ROCKY MOUNTAIN
JOURNAL OF MATHEMATICS
Volume 24, Number 4, Fall 1994

AN INTERPOLATION SCHEME WITH RADIAL BASIS IN SOBOLEV SPACES $H^s(R^n)$

JULIO G. DIX AND ROBERT D. OGDEN

ABSTRACT. This article concerns a global interpolation scheme that uses the radial basis $\exp(-|x-x_i|)P(|x-x_i|)$, where $P(\cdot)$ is a polynomial. These basis functions belong to a Sobolev space $H^s(R^n)$ and correspond to evaluations of reproducing kernels at data points. To balance the visual smoothness of interpolating graphs and the condition number of a resulting linear system, we use a distance scaling factor that depends on the data. We find a rate of convergence for approximations and show results of numerical experiments.

0. Introduction. The interpolation problem that we are interested in consists of finding a function $\tilde{f}$ that assumes given real values at given points $\{x_i\}$ of R^n. This problem has many applications in real life. The intended use of interpolants determines the properties that $\tilde{f}$ should have. For instance, it should have continuous derivatives, or should be easily computable, or should be continuously dependent on the data, etc. A summary of desirable properties of interpolants is found in Grosse [**12**]. A list of interpolation methods is found in Afeld [**1**], and a comparison of methods is found in Franke [**6**].

An interpolation scheme can be global or local. It is global if the interpolant at any given point depends on all the data and local if it depends only on data "near" the given point. Each type of interpolation is suitable for some applications. But no one scheme is best for all applications. The scheme here is global but its basis can be truncated to have a local character.

An interpolation method is called radial if its basis functions depend on x only through the distance to data points, $|x-x_i|$. The general form of radial interpolants is

$$\tilde{f}(x)=\sum_i w_i\phi_i(|x-x_i|), \tag{0.1}$$

Received by the editors on February 11, 1992, and in revised form on June 28, 1993.
1980 *Mathematics Subject Classification.* 41A05, 41A63, 65D05, 65D15.
Key words. Interpolation, scattered data, radial functions, reproducing kernel.

where ϕ_i's are the basis functions and w_i's are constants to be found. The main feature of radial interpolation is that it does not become more elaborate as the dimension of the domain increases. Well known cases are $\phi(r) = r$, $r^2 \log r$, $\sqrt{r^2 + c^2}$, $1/\sqrt{r^2 + c^2}$, r^3, $\exp(-r^2)$, which are called linear, thin plate spline, multiquadric, inverse quadric, cubic and Gaussian basis respectively. $\phi(r) = r$ and r^3 also have been called multiconic and pseudocubic, respectively, by Duchon [**9**].

In order to facilitate the calculation of w_i's, equation (0.1) can be transformed into the cardinal form

$$\tilde{f}(x) = \sum_i f(x_i) C_i(x),$$

where $C_i(x_j) = \delta_{i,j}$ is the delta Kronecker. In this case the constants are the same as the data given, but cardinal bases may be unpractical for computations, see, e.g., [**1, 5**].

We show that when $\phi(r) = \exp(-r)P(r)$, scheme (0.1) is suitable for interpolating scattered data that decays to zero at infinity, that it is computable, and that it has adequate theoretical foundations. Although this basis function appeared in the meteorological literature as early as 1972, Franke [**10**] did not test this interpolation scheme in his report, and to the best of our knowledge they have not appeared in the interpolation literature.

The reviewer has brought the work of Madych and Nelson to our attention [**14**]. Many of the ideas independently developed by us were developed by Madych and Nelson in a more general setting. Kernels here are conditionally positive definite functions of order zero in [**14**]. However, the examples that we have in mind are different from their primary examples, although they are in a sense dual to them. Thus, our interpolants are rapidly decaying and belong to ordinary $L^2(\mathbf{R}^n)$, but they are of limited smoothness, whereas primary examples in [**14**] are infinitely differentiable but of polynomial order at infinity.

In Section 1 we state well-known properties of reproducing kernel Hilbert spaces and of Sobolev spaces. In Section 2 we calculate the reproducing kernel for the Sobolev space $H^s(R^n)$ when $s = m + (n+1)/2$. In Section 3 we reduce the problem of calculating the approximation $\tilde{f}$ to a linear least-squares problem. Then we estimate the rate at which approximations converge. In Section 4 we discuss computational aspects and derive formulas for the derivatives and the Fourier transform of approximations. We also discuss the role of a distance scaling factor and show some computer experiments.

1. Reproducing kernel Hilbert spaces (RKHS). A linear space H of functions defined on a set Ω is said to be an RKHS under the inner product $\langle \cdot, \cdot \rangle_H$ if the following is true. There exists a function called reproducing kernel $k : \Omega \times \Omega \to R$ such that for each x in Ω and f in H we have $k_x(t) := k(x,t)$ is in H and

$$\langle k_x, f \rangle_H = f(x).$$

Given an RKHS it is known that $k_x(\cdot)$ is unique, that for each $t \in \Omega$ the evaluation functional $f \mapsto f(t)$ is continuous, that norm convergence implies pointwise convergence, that reproducing kernels are symmetric and nonnegative definite. See, for example, [**2, 7, 17, 18**].

