ON SPHERICAL AVERAGES OF RADIAL BASIS FUNCTIONS

B. J. C. BAXTER

Dedicated to Arieh Iserles on the occasion of his 60^{th} birthday.

Abstract. A radial basis function (RBF) has the general form

$$s(x) = \sum_{k=1}^{n} a_k \phi(x - b_k), \qquad x \in \mathbb{R}^d,$$

where the coefficients $a_1, \ldots, a_n$ are real numbers, the points, or centres, $b_1, \ldots, b_n$ lie in $\mathbb{R}^d$, and $\phi : \mathbb{R}^d \to \mathbb{R}$ is a radially symmetric function. Such approximants are highly useful and enjoy rich theoretical properties; see, for instance, Buhmann [3], Fasshauer [6], Light and Cheney [11] or Wendland [19]. The important special case of *polyharmonic splines* results when ϕ is the fundamental solution of the iterated Laplacian operator, and this class includes the Euclidean norm $\phi(x) = \|x\|$ when d is an odd positive integer, the thin plate spline $\phi(x) = \|x\|^2 \log \|x\|$ when d is an even positive integer, and univariate splines. Now B-splines generate a compactly supported basis for univariate spline spaces, but an analyticity argument implies that a nontrivial polyharmonic spline generated by (1.1) cannot be compactly supported when $d > 1$. However, a pioneering paper of Jackson [8] established that the *spherical average* of a radial basis function generated by the Euclidean norm can be compactly supported when the centres and coefficients satisfy certain moment conditions; Jackson then used this compactly supported spherical average to construct approximate identities, with which he was then able to derive some of the earliest uniform convergence results for a class of radial basis functions. Our work extends this earlier analysis, but our technique is entirely novel, and applies to all polyharmonic splines. Furthermore, we observe that the technique provides yet another way to generate compactly supported, radially symmetric, positive definite functions. Specifically, we find that the spherical averaging operator commutes with the Fourier transform operator, and we are then able to identify Fourier transforms of compactly supported functions using the Paley–Wiener theorem. Furthermore, the use of Haar measure on compact Lie groups would not have occurred without frequent exposure to Arieh Iserles' study of geometric integration.

1. Introduction

A paper on radial basis functions (RBFs) might initially strike the reader as being out of place in this *Festschrift*, since Arieh Iserles has not worked on RBFs directly. However, his influence on generations of Cambridge numerical analysis researchers has been enormous. In particular, the RBF research of Michael Powell's group has been greatly enhanced by Arieh's breadth of knowledge and encouragement. In this note, I demonstrate that the Fourier transform of a RBF, viewed as a

1991 *Mathematics Subject Classification.* Primary: 41A30, 43A70; Secondary: 43A35.
Key words and phrases. Radial Basis Functions, Spherical Average, Haar Measure, Paley-Wiener.

tempered distribution, possesses a certain natural form, which results from the fundamental fact that the spherical averaging operator commutes with the Fourier transform operator. This fact, in itself a nontrivial result, will be established using an alternative definition of spherical averaging obtained via Haar measure on the orthogonal group. We then apply the Paley-Wiener theorem, which characterizes the Fourier transforms of compactly supported distributions as entire functions of exponential type, from which certain moment conditions emerge quite naturally. These topics are particularly relevant to this *Festschrift*, since it is very much Arieh's influence on my mathematical development which is here evident, from the theory of Lie groups to classical complex analysis.

A radial basis function has the general form

$$s(x) = \sum_{k=1}^{n} a_k \phi(x - b_k), \qquad x \in \mathbb{R}^d, \tag{1.1}$$

where the coefficients $a_1, \ldots, a_n$ are real numbers, the points, or centres, $b_1, \ldots, b_n$ lie in $\mathbb{R}^d$, and $\phi : \mathbb{R}^d \to \mathbb{R}$ is a radially symmetric function. Such approximants are highly useful and enjoy rich theoretical properties; see, for instance, Buhmann [3] or Light and Cheney [11]. The approximation theory community began to study RBFs during the mid-1980s, following the seminal work of Micchelli [13]. In the course of his fundamental work on the convergence properties of RBFs, Jackson [8] discovered the following remarkable fact: if $\phi(x) = \|x\|$, the Euclidean norm, then the *spherical average* of (1.1) is compactly supported when the dimension d is odd and the coefficients and centres satisfy the relations

