega \cdot)^{\vee} * \varphi_c = \omega^{\lambda} \psi_\rho(\cdot/\omega). \tag{2.8}$$

Hence, from (2.6) and (2.8), a direct calculation using change of variables yields that

$$\begin{aligned} \sigma_\omega^{\vee} * f &= \omega^{-d} \psi_\rho(\cdot/\omega) * \left(\frac{\sigma_\omega \hat{f}}{\hat{\psi}_\rho(\omega \cdot)} \right)^{\vee} \\ &= \left[\psi_\rho * \left(\frac{\sigma_\omega \hat{f}}{\hat{\psi}_\rho(\omega \cdot)} \right)^{\vee} (\omega \cdot) \right] (\cdot/\omega). \end{aligned}$$

Thus, approximating the convolution kernel $\psi_\rho(\cdot/\omega - t)$ in the above identity by

$$K(\cdot, t) := \sum_{\alpha \in \mathbf{Z}^d} \mu(\alpha) \varphi_c(\cdot, (t + \alpha)\omega)/\omega^{\lambda}$$

with $\varphi_c(\cdot, t)$ in (1.8), we obtain the following scheme:

Definition 2.2. With φ_c, ψ_c, and $K(\cdot, t)$ as above, we define our approximation scheme R_Ξ by

$$R_\Xi f := \int_{\mathbf{R}^d} K(\cdot, t)\Lambda f(\omega t)\, dt, \tag{2.9}$$

where Λ is the operator

$$\Lambda\colon\ f \mapsto (\sigma_\omega/\hat{\psi}_\rho(\omega\cdot))^\vee * f. \tag{2.10}$$

Remark. From (2.10), we observe that

$$\Lambda f = \int_{\mathbf{R}^d} (\sigma/\hat{\psi}_\rho)^\vee(\cdot/\omega - y)\, f(\omega y)\, dy. \tag{2.11}$$

Here, we find that $\sigma/\hat{\psi}_\rho \in C^{2\lambda+2d-1}(\mathbf{R}^d)$ (see [DJLR]) and $(\sigma/\hat{\psi}_\rho)^\vee$ decays at some polynomial degree. Thus, we realize that the localization property of R_Ξ is due to the decaying properties of the kernel K and $(\sigma/\hat{\psi}_\rho)^\vee$ in the sense that the contribution to the approximant's value at a point x by the data value at $\xi \in \Xi$ decreases as the distance between x and ξ increases. In addition, the definition of $K(\cdot, t)$ leads to the explicit form

$$R_\Xi f(x) = \sum_{\xi\in\Xi} \varphi_c(x - \xi) \sum_{\alpha\in\mathbf{Z}^d} \mu(\alpha) c_{\xi,\alpha}(f),$$

where

$$c_{\xi,\alpha}(f) := \int_{\mathbf{R}^d} A(\omega(t + \alpha), \xi)\Lambda f(\omega t)\, dt/\omega^\lambda$$

with Λ in (2.10). It ensures that the approximant $R_\Xi f$ belongs to $S_\Xi(\varphi_c)$.

Remark. The scheme R_Ξ has a smoothing effect through the convolution

$$\Lambda f = (\sigma_\omega/\hat{\psi}_\rho(\omega\cdot)) * f,$$

in (2.9). The actual smoothing parameters are c and ω; they are adjusted according to the density of centers and the level of the noise. As $c, \omega \to 0$, the function Λf converges to a local interpolant of f. On the other hand, as c grows, the approximant becomes smoother; hence, it may lose some "details." A good choice for the parameters c and ω can be interpreted as a balanced compromise between smoothness and fidelity of the approximation to the data. Practical examples are discussed in [Y].

2.2. *The Admissible Coefficients* $(A(t, \xi))_{\xi\in\Xi}$ *for* $\varphi_c(\cdot, t)$

For a given discrete set Ξ in $\mathbf{R}^d$, our approximation scheme R_Ξ in (1.7) is based on the approximation of each shift $\varphi_c(\cdot - t)$, $t \in \mathbf{R}^d$, by a linear combination

$$\varphi_c(\cdot, t) := \sum_{\xi\in\Xi} A(t, \xi)\varphi_c(\cdot - \xi). \tag{2.12}$$

We refer to $\varphi_c(\cdot, t)$ as a "pseudo-shift" of φ_c. For every $t \in \mathbf{R}^d$, the coefficients $(A(t,\xi))_{\xi\in\Xi}$ for $\varphi_c(\cdot, t)$ are assumed to satisfy the following condition: for some $m_A > d$, there exists a constant c independent of x and t such that

$$|\varphi_c(x-t) - \varphi_c(x,t)| \le c(1+|x-t|)^{-m_A}, \qquad x \in \mathbf{R}^d. \tag{2.13}$$

This condition pertains to the most fundamental property of the basis functions φ_c that we study; while φ_c itself grows at ∞, a suitable linear combination of translates of φ_c should decay at ∞.