Some reproducing kernels are constructed from Green functions. For examples on finite domains see [**17**], and for an infinite domain consider the next example [personal communication with John Walsh]. The space $H^1(R)$ consisting of $L^2(R)$ functions whose first derivative is also in $L^2(R)$ is an RKHS with

$$k_x(t) = \frac{1}{2} \exp(-|t-x|).$$

This statement is proved using the fact that $k_x'(t) = k_x(t)$ for $t < x$, that $k_x'(t) = -k_x(t)$ for $x < t$, and that $\lim_{x\to\infty} e^{-x} f(x) = 0$. In fact

$$\begin{aligned}\langle k_x, f \rangle_{H^1} &= \langle k_x, f \rangle_{L^2} + \langle k_x', f' \rangle_{L^2} \\ &= \frac{1}{2}\int_{-\infty}^{x} e^{t-x} f(t)\,dt + \frac{1}{2}\int_{x}^{+\infty} e^{x-t} f(t)\,dt \\ &\quad + \frac{1}{2}\int_{-\infty}^{x} e^{t-x} f'(t)\,dt - \frac{1}{2}\int_{x}^{+\infty} e^{x-t} f'(t)\,dt \\ &= \frac{1}{2}\lim_{a\to-\infty} [e^{t-x} f(t)]_a^x - \frac{1}{2}\lim_{b\to+\infty} [e^{x-t} f(t)]_x^b = f(x).\end{aligned}$$

Note that $k_x(t)$ is the Green's function for the operator $(I - d^2/dt^2)$, in that

$$\frac{d^2}{dt^2} k_x(t) = k_x(t) - \delta(t-x).$$

Before defining Sobolev spaces we set some notation. The complex conjugate of a function f is denoted by $\bar{f}$. The inner product and norm in R^n are denoted by $y^T x$ and $|y|$, respectively. The set of tempered distributions on R^n is denoted by $\mathcal{S}'(R^n)$, and function spaces are considered to be vector subspaces of this space. The Fourier transform of a test function ϕ in $\mathcal{S}(R^n)$ is defined by

$$\mathcal{F}\phi(y) = \hat{\phi}(y) = \int_{R^n} e^{-2\pi i y^T x} \phi(x)\, dx.$$

The Fourier transform inverse is denoted by

$$\mathcal{F}^{-1}\phi(x) = \int_{R^n} e^{2\pi i y^T x} \phi(y)\, dy,$$

and the Fourier transform for a tempered distribution f by

$$\hat{f}(\phi) = f(\hat{\phi}).$$

The space L^2 can be defined as all tempered distributions f such that $|f(\phi)| \leq C_f ||\phi||_{L^2}$, where for a measurable function the L^2 norm is defined as usual, and C_f is a constant that depends on f. This definition is equivalent to the usual one of L^2 as equivalence classes of square-summable functions. The Fourier transform defines an automorphism of $\mathcal{S}(R^n)$ and therefore of $\mathcal{S}'(R^n)$, and when restricted to L^2 it is a unitary automorphism. Let s be a real number, and put

$$\eta(y) = (1 + |2\pi y|^2)^s. \tag{1.1}$$

Then the multiplication by η, $f \mapsto \eta f$, is an automorphism of $\mathcal{S}(R^n)$ and the adjoint operation on $\mathcal{S}'(R^n)$ defines an automorphism there on. Let Δ be the Laplacian defined in the usual way as a differential operator on tempered distributions. Then the operator

$$(I - \Delta)^s(f) = \mathcal{F}^{-1}(\eta \hat{f}),$$

can be extended to arbitrary real s, and this extension is always an automorphism of $\mathcal{S}'(R^n)$. For a general reference see Stein and Weiss [**16**].

Following Freidlander [**11**], we define the Sobolev space on R^n of arbitrary order s by

$$H^s = \{f \in \mathcal{S}' : (1+|2\pi y|^2)^{s/2}\hat{f}(y) \in L^2\}.$$

In this space for any $\lambda > 0$ the inner product

$$\langle f, g\rangle_{H^s_\lambda} = \int_{R^n} \overline{\hat{f}(y)}\hat{g}(y)(1+|2\pi\lambda y|^2)^s\, dy$$

induces a Hilbert space structure on H^s, that we denote by H^s_λ.

Properties of the H^s_λ norm, as a function of λ. If $s > 0$ then H^s_λ is a subspace of L^2, and so the distributions in H^s_λ are induced by ordinary functions. It is easy to see that all these norms are equivalent. Furthermore, as $\lambda \to +\infty$, the H^s_λ-norm increases to infinity with order λ^s, and as $\lambda \to 0$, the norm reduces to the L^2 norm. We shall mainly consider $\lambda = 1$, in which case we omit the subscript in H^s_λ.

Proposition 1.1. *The space H^s is invariant under translations and orthogonal transformations, and these act as unitary automorphisms of the spaces H^s_λ.*

Furthermore, H^s admits an orthogonal decomposition into spherical harmonics isomorphic to the classical decomposition of L^2, see, e.g., [**16**, *Lemma* 2.18].

H^s_λ is invariant under dilations, but the norm changes according to the rule

$$||f(cI)||^2_{H^s_\lambda} = ||f||^2_{H^s_{c\lambda}}, \quad \textit{for } c > 0.$$

Proposition 1.2. *Let f be in H^s with $s > 0$ and q a multi-index with $|q| \le s$. Then the* (*distributional*) *derivative $D^q(f)$ is in L^2, and*

$$||D^q(f)||_{L^2} \le C\lambda^{-|q|}||f||_{H^s_\lambda},$$

where C is a constant independent of λ and f.