$$\sum_{k=1}^{n} a_k \|b_k\|^{2\ell} = 0, \qquad \ell = 0, 1, \ldots, (d-1)/2, \tag{1.2}$$

the spherical average As being defined by

$$As(x) = \int_{S^{d-1}} s(\|x\|\theta)\, d\mu_d(\theta), \qquad x \in \mathbb{R}^d, \tag{1.3}$$

where μ_d denotes normalized $(d-1)$-dimensional Lebesgue measure on the unit sphere S^{d-1} in $\mathbb{R}^d$. Jackson's method was to expand the integrand of (1.3) for large $\|x\|$ and, in a *tour de force* of classical analysis, to identify the result as a certain hypergeometric function, from which he was then able to deduce the compact support of As when relations (1.2) hold. Jackson [8] was motivated by the construction of approximate identities using compactly supported spherical averages of radial basis functions; an ingenious construction, although it is not our primary interest in this paper. In contrast, the new technique presented here generalizes to *any* function $\phi : \mathbb{R}^2 \to \mathbb{R}$ whose (distributional) Fourier transform is of the form $\widehat{\phi}(\xi) = C\|\xi\|^{-2m}$, for some positive integer m. The class of all such functions is called the *polyharmonic splines*, following Madych and Nelson [12], for any such function is the fundamental solution of an iterated Laplacian operator. In particular, our analysis applies to the important case of *thin plate splines*, for which $\phi(x) = \|x\|^2 \log_e \|x\|$, for $x \in \mathbb{R}^2$. We also remark that a minor modification of our technique provides another way to construct compactly supported, radially symmetric, positive definite functions (cf. [19] for further details of such constructions).

2. Spherical averaging via the Fourier transform

Let $f : \mathbb{R}^d \to \mathbb{R}$ be any continuous function. Its *spherical average* $Af : \mathbb{R}^d \to \mathbb{R}$ is usually defined by the relation

$$Af(x) = \int_{S^{d-1}} f(\|x\|\theta)\, d\mu_d(\theta), \qquad x \in \mathbb{R}^d, \tag{2.1}$$

where μ_d denotes normalized Haar probability measure on the unit sphere $S^{d-1} = \{x \in \mathbb{R}^d : \|x\| = 1\}$; in other words, μ_d denotes ordinary $(d-1)$-dimensional Lebesgue measure scaled by the $(d-1)$-dimensional measure of the unit sphere S^{d-1}. Spherical averages have long had important applications in the theory of partial differential equations, as exemplified by the elegant monograph of John [9]. However, our calculations will be greatly simplified by use of the following equivalent definition.

Definition 2.1. *Let $f : \mathbb{R}^d \to \mathbb{R}$ be any continuous function. The* spherical average *$Af : \mathbb{R}^d \to \mathbb{R}$ can also be defined by the equation*

$$Af(x) = \int_{O(d)} f(Ux)\, d\sigma_d(U), \qquad x \in \mathbb{R}^d, \tag{2.2}$$

where $O(d)$ denotes the orthogonal group, that is,

$$O(d) = \{V \in \mathbb{R}^{d\times d} : V^T V = I\},$$

and σ_d denotes the normalized Haar probability measure on $O(d)$.

Haar measure σ_d on the compact metric group $O(d)$ is the unique measure possessing the invariance property

$$\sigma_d(Y) = \sigma_d(QY) = \sigma_d(YQ), \tag{2.3}$$

for any Borel subset $Y \subset O(d)$ and any $Q \in O(d)$, and further properties, together with an illuminating rigorous construction, may be found in the early chapters of Milman and Schechtman [14]. It may be easily checked that (2.3) implies the equivalence of (2.1) and (2.2).

Now Haar measure is, perhaps, somewhat unfamiliar to our audience, let us mention some salient facts for the convenience of the reader. We shall say that $M \in \mathbb{R}^{n\times n}$ is a *Gaussian random matrix* if its elements are independent Gaussian random variables with mean zero and unit variance; see, for instance, Baxter and Iserles [2] or Edelman and Raj Rao [5]. If we calculate the QR factorization of this Gaussian random matrix $M = QR$, where $Q \in \mathbb{R}^{n\times n}$ is orthogonal and $R \in \mathbb{R}^{n\times n}$ is upper triangular with positive diagonal entries, then Q is a (non-Gaussian) random matrix that is uniformly generated with respect to Haar measure on $O(d)$. Thus, generating N such orthogonal matrices $Q_1, \ldots, Q_N$ in this way, we can approximately integrate any continuous function $f : O(d) \to \mathbb{R}$ via the Monte Carlo sum

$$\int_{O(d)} f(Q)\sigma_d(Q) \approx N^{-1} \sum_{k=1}^{N} f(Q_k),$$

and the integral over the orthogonal group is the limit of this sum as $N \to \infty$.