Theorem 2.3. *Let $(A(t,\xi))_{\xi\in\Xi}$, $t \in \mathbf{R}^d$, be the coefficients for $\varphi_c(\cdot, t)$ in* (2.12). *Assume that*:

(a) *the set $\{A(t,\cdot)(t-\cdot)^j\colon t \in \mathbf{R}^d\}$ of functions defined on Ξ lies in $\ell_1(\Xi)$ and is bounded there for all $j < s$ for some nonnegative integer s; and*
(b) *for all $p \in \Pi_n$ with $n \in [\lambda + d, s)$ ($\lambda + d$ is the order of singularity of $\hat{\varphi}_c$ at the origin), the coefficients $(A(t,\xi))_{\xi\in\Xi}$ satisfy*

$$\sum_{\xi\in\Xi} A(t,\xi)p(\xi) = p(t).$$

Then we have the relation

$$|\varphi_c(x-t) - \varphi_c(x,t)| \le c(1+|x-t|)^{-m_A}$$

with $m_A = n - \lambda > d$ and c a constant independent of x and t.

Indeed, once the following simple lemma is established, this theorem is proved directly by Theorem 2.7.1 of [DR]:

Lemma 2.4. *The relation $\sum_{\xi\in\Xi} A(t,\xi)p(\xi) = p(t)$ holds for every $p \in \Pi_n$ if and only if*

$$\sum_{\xi\in\Xi} A(t,\xi)(t-\xi)^\alpha = \delta_{\alpha,0}. \tag{2.14}$$

Proof of the Lemma. The "only if" implication is trivial, since $(\cdot - \xi)^\alpha$ is a polynomial. As for the "if" implication, assuming (2.14) to hold, it is clear that

$$\sum_{\xi\in\Xi} A(t,\xi)p(\xi - t) = p(0)$$

for any $p \in \Pi_n$. This implies that

$$\begin{aligned} \sum_{\xi\in\Xi} A(t,\xi)p(\xi) &= \sum_{|\alpha|_1\le n} t^\alpha \sum_{\xi\in\Xi} A(t,\xi)D^\alpha p(\xi - t)/\alpha! \\ &= \sum_{|\alpha|_1\le n} t^\alpha D^\alpha p(0)/\alpha! = p(t), \end{aligned}$$

which completes our proof. ∎

For the given basis function φ_c, Theorem 2.3 states sufficient conditions on $(A(t,\xi))_{\xi\in\Xi}$ that imply (2.13). The highlight of these sufficient conditions is that they are actually independent of the basis function φ_c: the only information required on the basis function φ_c is the order of singularity of the Fourier transform $\hat{\varphi}_c$ at the origin. If a different basis function (e.g., the surface spline of order λ) has the same order of singularity (at the origin on the Fourier domain), we can use the same coefficient sequence $(A(t,\xi))_{\xi\in\Xi}$ to construct its pseudo-shifts. The general conditions of $(A(t,\xi))_{\xi\in\Xi}$ and basis functions for the successful construction of $\varphi_c(\cdot,t)$ in the sense of (2.13) are studied in the papers [DR] and [Y].

Here and in the sequel, we assume (without much loss) that, for any fixed $t\in\mathbf{R}^d$, the sequence $(A(t,\xi))_{\xi\in\Xi}$ is finitely supported, and we use the abbreviation

$$\Xi_t := \{\xi\in\Xi\colon\ A(t,\xi)\neq 0\}.$$

Of course, we choose the centers in the set Ξ_t to be some "close neighbors" of t. It is natural to require Ξ_t to have the nondegeneracy property for Π_n, i.e., any polynomial in Π_n which vanishes on Ξ_t must be identically zero. This also implies that the number of centers in Ξ_t should be no smaller than

$$\dim \Pi_n(\mathbf{R}^d) = \frac{(n+d)!}{n!\,d!}.$$

In view of the above discussion, we introduce the notion of "admissible coefficients" $(A(\cdot,\xi))_{\xi\in\Xi}$.

Definition 2.5. The coefficients $(A(\cdot,\xi))_{\xi\in\Xi}$ are termed **admissible** for Π_n if they satisfy the following three conditions:

(a) there exists $c_1>0$ such that, for any $t\in\mathbf{R}^d$, $A(t,\xi)=0$ whenever $|t-\xi|>c_1 h$, with h the density of Ξ as in (1.3);
(b) the set $\{(A(t,\xi))_{\xi\in\Xi}\colon\ t\in\mathbf{R}^d\}$ is bounded in $\ell_1(\Xi)$; and
(c) for every $t\in\mathbf{R}^d$, $\sum_{\xi\in\Xi} A(t,\xi)\delta_\xi=\delta_t$ on Π_n, i.e.,

$$\sum_{\xi\in\Xi} A(t,\xi)\,p(\xi) = p(t), \qquad \forall\, p\in\Pi_n. \tag{2.15}$$

Remark. When the coefficients $(A(\cdot,\xi))_{\xi\in\Xi}$ are admissible for Π_n, we realize that their key property is local reproduction of polynomials in Π_n. It is an important ingredient in our error estimates in the following Section 3. We also note that the linear system in (2.15) is invariant under the dilation and translation on $\mathbf{R}^d$ and Ξ. Hence, without loss of generality, we assume that the following condition holds in this study:

$$(A(\delta t,\xi))_{\xi\in\Xi} = (A(t,\xi/\delta))_{\xi\in\Xi}, \qquad \delta>0.$$