Proof. Using properties of the Fourier transform we have

$$\begin{aligned} ||D^q(f)||_{L^2} &= ||\widehat{D^q(f)}||_{L^2} = ||(2\pi i y)^{|q|}\hat{f}(y)||_{L^2} \\ &\le \sup_{v\ge 0} \frac{v^{|q|}}{(1+v^2)^{s/2}}\lambda^{-|q|}||f||_{H^s_\lambda}. \end{aligned}$$

□

Proposition 1.3 (Sobolev embedding lemma). *Let f be in H^s with $s > m + n/2$ and m a nonnegative integer. Then f is in $C^m(R^n)$, and if q is a multi-index with $|q| \leq m$, then*

$$||D^q(f)||_{L^\infty} \leq C\lambda^{-(|q|+n/2)}||f||_{H^s_\lambda},$$

where C is a constant independent of λ and f. Furthermore, as $|x|$ approaches infinity, $D^q(f)(x)$ approaches zero.

Proof. Properties of the Fourier transform, and the Cauchy-Schwarz inequality give

$$\begin{aligned} ||D^q(f)||_{L^\infty} &\leq ||\widehat{D^q(f)}||_{L^1} = ||(2\pi iy)^{|q|}\hat{f}(y)||_{L^1} \\ &\leq \left\| \frac{(2\pi iy)^{|q|}}{\mu} \right\|_{L^2} ||\mu\hat{f}||_{L^2}, \end{aligned}$$

where $\mu(y) = (1 + |2\pi\lambda y|^2)^{s/2}$. And C is obtained as follows:

$$\begin{aligned} \left\| \frac{(2\pi iy)^q}{\mu} \right\|^2_{L^2} &\leq \left\| \frac{(2\pi|y|)^{|q|}}{(1+|2\pi\lambda y|^2)^s} \right\|_{L^1} \\ &\leq (\text{area of } n-1 \text{ sphere})\ \lambda^{-2|q|} \int_0^\infty \frac{(2\pi\lambda r)^{2m}}{(1+|2\pi\lambda r|^2)^s} r^{n-1}\,dr \\ &= (\text{area of } n-1 \text{ sphere})\ \lambda^{-2|q|-n} \int_0^\infty \frac{v^{2m}}{(1+v^2)^s} v^{n-1}\,dv, \end{aligned}$$

which is finite because $|q| < m$ and $2m + n - 2s < 0$. To conclude this proof we see that the last statement in this proposition is a consequence of the Riemann-Lebesgue lemma of Fourier analysis. □

2. The reproducing kernel for H^s. For the rest of this article we assume that $s > n/2$. From this assumption and Proposition 1.3 it follows that functions in H^s are continuous and vanish at infinity; that the mapping $f \mapsto f(x)$ is a continuous linear functional on H^s; and that H^s is an RKHS.

Theorem 2.1. *If $s > n/2$, then H^s has reproducing kernel $k_x(t) = k(t-x)$, where k is the inverse Fourier transform of $(1 + |2\pi y|^2)^{-s}$.*

Furthermore, k_x is a Green's function for the (possibly fractional) differential operator $(I-\Delta)^s$. And the reproducing kernel for H^s_λ is

$$k_{\lambda,x}(t) = \lambda^{-n} k((t-x)/\lambda).$$

Proof. Since $(1+|2\pi y|^2)^{-s}$ is in L^1, its inverse Fourier transform exists, and since $|\hat{k}|^2(1+|2\pi y|^2)^s$ is integrable, $k \in H^s$. Using that $\hat{k}_x(y) = \exp(-2\pi i x^T y)\hat{k}(y)$, and the definition of $\hat{k}$, we have

$$\begin{aligned}\langle k_x, f\rangle_{H^s} &= \int_{R^n} \overline{\hat{k}_x(y)}\hat{f}(y)(1+|2\pi y|^2)^s\, dy \\ &= \int_{R^n} e^{2\pi i x^T y}\hat{f}(y)\, dy = f(x).\end{aligned}$$

Here we have applied the Fourier inversion formula for the L^1 function $\hat{f}$.

The assertion about the role of k_x as a Green's function follows trivially from the definition of $(I-\Delta)^s$ given earlier and the fact that $y \mapsto \exp(-2\pi i x^T y)$ is the Fourier transform of $t \mapsto \delta(t-x)$.

Finally, the formula for the reproducing kernel for H^s_λ is a consequence of the above and the fact that

$$\widehat{k_{\lambda,x}}(y) = \exp(-2\pi i x^T y)(1+2|\pi\lambda y|^2)^{-s}. \tag{2.1}$$ □

For calculating k we have the next lemma which is a special case of statement (4) in [**12**, p. 88].

Lemma 2.1. *If $s > n/2$, then the inverse Fourier transform of the function $(1+|2\pi y|^2)^{-s}$ is*

$$k(x) = \frac{2^{1-s}(2\pi)^{-n/2}}{\Gamma(s)}|x|^{s-n/2}\mathcal{K}_{s-n/2}(|x|),$$

where $\mathcal{K}$ is the MacDonald's function, also known as a modified Bessel function of the third kind.

In the special case $s = m+(n+1)/2$, the MacDonald function reduces to a polynomial times an exponential function. Specifically, we only need to compute the MacDonald function of order $m+1/2$, $(= s-n/2)$, which is part of the more general formula (40) in [**3**, p. 10].