All of the radial basis functions currently studied are continuous functions of at most polynomial growth, and their analysis makes great use of the Schwartz theory of tempered distributions. Our notation is fairly standard and follows that of Friedlander and Joshi [7] and Rudin [17]. In particular, we let $S(\mathbb{R}^d)$ denote the vector space of infinitely differentiable real-valued functions whose every derivative has supra-algebraic decay. This vector space becomes a locally convex topological vector space in the usual way, its dual $S'(\mathbb{R}^d)$ being the vector space of *tempered distributions*. In particular, every continuous function $\psi : \mathbb{R}^d \to \mathbb{R}$ of polynomial growth is a tempered distribution, and its action on $S(\mathbb{R}^d)$ is given by the integral relation

$$\langle \psi, f\rangle = \int_{\mathbb{R}^d} \psi(x) f(x)\, dx, \qquad f \in S(\mathbb{R}^d). \tag{2.4}$$

Further, the Fourier transform $\widehat{f}$ of $f \in S(\mathbb{R}^d)$ is defined by

$$\widehat{f}(\xi) = \int_{\mathbb{R}^d} f(x) \exp(-i\xi^T x)\, dx, \qquad \xi \in \mathbb{R}^d,$$

and the Fourier transform operator $F : f \mapsto \widehat{f}$ defines a linear bijection $F : S(\mathbb{R}^d) \to S(\mathbb{R}^d)$. The inverse Fourier transform is then given by the integral

$$f(x) = (2\pi)^{-d} \int_{\mathbb{R}^d} \widehat{f}(\xi) \exp(i\xi^T x)\, d\xi, \qquad x \in \mathbb{R}^d,$$

and the Fourier transform is defined on the dual space $S'(\mathbb{R}^d)$ by the requirement that

$$\langle F\psi, f\rangle = \langle \psi, Ff\rangle,$$

for any $f \in S(\mathbb{R}^d)$ and $\psi \in S'(\mathbb{R}^d)$. We refer the reader to Rudin [17] for further exposition.

Theorem 2.1. *The spherical averaging operator A and the Fourier transform operator F commute when applied to elements of the space $S(\mathbb{R}^d)$, that is we have the commutative diagram*

$$\begin{array}{ccc} S(\mathbb{R}^d) & \xrightarrow{A} & S_R(\mathbb{R}^d) \\ {\scriptstyle F}\downarrow & & {\scriptstyle F}\downarrow \\ S(\mathbb{R}^d) & \xrightarrow{A} & S_R(\mathbb{R}^d) \end{array}$$

where $S_R(\mathbb{R}^d)$ denotes the subspace of radially symmetric functions in $S(\mathbb{R}^d)$.

Proof. Applying Fubini's theorem yields the equations

$$\begin{aligned} \widehat{Af}(\xi) &= \int_{\mathbb{R}^d} \Big(\int_{O(d)} f(Ux)\, d\sigma_d(U) \Big) \exp(-i\xi^T x)\, dx \\ &= \int_{O(d)} \Big(\int_{\mathbb{R}^d} f(Ux) \exp(-i\xi^T x)\, dx \Big) d\sigma_d(U) \\ &= \int_{O(d)} \widehat{f}(U\xi)\, d\sigma_d(U) \\ &= A\widehat{f}(\xi). \qquad (2.5) \end{aligned}$$

Thus $F(Af) = A(Ff)$, for any $f \in S(\mathbb{R}^d)$. Finally, the well-known fact that the Fourier transform of any radially symmetric function is itself radially symmetric implies that F maps $S_R(\mathbb{R}^d)$ into itself.