3. The Approximation Power of the Scheme R_Ξ

3.1. *Basic Results*

As we discussed in the previous sections, our approximation is performed in two steps: first a function f is approximated by $\sigma_\omega^\vee * f$, and then we approximate this mollified function by the quasi-interpolant $R_\Xi f$. For this reason, in error analysis, it is useful to divide $f - R_\Xi f$ into two parts

$$f - R_\Xi f = (f - \sigma_\omega^\vee * f) + (\sigma_\omega^\vee * f - R_\Xi f). \tag{3.1}$$

To estimate $\sigma_\omega^\vee * f - R_\Xi f$, we first consider the following lemma:

Lemma 3.1. *Let R_Ξ be the scheme defined as in* (2.9). *Assume that the coefficients $(A(t,\xi))_{\xi\in\Xi}$ for $\varphi_c(\cdot,t)$ in* (1.8) *are admissible for Π_n with $n \ge \lambda + d$. Then, for every $f \in L_1(\mathbf{R}^d)$, we have the identity*

$$\sigma_\omega^\vee * f - R_\Xi f = \int_{\mathbf{R}^d} (\varphi_c(\cdot - t) - \varphi_c(\cdot, t))(\sigma_\omega \hat{f}/\hat{\varphi}_c)^\vee \, dt. \tag{3.2}$$

Proof. From (2.8), let us recall the equation

$$\psi_\rho(\cdot/\omega) = \omega^{-\lambda}\tau(\omega\cdot)^\vee * \varphi_c$$

with $\rho = c/\omega$. Then, by (2.7) and change of variables, we deduce that

$$\begin{aligned}(\sigma_\omega^\vee * f - R_\Xi f)(x) &= \omega^{-\lambda-d}\int_{\mathbf{R}^d} \tau(\omega\cdot)^\vee * (\varphi_c(x-\cdot) - \varphi_c(x,\cdot))(t)\Lambda f(t)\,dt \\ &= \omega^{-\lambda-d}\int_{\mathbf{R}^d} (\varphi_c(x-t) - \varphi_c(x,t))(\tau(\omega\cdot)^\vee * \Lambda f)(t)\,dt\end{aligned}$$

with Λ in (2.10). Using (2.7) again, we get

$$\omega^{-\lambda-d}\tau(\omega\cdot)^\vee * \Lambda f = (\sigma_\omega \hat{f}/\hat{\varphi}_c)^\vee,$$

which completes our proof. ■

Theorem 3.2. *Let R_Ξ be the scheme defined as in* (2.9). *Suppose the coefficients $(A(t,\xi))_{\xi\in\Xi}$ for $\varphi_c(\cdot,t)$ in* (1.8) *are admissible for Π_n with $n \ge \lambda + d$. Then, for every band-limited function $f \in L_1(\mathbf{R}^d)$, we have*

$$\|f - R_\Xi f\|_p \le \text{const}\, h^{n+1}.$$

Proof. For sufficiently small ω, the term $f - \sigma_\omega^\vee * f$ is identically zero because $\hat{f}$ is compactly supported. Thus, invoking (3.1), it suffices to estimate only the error $\sigma_\omega^\vee * f - R_\Xi f$. For this, we recall the condition $\sum_{\xi\in\Xi} A(t,\xi) = 1$ to derive

$$\varphi_c(x-t) - \varphi_c(x,t) = \sum_{\xi\in\Xi} A(t,\xi)(\varphi_c(x-t) - \varphi_c(x-\xi)).$$

Let $T_{x-t}\varphi_c$ be the Taylor expansion of degree n of φ_c about $x - t$. Then, due to the polynomial reproduction property $\sum_{\xi\in\Xi} A(\cdot,\xi)p(\xi) = p$ for any $p \in \Pi_n$, we get

$$\sum_{\xi\in\Xi} A(t,\xi)[\varphi_c(x-t) - T_{x-t}\varphi_c(x-\xi)] = \sum_{0<|\nu|_1\le n} D^\nu\varphi_c(x-t)\sum_{\xi\in\Xi} A(t,\xi)(t-\xi)^\nu/\nu! = 0.$$

Thus, it gives the identity

$$\varphi_c(x-t) - \varphi_c(x,t) = \sum_{\xi\in\Xi} A(t,\xi) R_n(t,\xi)$$

with the remainder in the integral form

$$R_n(t,\xi) := \sum_{|\nu|_1=n+1} \frac{(t-\xi)^\nu}{\nu!} \int_0^1 (n+1)(1-y)^n D^\nu \varphi_c(x-t+y(t-\xi))\, dy.$$

Moreover, since $A(t,\xi) = 0$ whenever $|t-\xi| \geq c_1 h$ for some constant $c_1 > 0$ (see Definition 2.5), we obtain the bound

$$\sum_{\xi\in\Xi} |A(t,\xi)||t-\xi|^\nu/\nu! \leq \text{const}\, h^{n+1} \sum_{\xi\in\Xi} |A(t,\xi)|$$

for $|\nu|_1 = n+1$. Therefore, it provides the inequality

$$\begin{aligned}|\varphi_c(x-t) - \varphi_c(x,t)| \;\leq\; & \text{const}\, h^{n+1} \sum_{\xi\in\Xi} |A(t,\xi)| \\ & \times \int_0^1 (n+1)(1-y)^n \sum_{|\nu|_1=n+1} |D^\nu \varphi_c(x-t+y(t-\xi))|\, dy.\end{aligned}$$