Theorem 2.2. *In the case where $s = m + (n+1)/2$, where m is a nonnegative integer, the reproducing kernel for H^s has the form*

$$k_x(t) = k(0)u_m(|t-x|),$$

where

$$k(0) = \frac{2^{1-2s}\pi^{(1-n)/2}(2m)!}{\Gamma(s)m!},$$

and where

$$u_m(z) = e^{-z}\sum_{l=0}^{m} C(m,l)z^l$$

with

$$C(m,l) = \frac{2^l}{l!}\binom{m}{l} \div \binom{2m}{l}.$$

We conclude this section by giving two recursion formulas for the calculation of the reproducing kernel when $s = m + (n+1)/2$. First, for a given m we have

$$C(m,l+1) = \frac{2(m-l)}{(l+1)(2m-l)}C(m,l).$$

And, second, if $u_m(z) = e^{-z}p_m(z)$ formula (25) in [**3**, II, 7.11] gives $p_0(z) = 1$, $p_1(z) = 1+z$, and

$$p_m(z) = p_{m-1}(z) + \frac{z^2}{(2m-1)(2m-3)}p_{m-2}(z), \quad \text{for } m > 1.$$

3. Approximation of functions. In this section we assume that the data points $x_1,\dots,x_q$ are distinct, and that $y_i = f(x_i)$, where f is a function in H^s with $s = m+(n+1)/2$, so that f has continuous derivatives up to order m. To simplify notation we shall denote the reproducing kernel by

$$k_i(t) := k_{\lambda,x_i}(t) = k(0)\lambda^{-n}u_m(|t-x_i|/\lambda).$$

Since $f(x_i) = \langle k_i, f\rangle_{H^s_\lambda}$, the question of finding $\tilde{f}$ becomes a least squares problem, i.e., $\tilde{f}$ is the H^s_λ-orthogonal projection of f onto the span of $\{k_i\}_{i=1}^q$. Therefore, the interpolant can be written as

$$\tilde{f}(t) = \sum_{i=1}^{q} w_i u_m(|t-x_i|/\lambda) = \lambda^n \sum_{i=1}^{q} \frac{w_i}{k(0)} k_i(t). \tag{3.1}$$

Let w and y denote the vectors of entries w_i and y_i, respectively. Then

$$Aw = y, \tag{3.2}$$

where A is the matrix of entries

$$a_{i,j} = u_m(|x_i - x_j|/\lambda) = \frac{\lambda^n}{k(0)}\langle k_i, k_j\rangle_{H^s_\lambda}.$$

Note that the matrix A is positive definite, symmetric, all its entries are positive, and the diagonal entries are ones. From the formula for u_m we see that the entries $a_{i,j}$ approach zero exponentially as the distance $|x_i - x_j|$ increases. We will investigate the condition number of the matrix A in a later section.

Convergence of approximations. We would hope that as the sampled points $\{x_i\}$ exhaust R^n, the interpolants $\tilde{f}$ would converge to f in the H^s norm. This is true, and we have

Theorem 3.1. *Let $E_1 \subset E_2 \subset \cdots$ be a sequence of finite sets whose union is dense in R^n. Let $f \in H^s_\lambda$ with $s > n/2$ and $\tilde{f}_E$ be the orthogonal projection of f onto span of $\{k_x : x \in E\}$. Then*

$$\lim_{i\to\infty} ||\tilde{f}_{E_i} - f||_{H^s_\lambda} = 0.$$

Proof. Let $D = \cup_{i=1}^{\infty} E_i$. Since H^s_λ is complete $\lim_{i\to\infty} \tilde{f}_{E_i}$ exists and equals the orthogonal projection of f onto the closure of the span

$\{k_x : x \in D\}$. If $x \in D$, then there exists an i such that $x \in E_j$ and $\langle k_x, f_{E_j}\rangle = \langle k_x, f\rangle$ for all $j \geq i$. Therefore,

$$\begin{aligned}\tilde{f}_D(x) &= \langle k_x, \tilde{f}_D\rangle_{H^s_\lambda} = \lim_{j\to\infty}\langle k_x, \tilde{f}_{E_j}\rangle \\ &= \langle k_x, f\rangle = f(x),\end{aligned}$$

which implies that the continuous function $\tilde{f}_D - f$ vanishes on the dense set D; hence, $f = \tilde{f}_D$. □

Rate of convergence for approximations. In this section we assume that data comes from a known function f in H^s. For positive numbers m, p, T we define the interpolation nodes

$$x_{l,j} = (l(2m-1) + j)h,$$

where $h = 1/(p(2m-1))$, $-pT \leq l \leq pT$ and $0 \leq j \leq 2m-1$. When holding m, T fixed as $p \to \infty$, we note that $h \to 0$ and that the set of interpolation nodes is eventually dense in $[-T, T]$.

Theorem 3.2. *Suppose that $f \in H^s(R)$, with $s = 1 + m$, and that f is the convolution $k * v$, where v in $L^2(R) \cap L^1(R)$. If $\tilde{f}$ is the H^s-projection of f onto* span $\{k_{x_{l,j}}\}$, *with $x_{l,j}$ as above, then*

$$\|f - \tilde{f}\|_{H^s} = O(h^m).$$

Furthermore, by Proposition 1.3 *we have* $\|(f - \tilde{f})^{(m)}\|_\infty = O(h^m)$.

Proof. It suffices to show that $\|f - g\|_{H^s} = O(h^m)$ for some function g in the span of $\{k_{x_{l,j}}\}$.