□

The tempered distributional definition of the spherical averaging operator is defined via its action on $S(\mathbb{R}^d)$, that is

$$\langle A\psi, f\rangle := \langle \psi, Af\rangle, \qquad f \in S(\mathbb{R}^d), \psi \in S'(\mathbb{R}^d).$$

Now the Fourier transform is also defined on $S'(\mathbb{R}^d)$ via its action on $S(\mathbb{R}^d)$. More precisely, we have

$$\langle F\psi, g\rangle = \langle \psi, Fg\rangle, \qquad g \in S(\mathbb{R}^d),$$

for any $\psi \in S'(\mathbb{R}^d)$. We use the same trick to extend the definition of our spherical averaging operator A to $S'(\mathbb{R}^d)$, that is

$$\langle A\psi, g\rangle = \langle \psi, Ag\rangle, \qquad g \in S(\mathbb{R}^d).$$

Theorem 2.2. *The spherical averaging operator A and the Fourier transform operator F commute when applied to tempered distributions, that is, we obtain the commutative diagram*

$$\begin{array}{ccc} S'(\mathbb{R}^d) & \xrightarrow{A} & S'_R(\mathbb{R}^d) \\ {\scriptstyle F}\big\downarrow & & {\scriptstyle F}\big\downarrow \\ S'(\mathbb{R}^d) & \xrightarrow{A} & S'_R(\mathbb{R}^d) \end{array}$$

where $S'_R(\mathbb{R}^d)$ denotes the subspace of rotation-invariant tempered distributions.

Proof. This is merely diagram chasing:

$$\langle AF\psi, g\rangle = \langle F\psi, Ag\rangle = \langle \psi, F(Ag)\rangle = \langle \psi, A(Fg)\rangle = \langle A\psi, Fg\rangle = \langle F(A\psi), g\rangle.$$

□

3. Spherically averaging radial basis functions

Theorem 3.1. *Let $\phi : \mathbb{R}^d \to \mathbb{R}$ be a radially symmetric continuous function of polynomial growth and define*

$$s(x) = \sum_{k=1}^{n} a_k \phi(x - b_k), \qquad x \in \mathbb{R}^d.$$

Then

$$\widehat{As}(\xi) = \widehat{\phi}(\xi) \sum_{k=1}^{n} a_k \Omega_d(\|b_k\|\|\xi\|), \qquad \xi \in \mathbb{R}^d, \tag{3.1}$$

where $\Omega_d : \mathbb{R} \to \mathbb{R}$ is defined by

$$\Omega_d(t) = \widehat{\mu_d}(tu), \qquad t \in \mathbb{R}, \tag{3.2}$$

for any fixed unit vector $u \in \mathbb{R}^d$. Thus Ω_d is essentially the Fourier transform of the Haar probability measure on the sphere.

Proof. Since ϕ is a tempered distribution, we deduce

$$\widehat{s}(\xi) = \widehat{\phi}(\xi) \sum_{k=1}^{n} a_k \exp(-ib_k^T \xi), \qquad \xi \in \mathbb{R}^d,$$

and we recall that $\widehat{\phi}$ is a radially symmetric function. Now, the spherical average of the exponential

$$E_b(\xi) := \exp(ib^T \xi), \qquad \xi \in \mathbb{R}^d,$$

is the radially symmetric function defined by the integral

$$AE_b(\xi) = \int_{S^{d-1}} \exp(i\|\xi\| b^T \theta) d\mu_d(\theta).$$

Since we may rotate the coordinate system without modifying the value of this integral, we deduce

$$AE_b(\xi) = \int_{S^{d-1}} \exp(i\|\xi\| \|b\| u^T \theta) d\mu_d(\theta),$$

where u can be any unit vector. Thus we have

$$AE_b(\xi) = \Omega_d(\|b\|\|\xi\|), \qquad \xi \in \mathbb{R}^d,$$

and the spherical average is given by

$$\widehat{As}(\xi) = \widehat{\phi}(\xi) \sum_{k=1}^{n} a_k \Omega_d(\|b_k\|\|\xi\|).$$

□

We shall need the Paley-Wiener theorem in the form given by Rudin [17] and provide some background material for the reader.

Definition 3.1. *A continuous function $f\colon \mathbb{C}^d \to \mathbb{C}$ is said to be an entire function (of several complex variables) if the maps*

$$\{z \mapsto f(a_1, \ldots, a_{j-1}, z, a_{j+1}, \ldots, a_d) : z \in \mathbb{C}\}$$

are entire functions (of one complex variable) for any complex numbers $a_1, \ldots, a_d$ and $1 \le j \le d$.