Now, in order to bound the error $\|f - R_\Xi f\|_p$ with $1 \leq p \leq \infty$, we apply the above inequality to the right-hand side of the identity (3.2). Then, a direct calculation using the Minkowski's inequality (for the case $1 \leq p < \infty$) yields

$$\|f - R_\Xi f\|_p \leq \text{const}\, h^{n+1} \|(\sigma_\omega \hat{f}/\hat{\varphi}_c)^\vee\|_1 |\varphi_c|_{n+1,p}, \qquad 1 \leq p \leq \infty. \tag{3.3}$$

Consequently, to complete the proof of the theorem, it remains to show that the term $\|(\sigma_\omega \hat{f}/\hat{\varphi}_c)^\vee\|_1$ is bounded by a constant independent of ω. Let σ_f be a compactly supported C^∞-cutoff function such that $\sigma_f = 1$ on the support of $\hat{f}$. Then, for sufficiently small ω, it is clear that

$$(\sigma_\omega \hat{f}/\hat{\varphi}_c)^\vee = (\sigma_f/\hat{\varphi}_c)^\vee * f,$$

and it follows

$$\|(\sigma_\omega \hat{f}/\hat{\varphi}_c)^\vee\|_1 \leq \|(\sigma_f/\hat{\varphi}_c)^\vee\|_1 \|f\|_1.$$

Combining this bound with (3.3), we establish the required result. ∎

We now estimate the error $f - \sigma_\omega^\vee * f$:

Lemma 3.3. *Let σ_ω be the cutoff function defined as in Section* 2.1. *Then, for any $\nu \in \mathbf{Z}_+^d$,*

$$\int_{\mathbf{R}^d} \theta^\nu \sigma_\omega^\vee(\theta)\, d\theta = \delta_{\nu,0}.$$

Proof. Using the property $\theta^\nu \sigma_\omega^\vee(\theta) = (-i\omega)^{|\nu|_1}(D^\nu\sigma(\omega\cdot))^\vee(\theta)$ for any $\nu \in \mathbf{Z}_+^d$, we have

$$\int_{\mathbf{R}^d} \theta^\nu \sigma_\omega^\vee(\theta)\, d\theta = (-i\omega)^{|\nu|_1}(D^\nu\sigma(\omega\cdot))(0) = \delta_{\nu,0}.$$

Lemma 3.4. *Let σ_ω be the cutoff function defined as in Section* 2.1. *Then, for every $f \in W_p^k(\mathbf{R}^d)$, we have the convergence property*

$$\| f - \sigma_\omega^\vee * f\|_p = o(\omega^k), \qquad 1 \le p \le \infty,$$

as ω tends to 0.

Proof. From Lemma 3.3, it is easy to check that $\int_{\mathbf{R}^d} \sigma_\omega^\vee(\theta)\, d\theta = 1$ for any $\omega > 0$. Then it leads to the identity

$$(f - \sigma_\omega^\vee * f)(t) = \int_{\mathbf{R}^d} \sigma_\omega^\vee(\theta)(f(t) - f(t-\theta))\, d\theta.$$

Here, Taylor expansion of $f(t-\theta)$ about t gives the expression

$$f(t) - f(t-\theta) = \sum_{0<|\nu|_1<k} (-\theta)^\nu D^\nu f(t)/\nu! + R_k f(t,\theta) \tag{3.4}$$

with

$$R_k f(t,\theta) = \sum_{|\nu|_1=k} (-\theta)^\nu \int_0^1 k(1-y)^{(k-1)} D^\nu f(t-y\theta)\, dy/\nu!.$$

Due to the fact $\int_{\mathbf{R}^d} \sigma_\omega^\vee(\theta)\theta^\nu\, d\theta = 0$ for $\nu \ne 0$ (see Lemma 3.3), we find that the integral of $\sigma_\omega^\vee$ multiplied by the first term in the right-hand side of (3.4) is identically zero. Thus, we get

$$\begin{aligned}(f - \sigma_\omega^\vee * f)(t) \;&=\; \int_{\mathbf{R}^d} \sigma_\omega^\vee(\theta) R_k f(t,\theta)\, d\theta \\ &=\; (i\omega)^k \int_{\mathbf{R}^d} \omega^{-d} \sum_{|\nu|_1=k} (D^\nu\sigma)^\vee(\theta/\omega) \\ &\qquad \times \int_0^1 k(1-y)^{(k-1)} D^\nu f(t-y\theta)\, dy\, d\theta/\nu!.\end{aligned}$$

One can in fact prove by using Minkowski's inequality that

$$\omega^{-k}\| f - \sigma_\omega^\vee * f\|_p \le \text{const} \sum_{|\nu|_1=k} \|D^\nu f\|_p \int_{\mathbf{R}^d} \omega^{-d} |(D^\nu\sigma)^\vee(\theta/\omega)|\, d\theta$$

with $1 \le p \le \infty$. Note that $(D^\nu\sigma)^\vee(0) = 0$ for $\nu \ne 0$. Also, for $\theta \ne 0$,

$$\omega^{-d}(D^\nu\sigma)^\vee(\theta/\omega) \to 0, \qquad \omega \to 0.$$

Therefore, this lemma is true by the Lebesgue Dominated Convergence Theorem. ■

Though asymptotic approximation properties are usually quantified by approximation orders, an error estimate can be carried out in terms of a user's requirement, say "tolerance." By taking a sufficiently small ω, we can make the error $f - \sigma_\omega^\vee * f$ small enough to satisfy a (given) tolerance. Then, with the fixed ω, an approximand is sampled densely enough to make the final error satisfy the required tolerance. The following corollary is related to this issue:

Corollary 3.5. *Let R_Ξ be the scheme defined as in* (2.9). *Suppose the coefficients $(A(t,\xi))_{\xi\in\Xi}$ for $\varphi_c(\cdot,t)$ in* (1.8) *are admissible for Π_n with $n \geq \lambda + d$. Then, for every $f \in W_p^k(\mathbf{R}^d) \cap L_1(\mathbf{R}^d)$, we have*

$$\| f - R_\Xi f \|_p \leq \text{const}\, h^{n+1} + o(\omega^k), \qquad 1 \leq p \leq \infty,$$

where const *is dependent on the parameter c in φ_c.*

3.2. *The Approximation Order of R_Ξ*

As a set Ξ becomes dense, we want $R_\Xi f$ to better approximate the approximand f. The basic criterion of the approximation properties of R_Ξ is that of approximation orders. In this section, we will observe that R_Ξ provides *spectral approximation orders*: the convergence rate of the scheme will be determined by the decay at infinity of $\hat f$ (i.e., by the smoothness of the approximand f).

As a matter of fact, remembering the expression

$$f - R_\Xi f = (\sigma_\omega^\vee * f - R_\Xi f) + (f - \sigma_\omega^\vee * f)$$

in (3.1), the error $f - \sigma_\omega^\vee * f$ is computed directly by Lemma 3.4 regardless of the scheme R_Ξ. Hence, we focus our estimate on the term $\sigma_\omega^\vee * f - R_\Xi f$. For this purpose, we recall the expression in (3.3):

$$\|\sigma_\omega^\vee * f - R_\Xi f\|_p \leq \text{const}\, h^{n+1} \| f^* \|_1 |\varphi_c|_{n+1,p}$$

with

$$f^* := \left(\frac{\sigma_\omega \hat f}{\hat\varphi_c} \right)^\vee. \tag{3.5}$$

Here, we see that the function f^* cannot be kept, in general, bounded as ω tends to zero because $\hat\varphi_c^{-1}(\theta)$ increases exponentially fast as $\theta \to \infty$ (see (2.1) and (2.2)). However, it will be shown in the following analysis that this phenomenon is overcome by choosing a fixed ratio $\rho = c/\omega > 0$.

Exploiting the relation $(gh)^\vee = g^\vee * h^\vee$, for a function $f \in L_1(\mathbf{R}^d)$, we have the identity

$$\begin{aligned} f^*(t) &= \frac{1}{c(\lambda,d)} \int_{\mathbf{R}^d} f(t-\theta) \left(\frac{\sigma_\omega |\cdot|^{\lambda+d}}{\tilde K_{(\lambda+d)/2}(c\cdot)} \right)^\vee (\theta)\, d\theta \\ &= \frac{1}{c(\lambda,d)} \int_{\mathbf{R}^d} f(t-c\theta) \left(\frac{\sigma_{\omega/c} |\cdot|^{\lambda+d}}{c^{\lambda+d} \tilde K_{(\lambda+d)/2}} \right)^\vee (\theta)\, d\theta \end{aligned} \tag{3.6}$$

where $c(\lambda, d)$ is in (2.1) and $\tilde{K}_\nu(|t|) := |t|^\nu K_\nu(|t|)$ with $K_\nu(|t|)$ the modified Bessel function of order ν, see (2.2). Now, let us first consider the case $f \in W_1^{\lambda+d}(\mathbf{R}^d)$. Denoting

$$p_{\lambda+d}(x) := \sum_{|\alpha|_1=\lambda+d} c_\alpha x^\alpha := |x|^{\lambda+d},$$

a variant of the expression in (3.6) is as follows:

$$f^* = p_{\lambda+d}(D) f * (\sigma_\omega / \tilde{K}_{(\lambda+d)/2}(c\cdot))^\vee.$$

It implies that

$$\| f^* \|_1 \le | f |_{1,\lambda+d} \| (\sigma_{1/\rho} / \tilde{K}_{(\lambda+d)/2})^\vee \|_1.$$

In a similar fashion, for a function $f \in W_1^k(\mathbf{R}^d)$ with $k < \lambda + d$, we have the identity

$$f^* = \sum_{|\alpha|_1=\lambda+d} c_\alpha \left(\frac{\sigma_\omega(\ )^\alpha}{\tilde{K}_{(\lambda+d)/2}(c\cdot))} \right)^\vee * f.$$

For each $\alpha \in \mathbf{Z}_+^d$ with $|\alpha|_1 = \lambda + d$, we can write $\alpha = \beta + (\alpha - \beta)$ with $|\beta|_1 = k < \lambda + d$ and $\beta_i \le \alpha_i$ $(i = 1, \dots, d)$. Then, it follows that

$$f^* = \sum_{|\alpha|_1=\lambda+d} c_\alpha (-i)^k D^\beta f * \left(\frac{\sigma_\omega(\ )^{\alpha-\beta}}{\tilde{K}_{(\lambda+d)/2}(c\cdot)} \right)^\vee. \tag{3.7}$$