Since v is in L^1 there exists a positive T such that

$$\int_{R\setminus[-T,T]} |v(x)|\, dx \leq h^{2m}. \tag{3.3}$$

Now we define g and calculate its Fourier transform

$$\begin{aligned}g(y) &= \sum_{l,j} k_{x_{l,j}}(y) \int_{-T}^{T} L_{l,j}(z) v(z)\, dz \\ \hat{g}(y) &= \hat{k}(y) \int_{-T}^{T} \sum_{l,j} e^{-2\pi i y x_{l,j}} L_{l,j}(z) v(z)\, dz,\end{aligned}$$

where $L_{l,j}(z)$'s are the Lagrange polynomials $\prod_{i=1;i\neq j}^{2m-1}(z-x_{l,i})/(x_{l,j}-x_{l,i})$ on the interval $[x_{l,0},x_{l+1,0})$, and 0 elsewhere. From this definition it follows that each real number z belongs to at most one interval $[x_{l,0},x_{l+1,0})$, and that the denominator of $|L_{l,j}(z)|$ equals $j!(2m-1-j)!h^{2m-1}$ while the numerator is bounded by $(2m-1)!h^{2m-1}$; consequently,

$$\sum_{l,j}|L_{j,l}(z)| \le \sum_{j=0}^{2m-1}\binom{2m-1}{j} = 2^{2m-1}. \tag{3.4}$$

Since $||f-g||_{H^s}^2 = \langle f, f-g\rangle_{H^s} - \langle g, f-g\rangle_{H^s}$, all we need to show is that each of these terms is $O(h^{2m})$. To estimate the first term, we have
(3.5)

$$\begin{aligned}\langle f, f-g\rangle_{H^s} &= \int_R \overline{\hat{f}(y)}(\hat{k}\hat{v}-\hat{g})(y)/\hat{k}(y)\,dy\\ &= \int_R \overline{\hat{f}(y)}\int_{-T}^{T}\bigg(e^{-2\pi iyz} - \sum_{l,j} e^{-2\pi iyx_{l,j}}L_{l,j}(z)\bigg)v(z)\,dz\,dy\\ &\quad + \int_R \overline{\hat{f}(y)}\int_{R\setminus[-T,T]} e^{-2\pi iyz}v(z)\,dz\,dy.\end{aligned}$$

Since the sum in the above parenthesis is a polynomial interpolation for $\exp(-2\pi iyz)$, a standard error estimate, depending on the $2m$-th derivative with respect to z, gives [**6**, p. 132]

$$\bigg|e^{-2\pi iyz} - \sum_{l,j} e^{-2\pi iyx_{l,j}}L_{l,j}(z)\bigg| \le \frac{h^{2m}|2\pi y|^{2m}}{8m}, \qquad \forall z \in [-T,T].$$

From this inequality, (3.3) and (3.5), it follows that

$$|\langle f, f-g\rangle_{H^s}| \le h^{2m}\int_R |\hat{f}(y)|\bigg(|2\pi y|^{2m}\frac{1}{8m}\int_{-T}^{T}|v(z)|\,dz + 1\bigg)\,dy.$$

Since $v \in L^1$ and $f \in H^s(R)$ with $s = 1+m$, the above integrals are finite; therefore, $\langle f, f-g\rangle_{H^s} = O(h^{2m})$.

An argument analogous to the one above gives the estimate

$$|\langle g, f-g\rangle_{H^s}| \le h^{2m} \int_R |\hat{g}(y)| \bigg(|2\pi y|^{2m} \frac{1}{8m} \int_{-T}^{T} |v(z)|\, dz + 1 \bigg)\, dy.$$

Using the formula for $\hat{g}$ and (3.4), we find that

$$\int_R |\hat{g}(y)|\, |2\pi y|^{2m} \le 4^{2m-1} ||v||_{L^1} \int_R \hat{k}(y) |2\pi y|^{2m}\, dy.$$

Since this integral is finite because $\hat{k}(y) = (1+|2\pi y|^2)^{-s}$ with $s = m+1$, it follows that $\langle g, f-g\rangle_{H^s} = O(h^{2m})$. □

Remarks about Theorem 3.2. The nodes in the partition do not need to be equally spaced. But it is required that the quotient maximum over minimum distance between consecutive nodes be bounded as $\min|x_i - x_j|$ approaches zero, and that these nodes eventually become dense in $[-T, T]$. The assumptions on f are satisfied by a large family of functions. In fact convolutions of k with continuous functions of compact support satisfy these conditions and are dense in L^2.

A two-dimensional order of convergence can be proved as follows. Define the interpolation nodes $(x^1_{l,j}, x^2_{s,t})$ on a grid of squares whose side is $h = 1/(p(m-1))$. Define T as before, but for a double integral. Modify the sum on the definition of g to be $\sum L_{l,j}(z^1) L_{s,t}(z^2)$. After some computations we note that the bound (3.4) becomes 4^{m-1} and the bound for the polynomial interpolation in (3.5) becomes

$$2\frac{h^m}{4m}|2\pi y|^m + \frac{h^{2m}}{(4m)^2}|2\pi y|^{2m}.$$

Then proceed as in Theorem 3.2 to conclude that $|\langle f, f-g\rangle|$ and that $|\langle g, f-g\rangle|$ are $O(h^m)$; therefore, $||f - \tilde{f}|| = O(h^{m/2})$. By this same argument we can show that in R^n the order of convergence is $O(h^{m/n})$.

4. Applications of $\tilde{f}$. Since the linear operator $f \mapsto D^q(f)$, from H^s to $C_0(R^n)$ is continuous, it seems reasonable to approximate $D^q(f)$ by $D^q(\tilde{f})$. Here $C_0(R^n)$ denotes the space of continuous functions vanishing at infinity with the supremum norm topology.