Theorem 3.2 (Paley–Wiener)**.** *If μ is a signed measure on $\mathbb{R}^d$ supported by $\{x \in \mathbb{R}^d : \|x\| \le r\}$, then its Fourier transform $\widehat{\mu}$ is an entire function of exponential type r, that is*

$$\widehat{\mu}(z) \le C \exp(r|\operatorname{Im} z|), \qquad z \in \mathbb{C}^d,$$

for some constant C. Every entire function of exponential type r arises in this way.

Proof. This is Theorem 7.22 of Rudin [17]. □

The definition is a special case of Rudin [17], Definition 7.20, and

$$|\operatorname{Im} z| := \left((\operatorname{Im} z_1)^2 + \cdots + (\operatorname{Im} z_d)^2\right)^{1/2}, \qquad z = (z_1, \ldots, z_d) \in \mathbb{C}^d.$$

We have used the fact that a compactly supported distribution ψ of order zero can be identified with a signed measure. Indeed, we have the inequality $|\psi(f)| \le C\|f\|_\infty$ for every continuous function $f\colon K \to \mathbb{R}$, where K denotes the compact set supporting ψ. Thus ψ is a continuous linear functional on $(C(K), \|\cdot\|_\infty)$, and the Riesz representation theorem allows us to identify ψ with a measure μ via the formula

$$\psi(f) = \int_K f(x)\, d\mu(x), \qquad f \in C(K).$$

We observe that Ω_d is therefore an entire function of exponential type 1, because of (3.2). Further, since the Fourier transform of a rotation-invariant measure is radially symmetric, we deduce that Ω_d is an even function. Hence its Taylor series takes the form

$$\Omega_d(z) = \sum_{\ell=0}^{\infty} c_\ell^{(d)} z^{2\ell}, \qquad z \in \mathbb{C}. \tag{3.3}$$

Further, it can be shown that

$$\Omega_d(t) = \Gamma(d/2)(2/t)^{(d-2)/2} J_{(d-2)/2}(t);$$

see equation (1.8) of Schoenberg [18] or Section 41 of Donoghue [4]. Since not all readers will be fans of Bessel functions, let us mention that using spherical polar coordinates provides the alternative formula for Ω_d is

$$\Omega_d(t) = \frac{\widehat{f_d}(t)}{\widehat{f_d}(0)}, \qquad t \in \mathbb{R}^d,$$

where $f_d : \mathbb{R} \to \mathbb{R}$ is the compactly supported univariate function given by

$$f_d(x) = \begin{cases} (1-x^2)^{(d-3)/2} & \text{for } |x| \le 1, \\ 0 & \text{otherwise.} \end{cases}$$

Theorem 3.3. *Let $\phi : \mathbb{R}^d \to \mathbb{R}$ be a polyharmonic spline, that is, a tempered distribution whose Fourier transform takes the form*

$$\widehat{\phi}(\xi) = C\|\xi\|^{-2m}, \qquad \xi \in \mathbb{R}^d \setminus \{0\}, \tag{3.4}$$

for some positive integer m. Then the spherical average As of

$$s(x) = \sum_{k=1}^{n} a_k \phi(x - b_k), \qquad x \in \mathbb{R}^d. \tag{3.5}$$

is compactly supported if and only if

$$\sum_{k=1}^{n} a_k \|b_k\|^{2\ell} = 0, \qquad \text{for } \ell = 0, 1, \ldots, m-1. \tag{3.6}$$

Proof. We have

$$\begin{aligned} \widehat{As}(\xi) &= C\|\xi\|^{-2m} \sum_{k=1}^{n} a_k \Omega_d(\|b_k\|\|\xi\|) \\ &= C\|\xi\|^{-2m} \sum_{k=1}^{n} a_k \sum_{\ell=0}^{\infty} c_\ell^{(d)} \|b_k\|^{2\ell} \|\xi\|^{2\ell} \\ &= C\|\xi\|^{-2m} \sum_{\ell=0}^{\infty} c_\ell^{(d)} \|\xi\|^{2\ell} \Big(\sum_{k=1}^{n} a_k \|b_k\|^{2\ell} \Big) \\ &=: C\|\xi\|^{-2m} g(\xi). \end{aligned} \tag{3.7}$$