We note that

$$\int_{\mathbf{R}^d} \left(\frac{\sigma_\omega(\ )^{\alpha-\beta}}{\tilde{K}_{(\lambda+d)/2}(c\cdot)} \right)^\vee (\theta)\, d\theta = 0,$$

and hence, it implies the equation

$$D^\beta f * \left(\frac{\sigma_\omega(\ )^{\alpha-\beta}}{\tilde{K}_{(\lambda+d)/2}(c\cdot)} \right)^\vee = \int_{\mathbf{R}^d} \left(\frac{\sigma_\omega(\ )^{\alpha-\beta}}{\tilde{K}_{(\lambda+d)/2}(c\cdot)} \right)^\vee (\theta)(D^\beta f(t-\theta) - D^\beta f(t))\, d\theta.$$

From (3.7), it leads to an estimate of f^* as follows:

$$\|c^{\lambda+d-k} f^*\|_1 \le \sum_{|\alpha|_1=\lambda+d} |c_\alpha| \int_{\mathbf{R}^d} c^{-d} \left| \left(\frac{\sigma_{1/\rho}(\ )^{\alpha-\beta}}{\tilde{K}_{(\lambda+d)/2}} \right)^\vee (\theta/c) \right| \|D^\beta f(\cdot - \theta) - D^\beta f\|_1\, d\theta.$$

We observe that the above integrand is zero if $\theta = 0$, and for $\theta \ne 0$,

$$c^{-d} \left(\frac{\sigma_{1/\rho}(\ )^{\alpha-\beta}}{\tilde{K}_{(\lambda+d)/2}} \right)^\vee (\theta/c) \to 0, \qquad c \to 0,$$

because $\tilde{K}_{(\lambda+d)/2} \in C^{\lambda+d-1}(\mathbf{R}^d)$. Thus, the Lebesgue Dominated Convergence Theorem implies that

$$\| f^* \|_1 = o(c^{k-\lambda-d}), \qquad c \to 0.$$

Therefore, the following lemma is established:

Lemma 3.6. *Let f^* be defined as in* (3.5). *Assume that the parameters c and ω satisfy the relation $c = \rho\omega$ for a fixed $\rho > 0$. Then, for every $f \in W_1^k(\mathbf{R}^d)$, we have*

$$\| f^* \|_1 \leq \text{const} \begin{cases} 1 & \text{if } k \geq \lambda + d, \\ o(c^{k-\lambda-d}) & \text{if } k < \lambda + d, \end{cases}$$

with const *independent of c and ω, but dependent on ρ.*

Now we are ready to present the following theorem:

Theorem 3.7. *Let R_Ξ be the scheme defined as in* (2.9). *Let the coefficients $(A(t,\xi))_{\xi\in\Xi}$ for $\varphi_c(\cdot,t)$ in* (1.8) *be admissible for Π_n with $n \geq \lambda + d$. Assume that $\omega(h) = h^r$ with $0 < r \leq 1$, and that $c = \rho\omega$ for a fixed $\rho > 0$. Then, for every function $f \in W_p^k(\mathbf{R}^d) \cap W_1^m(\mathbf{R}^d)$ with $m := \min(k, \lambda + d)$, we have*

$$\| f - R_\Xi f \|_p = o(h^{rk}) + \begin{cases} O(h^{(1-r)(n+1)+r(\lambda+d)}) & \text{if } k \geq \lambda + d, \\ o(h^{(1-r)(n+1)+rk}) & \text{if } k < \lambda + d. \end{cases}$$

Proof. Taking $\omega(h) = h^r$ with $0 < r \leq 1$, it is immediate from Lemma 3.4 that

$$\| f - \sigma_\omega^\vee * f \|_p = o(h^{rk})$$

for any function $f \in W_p^k(\mathbf{R}^d)$. Hence, invoking (3.1), it remains to estimate only the term $\sigma_\omega^\vee * f - R_\Xi f$. To this end, let us recall the inequality in (3.3):

$$\|\sigma_\omega^\vee * f - R_\Xi f\|_p \leq \text{const}\, h^{n+1} \| f^* \|_1 |\varphi_c|_{n+1,p}. \tag{3.8}$$

For any $\nu \in \mathbf{Z}^d{}_+$ with $|\nu|_1 = n + 1$, it is obvious that

$$D^\nu \varphi_c = c^\lambda D^\nu(\varphi_1(\cdot/c)) = c^{\lambda-n-1}(D^\nu\varphi_1)(\cdot/c).$$

It follows from Lemma 2.1 that, for any $|\nu|_1 = n + 1$,

$$\|D^\nu\varphi_c\|_p = c^{\lambda+d-n-1}\|D^\nu\varphi_1\|_p < \infty.$$

Therefore, choosing $c = \rho h^r$ with $0 < r \leq 1$, the inequality (3.8) implies the bound

$$\|\sigma_\omega^\vee * f - R_\Xi f\|_p \leq \text{const}\, h^{(1-r)(n+1)+r(\lambda+d)} \| f^* \|_1$$

with const independent of the parameters c and ω. Applying Lemma 3.6 to this inequality, we get the desired result. ∎

Remark. The reason for our choice $r \in (0, 1)$ in the above theorem is as follows: when $k \geq \lambda + d$ and $\omega(h) = h$, the approximation scheme becomes stationary; the approximation order is $\lambda + d$. However, the choice $\omega(h) = h^r$ with $0 < r < 1$ induces the nearly optimal approximation order $o(h^{rk})$ by taking sufficiently large n for a given

r. Of course, if r is getting closer to 1, practically, we need to solve larger linear systems to get the approximation power $o(h^{rk})$. In contrast, for the case $k < \lambda + d$, there is no advantage in the choice of $r \in (0, 1)$. Hence, taking $\omega(h) = h$ brings the approximation order $o(h^k)$, which is in fact the best possible convergence rate. Also, in the stationary case, the approximation power does not depend on the choice of the number n $(\geq \lambda + d)$.