Proposition 4.1. *Let $m \geq 2$ and let $z = |x-t|/\lambda$. Then the mapping $t \mapsto u_m(z)$ has gradient given by*

$$\nabla u_m(z) = \frac{1}{\lambda^2(2m-1)} u_{m-1}(z)(x-t) = \frac{e^{-z}}{\lambda(2m-1)} p_{m-1}(z) z v,$$

where $u_m(z) = e^{-z} p_m(z)$ and v is the unit vector in the direction from t to x. The second partial derivatives are given by

$$\frac{\partial^2}{\partial t^j t^{j'}} u_m(z) = \frac{e^{-z}}{\lambda^2(2m-1)} \left(\frac{p_{m-2}(z)}{\lambda^2(2m-3)} z^2 v^{j'} v^j - p_{m-1}(z) \delta_{j,j'} \right).$$

From this proposition and the linearity of differentiation, with x_i and w_i as in (3.1) it follows that

$$\Delta \tilde{f}(t) = \sum_{i=1}^{q} \frac{w_i e^{-z_i}}{\lambda^2(2m-1)} \left(\frac{p_{m-2}(z_i)}{\lambda^2(2m-3)} z_i^2 - n p_{m-1}(z_i) \right).$$

Here we note that the approximation $\Delta \tilde{f}$ can be computed in a parallel environment, which makes $\Delta \tilde{f}$ valuable for applications.

Since the Fourier transform maps H^s unitarily onto $L^2(R^n, \mu^2, dy)$, it seems reasonable to approximate the Fourier transform of f by the Fourier transform of $\tilde{f}$. Here $\mu(y) = (1 + |2\pi y|^2)^{s/2}$.

Proposition 4.2. *Let $s = m + (n+1)/2$ as before. Then*

$$\hat{\tilde{f}}(y) = (1 + |2\pi\lambda y|^2)^{-s} \frac{\lambda^n}{k(0)} \sum_{j=1}^{q} w_j \exp(-2\pi i x_j^T y).$$

Thus $\hat{\tilde{f}}$ is a linear combination of sinusoids modulated by $\mu^{-2}(\cdot)$. When the sampling points are not equally spaced, computations of $\hat{\tilde{f}}$ cannot be simplified as in the fast Fourier transform algorithm.

Another application that follows from Proposition 4.2 is to approximate $\hat{f}(0) = \int_{R^n} f(x)\, dx$ by

$$\hat{\tilde{f}}(0) = \int_{R^n} \tilde{f}(x)\, dx = \frac{\lambda^n}{k(0)} \sum_{i=1}^{q} w_i.$$

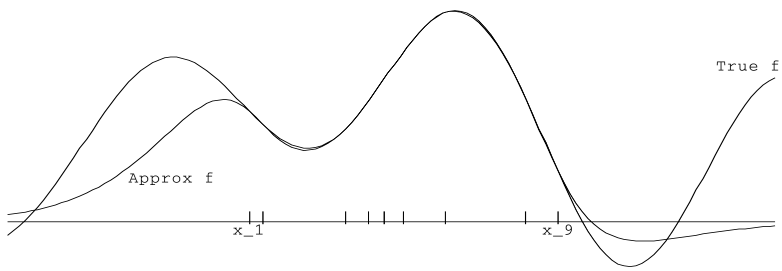

FIGURE 4.1.

As a last application, the least squares formulation (3.1) allows $||f||^2_{H^s_\lambda}$ being approximated by

$$||\tilde{f}||^2_{H^s_\lambda} = \frac{\lambda^n}{k(0)} \sum_{i=1}^{q} w_i f(x_i).$$

Choices for the parameter λ. In this section we study the role of the parameter λ in providing a tradeoff between "visual smoothness of interpolants" and easing of computations.

On the one hand, as $\lambda \to 0$, for any m and $t \neq t'$, $u_m(|t - t'|/\lambda)$ approaches 0. This improves the condition number of A, but the interpolant will look like a string with lots of slack, leading to sharp peaks near data values. On the other hand, as $\lambda \to \infty$ each $u_m(|t-t'|/\lambda)$ approaches 1 and the matrix A becomes arbitrarily ill conditioned.

A strategy that leads to a reasonable interpolant is to select λ large, but within a range where solutions to (3.2) have a certain number of accurate digits. This is because inaccurate solutions to (3.2) would lead to inaccurate interpolants. However, numerical solutions of any accuracy can lead to $\tilde{f}(x_i) \neq f(x_i)$ at some data point.

A way to measure accuracy in solving (3.2) is the condition number of matrix A. For instance, if $\text{cond}(A)$ is 10^t, then the solution w to (3.2) computed in v decimal digit arithmetic may have no more than $v - t$ accurate significant figures, see [**8**, p. I.8].

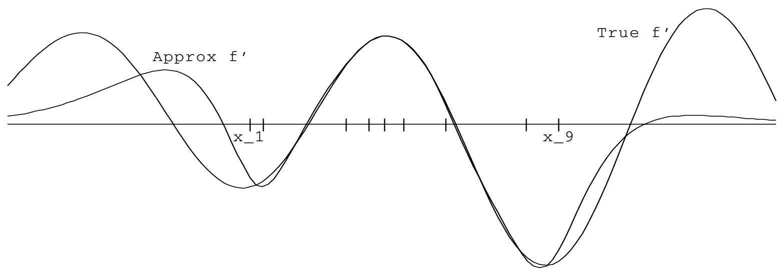

FIGURE 4.2.

Estimates for cond(A) can be done by the Gerschgorin theorem or numerically when solving (3.2). Numerical experiments show that cond(A) depends mostly on the minimum and in many cases is independent of the maximum distance between consecutive points. Therefore, an a priori choice for λ is $\min|x_i - x_j|$ times a constant that depends on the dimension of the domain. To find this constant we did the experiments summarized below. First, on equally spaced nodes in $[-1, 1]$, and then with some points clustered around zero and some around 200. In both cases we kept the minimum distance between points equal to $1/128$, and estimated cond(A). For this estimate we used the subroutine DPPCO from Linpack [**8**].