Thus $\widehat{As}$ is an entire function of exponential type 1 if and only if the entire function g possesses a zero at the origin of order $2m$, so cancelling the pole of order $2m$. Therefore we deduce that $\widehat{As}$ is an entire function of exponential type, and hence As is compactly supported, if and only if the first $2m$ coefficients of the Taylor series of g vanish, that is,

$$\sum_{k=1}^{n} a_k \|b_k\|^{2\ell} = 0, \qquad \ell = 0, 1, \ldots, m-1.$$

□

Finally, we observe that an easy modification of the above construction yields another way to construct compactly supported positive definite functions. Indeed, if we define $s : \mathbb{R}^d \to \mathbb{R}$ by (3.5) and (3.6), then the convolution of its spherical average with itself, that is,

$$\psi(x) = (As) * (As)(x), \qquad \text{for } x \in \mathbb{R}^d,$$

is a compactly supported radially symmetric function. Further, it is also a positive definite function, since its Fourier transform satisfies $\widehat{\psi}(\xi) = |\widehat{\phi}(\xi)|^2$, for all $\xi \in \mathbb{R}^d$. Furthermore, if we add the extra moment condition

$$\sum_{k=1}^{n} a_k \|b_k\|^{2m} \neq 0, \tag{3.8}$$

then we deduce that $\widehat{\psi}(0) > 0$, which implies that it is a strictly positive definite function.

4. Applications

We shall now address the important special cases of the Euclidean norm $\phi(x) = \|x\|$, for $x \in \mathbb{R}^d$ and d odd, and the thin plate spline $\phi(x) = \|x\|^2 \log_e \|x\|$, for $x \in \mathbb{R}^d$ and d even, for which we first establish that these are, indeed, polyharmonic radial basis functions. It is usual to refer the reader to standard texts such as Jones [10] for the explicit formulae giving the Fourier transforms of these tempered distributions. However, we prefer to sketch a simple derivation based on the Schoenberg theory of positive definite functions for completeness. Further details are given in the excellent textbooks Fasshauer [6] and Wendland [19].

The following integrals occur in almost all calculations of this form.

Lemma 4.1. *Define*

$$I(\lambda, \mu) = \int_0^\infty e^{-\lambda/t} t^{-\mu}\, dt, \qquad \text{for } \lambda > 0, \mu > 1. \tag{4.1}$$

Then

$$I(\lambda, \mu) = \frac{\Gamma(\mu - 1)}{\lambda^{\mu-1}}. \tag{4.2}$$

Proof. We simply set $\tau = \lambda/t$; the convergence of the integrals is elementary. □

Lemma 4.2. *We have*

$$\sum_{j,k=1}^{n} y_j y_k e^{-t\|x_j - x_k\|^2} = (2\pi)^{-d} \int_{\mathbb{R}^d} \Big| \sum_{k=1}^{n} y_k e^{i x_k^T \xi} \Big|^2 (\pi/t)^{d/2} e^{-\|\xi\|^2/4t}\, dt. \tag{4.3}$$

Proof. Simply apply the Fourier inversion theorem. □

4.1. **The Euclidean norm.** The Euclidean norm is given by the dimension-independent formula

$$\phi(x) = \int_0^\infty \Big(1 - e^{-t\|x\|^2}\Big)(4\pi)^{-1/2} t^{-3/2}\, dt. \tag{4.4}$$

We obtain this expression by setting $\tau = t\|x\|^2$ in the Gamma function integral

$$\Gamma(-1/2) = \int_0^\infty \Big(e^{-t} - 1\Big) t^{-3/2}\, dt,$$

which is an analytic continuation of the usual integral relation; see, for instance, Whittaker and Watson [20], Section 12.21.