Corollary 3.8. *Suppose that* $f \in W_p^k(\mathbf{R}^d) \cap W_1^{\lambda+d}(\mathbf{R}^d)$ *with* $k \geq \lambda + d$. *For a given* $r \in (0, 1)$, *let the number* n $(\geq \lambda + d)$ *be chosen to satisfy the condition* $(1 - r)(n + 1) + r(\lambda + d) > rk$. *Under the same conditions and notations of Theorem* 3.7, *we have*

$$\| f - R_\Xi f \|_p = o(h^{rk}).$$

Corollary 3.9. *Assume that* $\omega(h) = h$ *and* $f \in W_p^k(\mathbf{R}^d) \cap W_1^q(\mathbf{R}^d)$ *with* $q = \min(k, \lambda + d)$. *Then, under the same conditions and notations of Theorem* 3.7, *we have*

$$\| f - R_\Xi f \|_p = \begin{cases} O(h^{\lambda+d}) & \text{if } k \geq \lambda + d, \\ o(h^k) & \text{if } k < \lambda + d. \end{cases}$$

4. Numerical Results

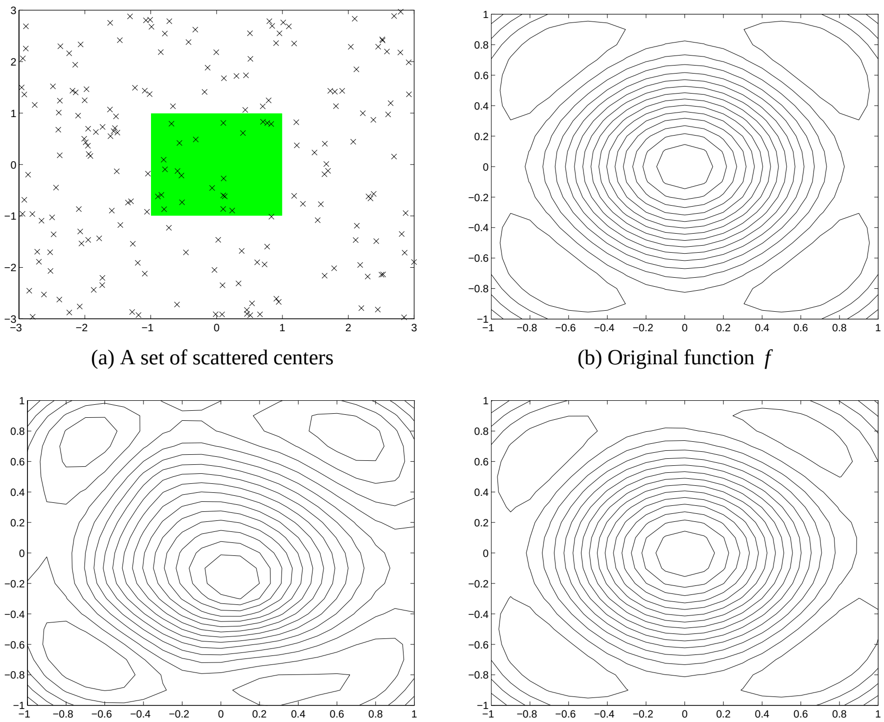
In this section, we provide a numerical example concerning our scheme R_Ξ. In this example, we approximate the function

$$f(x, y) = -\exp(-(x^2 + y^2)) + \left[\frac{\sin(x)\sin(y)}{xy}\right]^5. \tag{4.1}$$

Note that f is a smooth function, hence should be suitable for *interpolation methods*. Since interpolation methods have a known deteriorating performance near the boundary of the given domain, we confined our experiment to a "central portion" of the domain.

Here are the details. A set of 200 scattered centers Ξ is chosen randomly in the square $[-3, 3]$ (see Figure 4.1(a)). Then, we approximate the given function f in (3.9) from the space $S_\Xi(\varphi_c)$ (i.e., we employ a suitable scheme R_Ξ, as defined in (2.9) and used in the form given by the display following (2.11)). Finally, we measure the error only on $[-1, 1]^2$. We compare then the approximation with the benchmark method of the surface spline interpolation. Figures 4.1(b), (c), and (d) show the contour lines of the original function and the approximants by the surface spline interpolation and R_Ξ. The differences are obvious from the contour lines and the maximal absolute errors: 0.1682 by the surface spline interpolation and 0.0397 by R_Ξ. In addition, one major advantage of our scheme is that it is local in the sense that its value at a point mainly depends on values of $\varphi_c(\cdot - \xi)$, for those $\xi \in \Xi$ which are close to the point.