When $k(z) = \exp(-z)$, if $\lambda < .003$, then cond$(A) = 1$. If $\lambda = 1$, then cond$(A) = 120$ with rate of increase λ^2. If $\lambda = 10$, then cond$(A) = 2.4 \times 10^4$ with rate of increase λ.

When $k(z) = e^{-z}(1 + z)$ if $\lambda < .003$, then cond$(A) = 1$. If $\lambda = .134$, then cond$(A) = 8260$ with rate of increase λ^4. If $\lambda = 3.05$, then cond$(A) = 3.6 \times 10^9$ with rate of increase λ^3.

When $k(z) = e^{-z}(1 + z + z^2/3)$, if $\lambda < .003$, then cond$(A) = 1$. If $\lambda = .107$, then cond$(A) = 2.78 \times 10^5$ with rate of increase λ^6. If $\lambda = 1.95$, then cond$(A) = 1.85 \times 10^{12}$ with rate of increase λ^5.

When $k(z) = e^{-z}(1 + z + 2z^2)/5 + z^3/15)$, if $\lambda < .0024$, then cond$(A) = 1$. If $\lambda = .134$, then cond$(A) = 1.22 \times 10^8$ with rate of increase $\lambda^{7.8}$. If $\lambda = .8$ then cond$(A) = 7.14 \times 10^{13}$ with rate of increase λ^7.

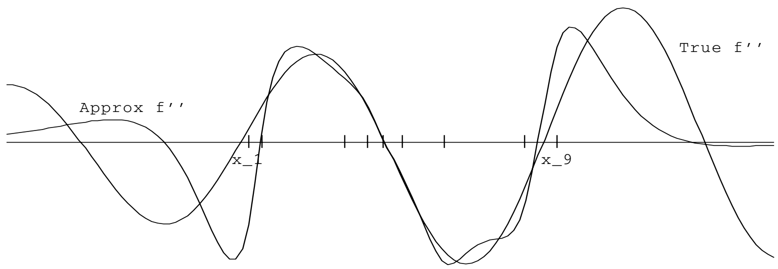

FIGURE 4.3.

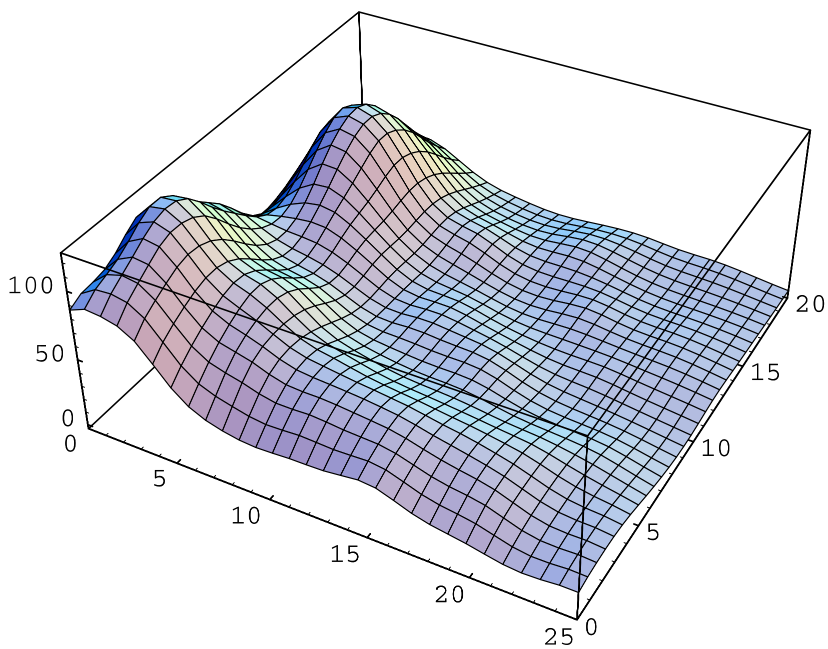

FIGURE 4.4.

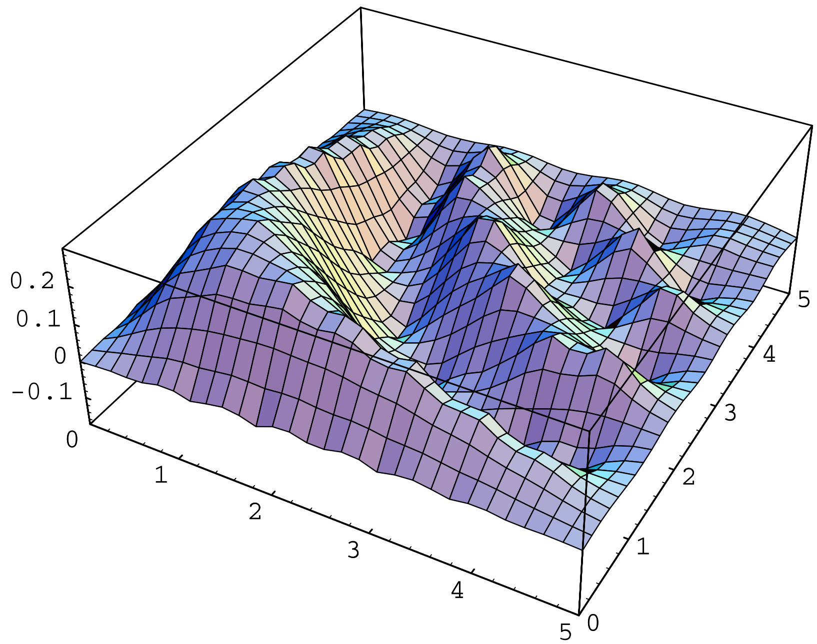

FIGURE 4.5.