If $y_1, \ldots, y_n \in \mathbb{R}$ satisfy $\sum_{j=1}^n y_j = 0$, then (4.4) implies the relation

$$\sum_{j,k=1}^n y_j y_k \|x_j - x_k\| = -\int_0^\infty \Big(\sum_{j,k=1}^n y_j y_k e^{-t\|x_j - x_k\|^2}\Big)(4\pi)^{-1/2} t^{-3/2}\, dt. \tag{4.5}$$

Applying the Fourier inversion theorem to the Gaussian quadratic form and swapping the order of integration (which is valid by Fubini's theorem), we obtain

$$\sum_{j,k=1}^n y_j y_k \|x_j - x_k\| = (2\pi)^{-d} \int_{\mathbb{R}^d} \Big|\sum_{k=1}^n y_k e^{i x_k^T \xi}\Big|^2 \psi_d(\xi)\, d\xi, \tag{4.6}$$

where

$$\begin{aligned} \psi_d(\xi) &= -(4\pi)^{-1/2} \int_0^\infty e^{-\|\xi\|^2/4t} (\pi/t)^{d/2} t^{-3/2}\, dt \\ &= -(4\pi)^{-1/2} \pi^{d/2} I(\|\xi\|^2/4, (d+3)/2). \end{aligned} \tag{4.7}$$

Substituting (4.2) in (4.7) yields

$$\psi_d(\xi) = -\Big(\frac{2^d \pi^{(d-1)/2} \Gamma(\frac{d+1}{2})}{\|\xi\|^{d+1}}\Big). \tag{4.8}$$

All of this is entirely classical, which is useful when explaining Fourier transform arguments to audiences suspicious of distribution theory[1]. Of course, (4.8) is the Fourier transform of the Euclidean norm in d-dimensional space, that is,

$$\widehat{\phi}(\xi) = -\Big(\frac{2^d \pi^{(d-1)/2} \Gamma(\frac{d+1}{2})}{\|\xi\|^{d+1}}\Big). \tag{4.9}$$

We observe that $\widehat{\phi}$ extends to an analytic function on $\mathbb{C}^d \setminus \{0\}$ when d is odd, and $\widehat{\phi}$ would therefore be an exponential function of exponential type were it not for its pole at the origin, and this is the crux of our analysis. Hence Theorem 3.3 implies that, when d is odd, the spherical average of

$$s(x) = \sum_{k=1}^n a_k \|x - b_k\|, \qquad x \in \mathbb{R}^d,$$

is compactly supported if and only if

$$\sum_{k=1}^n a_k \|b_k\|^{2\ell} = 0, \qquad \text{for } \ell = 0, 1, \ldots, (d-1)/2.$$

[1]The integral formula (4.9) is also derived in Baxter [1] in a more general context.

Jackson's construction had, as its end, the construction of an approximate identity As, i.e. a spherical average that was compactly supported with nonzero integral. The last condition therefore requires the further moment condition

$$\sum_{k=1}^{n} a_k \|b_k\|^{d+1} \neq 0.$$

4.2. **The thin plate spline.** We shall find it convenient to use the slightly nonstandard definition $\phi(x) = \|x\|^2 \log \|x\|^2$ for the thin plate spline. However, since $\phi(x) = 2\|x\|^2 \log \|x\|$, there should be no confusion. The integral representation is

$$\phi(x) = \|x\|^2 - 1 + \int_0^\infty \Big(e^{-t\|x\|^2} - e^{-t} + t(\|x\|^2 - 1)e^{-t}\Big) t^{-2}\, dt. \tag{4.10}$$

We obtain (4.10) by setting $f(t) = t \log t$, using the formula

$$f''(t) = \frac{1}{t} = \int_0^\infty e^{-\alpha t}\, d\alpha$$

and integrating twice. Then $\phi(x) = f(\|x\|^2)$.

If $y_1, \ldots, y_n \in \mathbb{R}$ and $x_1, \ldots, x_n \in \mathbb{R}^d$ satisfy

$$\sum_{j=1}^{n} y_j = 0 \quad \text{and} \quad \sum_{j=1}^{n} y_j x_j = 0,$$

then (4.10) implies the quadratic form relation

$$\sum_{j,k=1}^{n} y_j y_k \phi(x_j - x_k) = \int_0^\infty \Big(\sum_{j,k=1}^{n} y_j y_k e^{-t\|x_j - x_k\|^2} \Big) t^{-2}\, dt. \tag{4.11}$$

Just as before, we obtain

$$\sum_{j,k=1}^{n} y_j y_k \phi(x_j - x_k) = (2\pi)^{-d} \int_{\mathbb{R}^d} \Big| \sum_{k=1}^{n} y_k e^{i x_k^T \xi} \Big|^2 \psi_d(\xi)\, d\xi, \tag{4.12}$$

where

$$\psi_d(\xi) = \int_0^\infty (\pi/t)^{d/2} e^{-\|\xi\|^2/4t} t^{-2}\, dt = \pi^{d/2} I(\|\xi\|^2/4, 2 + d/2). \tag{4.13}$$