The above comparison is meaningful, since the two schemes, i.e., the surface spline interpolation and our scheme R_Ξ both select an approximant from a space spanned by the Ξ shifts of a basis function. Nonetheless, we need to stress that the above comparison, which heavily favors our scheme, is not completely "fair": the interpolation method uses as input only the scattered values of the given function at Ξ. Our method, in contrast, incorporates any needed information about f. We have conducted comparative



(a) A set of scattered centers (b) Original function f

(c) Surface spline interpolant (d) $R_\Xi f$

Fig. 4.1. The dark spot in (a) indicates the area on which we want to approximate. Note that the absolute maximal errors are 0.1682 by surface spline interpolation and 0.0397 by the scheme R_Ξ.

experiments where both schemes use only the values of f at Ξ; our approach in those experiments was still found to perform better than existing methods. In those other experiments, however, we have tackled the “boundary effect” problem as well, thereby implementing a more sophisticated version of the R_Ξ scheme.

In fact, the approximation scheme R_Ξ can be applied to noisy data as well as noiseless data. We refer the reader to [Y] for more examples, especially for noisy data approximation and the details of an algorithm. From the theory and experience gained with this approximation, and in comparison with other methods, we are convinced that the new scheme performs at least on a par with, and in many instances better than, the currently used methods.

Acknowledgments. I would like to thank Amos Ron for the advice and help which he offered during this work. This work was supported by Research Assistantships of Professor Ron at the University of Wisconsin-Madison (NSF Grant DMS-9626319 and the U.S. Army Research Office Contract DAAH04-95-1-0089). I also thank the referee who helped to make this paper more readable.

References

[A] M. A. AL-GWAIZ (1992): Theory of Distributions. New York: Marcel Dekker.

[AS] M. ABRAMOWITZ, I. STEGUN (1970): A Handbook of Mathematical Functions. New York: Dover.

[BR] C. DE BOOR, A. RON (1992): *Fourier analysis of the approximation power of principal shift-invariant spaces*. Constr. Approx. **8**:427–462.

[Bu] M. D. BUHMANN (1993): *New developments in the theory of radial basis functions interpolation*. In: Multivariate Approximation: From CAGD to Wavelets (K. Jetter and F. I. Utreras, eds.). Singapore: World Scientific, pp. 35–75.

[BuD] M. D. BUHMANN, N. DYN (1993): *Spectral convergence of multiquadric interpolation*. Proc. Edinburgh Math. Soc., **36**:319–333.

[BuDL] M. D. BUHMANN, N. DYN, D. LEVIN (1995): *On quasi-interpolation with radial basis functions with scattered centers*. Constr. Approx., **11**:239–254.

[BuR] M. D. BUHMANN, A. RON (1995): *Radial basis functions*: *L^p-approximation orders with scattered centers*. In: Wavelets, Images and Surface Fitting (P. J. Laurent, A Le Méhaué, and L. L. Schumaker, eds.). Boston: A. K. Peters, pp. 93–112.

[D] N. DYN (1989): *Interpolation and approximation by radial and related functions*. In: Approximation Theory VI (C. K. Chui, L. L. Schumaker, and J. Ward, eds.). New York: Academic Press, pp. 211–234.

[DJLR] N. DYN, I. R. H. JACKSON, D. LEVIN, A. RON (1992): *On multivariate approximation by integer translates of a basis function*. Israel J. Math., **78**:95–130.

[DR] N. DYN, A. RON (1995): *Radial basis function approximation*: *From gridded centers to scattered centers*. Proc. London Math. Soc., **71**:76–108.

[GMM] A. GRAY, G. B. MATHEWS, T. M. MACROBERT (1992): A Treatise on Bessel Function and Their Application to Physics. London: Macmillan.

[GS] I. M. GELFAND, G.E. SHILOV (1964): Generalized Functions, Vol. 1. New York: Academic Press.

[J] M. JOHNSON (1994): *Bound on the approximation order of surface splines*. Constr. Approx., **14**:429–438.

[M] C. A. MICCHELLI (1986): *Interpolation of scattered data*: *Distance matrices and conditionally positive functions*. Constr. Approx., **2**:11–22.

[MN1] W. R. MADYCH, S. A. NELSON (1988): *Multivariate interpolation and conditionally positive function I*. Approx. Theory Appl., **4**:77–89.

[MN2] W. R. MADYCH, S. A. NELSON (1990): *Multivariate interpolation and conditionally positive function II*. Math. Comp., **54**:211–230.

[P] M. J. D. POWELL (1992): *The theory of radial basis functions approximation in* 1990. In: Advances in Numerical Analysis, Vol. II: Wavelets, Subdivision Algorithms and Radial Basis Functions (W. A. Light, ed.). New York: Oxford University Press, pp. 105–210.

[R] A. RON (1994): *Negative observations concerning approximations from spaces generated by scattered shifts of functions vanishing at* ∞. J. Approx. Theory, **78**:364–372.

[Ru] W. RUDIN (1973): Functional Analysis, New York: McGraw-Hill.

[Y] J. YOON (1998): *Approximation to scattered data by radial basis function*. PhD thesis. Department of Mathematics, University of Wisconsin, Madison.

J. Yoon
Department of Applied Mathematics
School of Mathematical Sciences
Tel-Aviv University
Ramat Aviv, Tel-Aviv, 69978
Israel
yoon@post.tau.ac.il