An a posteriori parameter choice that comes to our minds is based on the cross validation principle, also known as "withheld data." Omit some data points, construct $\tilde{f}$ for several values of λ and select the λ for which the interpolant best predicts the data omitted. Cross validation has been implemented for choosing a parameter in thin plate splines interpolation, and it is available under the title GCVPACK and RKPACK from netlib; for details see [**4**].

Visual test for interpolants. To this end we use data values from a known function to construct $\tilde{f}$. Then graph the approximate and the true functions, or their cross sections, on the same coordinate system. Among the functions approximated we have

$$f(t) = \sum_j f\left(\frac{j}{2w}\right)\operatorname{sinc}\left(2\pi w\left(t - \frac{j}{2w}\right)\right) \tag{4.1}$$

where $\operatorname{sinc}(x) = \sin(x)/x$, j is in a subset of integers, and $w > 0$. It is known that these functions are dense in H^s and that their Fourier transform has compact support. By setting $f(j/(2w))$ to be a standard normal distribution we are interpolating a "function chosen at random," which is a plausible and practical interpretation of white noise functions [**15**, p. 86].

We conclude this article by showing some graphs of interpolants. Figures 4.1–4.3 correspond to an approximation of (4.1) with $w = 3$ and $j = -3, \dots, 3$. There are nine irregularly spaced sampling points. The kernel is $e^{-z}(1 + z + z^2/3)$ and $\lambda = .05$. The approximations $\tilde{f}$ and $\tilde{f}'$ appear to fit f and f' quite well considering the small number of data points, but f'' was poorly approximated.

Figure 4.4 shows an approximation of $\sin(xy)/(1 + |x| + |y|)$. There are 81 sampling points on the domain $[0, 5] \times [0, 5]$. The kernel is e^{-z} and $\lambda = 1$ that keeps $\operatorname{cond}(A)$ around 40.

Figure 4.5 shows an interpolant for Akima's data, experimental data provided by Richard Franke. There are 50 data points on the domain $[0, 25] \times [0, 20]$. The kernel is $e^{-z}(1 + z + z^2/3)$ and $\lambda = 2.5$ that keeps $\operatorname{Cond}(A)$ around 10.

Acknowledgment. We would like to thank the anonymous reviewer for the helpful suggestions and much useful information.

REFERENCES

1. P. Alfeld, *Scattered data interpolation in three or more variables*, in *Mathematical method in computer aided geometric design* (Lyche & Schumaker, eds.), Academic Press, Boston, 1989.

2. N. Aronszajn, *Theory of reproducing kernels*, Trans. Amer. Math. Soc. **68** (1950), 337–404.

3. H. Bateman, *Higher transcendental functions*, Volume II, McGraw-Hill, New York, 1953.

4. D.M. Bates, F. Reames and G. Wahba, *Getting better contour plots with S and Gcvpack*, Technical Report 865, Dept. of Stat., University of Wisconsin.

5. M.D. Buhman and M.J.D. Powell, *Radial basis function interpolation on an infinite regular grid*, in *Algorithms for approximation* II (Mason & Cox, eds.), Chapman & Hall, London, 1990.

6. W. Cheney and D. Kincaid, *Numerical methods and computing*, Second Edition, Brooks & Cole, 1985.

7. M.G. Cui and Z.X. Deng, *Two dimensional reproducing kernel and surface interpolation*, J. Comput. Math. **4** (1986), 177–181.

8. J.J. Dongarra, C.B. Moler, J.R. Bunch and G.W. Stewart, *Linpack users' guide*, SIAM, Philadelphia, 1979.

9. J. Duchon, *Splines minimizing rotation-invariant seminorms in Sobolev spaces*, in *Multivariate approximation theory* (Schemp and Zeller, eds.), Birkhauser, 1975.

10. R. Franke, *Scattered data interpolation*: *Test of some methods*, Math. Comp. **38** (1982), 181–200.

11. F.G. Friedlander, *Introduction to the theory of distributions*, Cambridge U. Press, New York, 1982.

12. I.M. Gelfand and G.E. Silov, *Generalized functions and calculations with them*, Volume I, 1958.

13. E. Grosse, *A catalogue of algorithms for approximation*, in *Algorithms for approximation* II, (Mason & Cox, eds.), Chapman & Hall, London, 1990.

14. W.R. Madcych and S.A. Nelson, *Multivariable interpolation and conditionally positive definite functions* II, Math. Comp. **54** (1990), 211–230.

15. Shanon and Weaver, *The Mathematical theory of communication*, The University of Illinois Press, 1964.

16. E.M. Stein and G. Weiss, *Introduction to Fourier analysis on Euclidean spaces*, Princeton University Press, Princeton, New Jersey, 1971.

17. G. Wahba, *Convergence rates of certain approximate solutions to Fredholm integral equations of the first kind*, J. Approx. Theory **7** (1973), 167–185.

18. ———, *Interpolating surfaces*: *High order convergence rates and their associated designs, with application to X-ray image restoration*, Technical Report No. 523, Dept. of Stat. University of Wisconsin (1978).

19. E.J. Wegman, *Reproducing kernel Hilbert spaces*, Encyclopedia of Statistical Sciences, John Wiley & Sons, New York, 1988.

Department of Mathematics and Computer Sciences, Southwest Texas State University, 601 University Drive, San Marcos, TX 78666

E-mail: jd01@academia.swt.edu; and rd01@academia.swt.edu