Thus

$$\psi_d(\xi) = \frac{2^{d+2}\pi^{d/2}\Gamma(1 + d/2)}{\|\xi\|^{d+2}}. \tag{4.14}$$

As for the Euclidean norm, this is the Fourier transform of ϕ in d-dimensional space. For example

$$\psi_2(\xi) = \frac{16\pi}{\|\xi\|^4}. \tag{4.15}$$

We observe that, when the dimension d is even, $\widehat{\phi}$ would again be an entire function of exponential type were it not for the pole at the origin. Hence Theorem 3.3 implies that, for even dimension d,

$$s(x) = \sum_{k=1}^{n} a_k \|x - b_k\|^2 \log_e \|x - b_k\|, \qquad x \in \mathbb{R}^d,$$

has a compactly supported spherical average if and only if the moment conditions

$$\sum_{k=1}^{n} a_k \|b_k\|^{2\ell} = 0$$

are satisfied for $\ell \leq d/2$.

Acknowledgments

I am particularly grateful to the anonymous referees for their comments and suggestions.

References

[1] B. J. C. Baxter (1994), "Norm estimates for inverses of Toeplitz distance matrices", *J. Approx. Theory* **79**:222–242.

[2] B. J. C. Baxter and A. Iserles (2003), "On the foundations of computational mathematics", in Handbook of Numerical Analysis XI (P.G. Ciarlet and F. Cucker, eds), North-Holland, Amsterdam, 3–34.

[3] M. D. Buhmann (2003), *Radial Basis Functions: Theory and Implementations*, Cambridge Monographs on Applied and Computational Mathematics, Cambridge University Press.

[4] W. F. Donoghue, Jr., *Distributions and Fourier Transforms*, Academic Press, New York.

[5] A. Edelman and N. Raj Rao (2005), "Random matrix theory", *Acta Numerica*: **14**: 233–297.

[6] G. Fasshauer (2007), *Meshfree Approximation Methods with Matlab*, World Scientific Publishing.

[7] F. G. Friedlander and M. Joshi (1999), *Introduction to the Theory of Distributions*, Cambridge University Press.

[8] I. R. H. Jackson (1988), "Convergence Properties of Radial Basis Functions", *Constr. Approx.* **4**: 243–264.

[9] F. John (1955), *Plane waves and spherical means applied to partial differential equations*, Interscience, New York.

[10] D. S. Jones (1982), *The Theory of Generalised Functions*, Cambridge University Press.

[11] W. A. Light and E. W. Cheney (2000), *A Course in Approximation Theory*, Brooks/Cole.

[12] W. R. Madych and S. A. Nelson (1990), "Polyharmonic cardinal splines", *J. Approx. Theory* **60**: 141–156.

[13] C. A. Micchelli (1986), "Interpolation of scattered data: distance matrices and conditionally positive functions", *Constr. Approx.* **2**: 11-22.

[14] V. Milman and G. Schechtman (1986), *Asymptotic theory of finite dimensional normed spaces*, Lecture Notes in Mathematics 1200, Springer.

[15] T. Poggio and S. Smale (2003), "The Mathematics of Learning: Dealing with Data", Notices of the AMS, May 2003: 537–544.

[16] M. J. D. Powell (1992), "The Theory of Radial Basis Functions in 1990", in *Advances in Numerical Analysis*, vol. II, ed. W. A. Light, Oxford University Press, pp. 105–210.

[17] W. Rudin (1991), *Functional Analysis*, McGraw-Hill.

[18] I. J. Schoenberg (1938), "Metric spaces and completely monotone functions", *Ann. of Math.* **39**: 811-841.

[19] H. Wendland (2004), *Scattered Data Approximation*, Cambridge University Press.

[20] E. T. Whittaker and G. N. Watson (1927), *A Course of Modern Analysis*, Cambridge University Press.

[21] D. V. Widder (1946), *The Laplace Transform*, Princeton University Press.

School of Economics, Mathematics and Statistics,, Birkbeck College, University of London,, Malet Street, London WC1E 7HX, England

E-mail address: b.baxter@bbk.ac.uk