JOURNAL OF APPROXIMATION THEORY 64, 69–94 (1991)

Norms of Inverses and Condition Numbers for Matrices Associated with Scattered Data

FRANCIS J. NARCOWICH AND JOSEPH D. WARD*

Department of Mathematics, Texas A & M University, College Station, Texas 77843-3368

Communicated by Charles K. Chui

Received October 9, 1989; revised February 28, 1990

For interpolation matrices arising in connection with translates of a conditionally negative definite, radially symmetric $\mathbb{R}^s$-function of order 1, we give a general method for obtaining bounds both on the norm of the inverse of the interpolation matrix and on the condition number of that matrix. We apply our method to obtain these bounds in several cases, including those associated with the functions $\sqrt{1+\|x\|_2^2}$ and $\log(1+\|x\|_2^2)$. © 1991 Academic Press, Inc.

I. INTRODUCTION

Data fitting in two or more dimensions is a practical problem that has many important applications—solution of computer aided design problems, for example. Recently, progress has been made in solving multi-dimensional data fitting problems. Of the latest approaches to multi-dimensional data fitting, the two most important are the method of thin plate splines, as developed by Duchon [3, 4], and Hardy's method of multi-quadric surfaces [8]. Hardy's approach has undergone rapid development and now provides an elegant, convenient tool for interpolating scattered, multivariate data. Instrumental in this development were results of Madych and Nelson [10, 11] and of Micchelli [12]. Among other things, they answered a question of Franke [7] by showing that N arbitrary, distinct points $\{x_j\}_{j=1}^N$ in $\mathbb{R}^2$ endowed with the Hilbert-space norm could always be interpolated by linear combinations of the N functions $\{\sqrt{1+\|x-x_j\|_2^2}\}_{j=1}^N$. Their research has stimulated much work on the Hardy approach to multi-dimensional data fitting. For a review of these and other developments, we refer the reader to Dyn's survey article [5].

Franke's interpolation problem is a special case of the more general

* Supported by the National Science Foundation under Grant DMS-8901345.

0021-9045/91 $3.00

problem of interpolating data $\{x_j\}_{j=1}^N$ in $\mathbb{R}^s$ by translates of a given function $h(x)$. (See Section II.) Solving the more general problem entails showing that the $N \times N$ interpolation matrix A with j, k-entry $A_{jk} = h(x_j - x_k)$ is invertible. In [10–12], the researchers mentioned above showed that a *sufficient* condition for A to be invertible was that the function h be *strictly* conditionally negative definite of order 1. (See Section II, Definition 2.1. The "order" just mentioned arises in connection with conditions one puts on h.) They then described a wide class of *radial* functions that fit into this category. However, they did not give quantitative estimates for either $\|A^{-1}\|$ or for the condition number of A.

Very recently, Ball [1] gave just such estimates in the case of the function $h(x) = \|x\|_2$, which is a strictly conditionally negative definite function of order 1. We will discuss a precise statement of his results in Sections V and VII. The difficult estimates arise in connection with bounding the norm of the inverse of the interpolation matrix. Once these estimates are made, a few simple facts from matrix analysis can be used to bound the condition number of an interpolation matrix. We wish to add that the estimate he obtained (see (5.5)) depends *only* on the minimal separation distance for the data, and not on the number of data points or on any other details of the distribution.

Every F that is a conditionally negative definite *radial* function of order 1 on $\mathbb{R}^s$ is generated by a "Bessel" transform of some nonnegative measure. In Section V, we will show that if the measure generating F decays polynomially, then a simple adaptation of the method employed by Ball can be used to obtain estimates on $\|A^{-1}\|$ whenever the function h has the form $h(x) = F(\|x\|_2)$. In that section we also show that it is not possible to further adapt this method to cover cases for those h coming from F that are generated by measures that decay exponentially. Since both $\sqrt{1 + \|x\|^2}$ and $\log(1 + \|x\|^2)$ arise from F's that are generated by measures with exponential decay (see Section III), further progress can only come from a different method.

It is our purpose in writing this paper to give a general method for obtaining quantitative estimates both for $\|A^{-1}\|$ and for the condition number of A when the corresponding h is a conditionally negative definite function of the form $h(x) = F(\|x\|_2)$, provided that the integral representation for F (see Section III) is such that the measure involved can be estimated from below. Our method was inspired by a theorem in Zygmund [17]. As applications of this method, we obtain quantitative estimates both in the case where h is the function $\sqrt{1 + \|x\|_2^2}$ studied by Franke [7] and in the case where it is the function $\log(1 + \|x\|_2^2)$ analyzed by Dyn [5]. As in the case of $h(x) = \|x\|_2$, our estimates for the norm of the inverse of the interpolation matrix depend only on the minimal separation distance of the distribution of data, and not on any other details of that distribution.

Summary and outline of the paper. In Section II, we discuss the interpolation problem that we want to solve. In doing so, we introduce relevant terms and notation, and state precisely a lemma that is due to Ball and that is crucial to our method.

In Section III, we review the well-known integral representation for order 1 conditionally negative definite radial $\mathbb{R}^s$-functions. Applying this representation, we next derive bounds on such functions; these bounds we use in Section VII in connection with estimating condition numbers of interpolation matrixes. We also use them, along with the theory of tempered distributions, to give a formula that simplifies computing the measure appearing in the integral representation when the underlying space is $\mathbb{R}^3$. Using this formula, we compute the measures generating both $\sqrt{1+r^2}$ and $\log(1+r^2)$.

In Section IV, we develop our method for estimating the minimum of the quadratic form introduced in Section II and represented via the formulas of Section III.

In Section V, we begin by showing that in cases where the radial function F is generated by a measure that is *polynomially* bounded below, the corresponding interpolation matrix is invertible; we also give a bound for the norm of its inverse. We point out the simple method that worked for the polynomial case will fail to provide bounds in cases where measures have faster decay. For purposes of illustration, we then use our method to obtain estimates on $\|A^{-1}\|$ when $h(x)=\|x\|_2$, the case studied by Ball [1].

In Section VI, we demonstrate the effectiveness of our method *even* when the associated generating measures have exponential decay by doing two examples. Namely, we use our method to obtain estimates on $\|A^{-1}\|$ when h is either $\sqrt{1+\|x\|_2^2}$ or $\log(1+\|x\|_2^2)$.

In Section VII, we conclude by using the results from the two previous sections to obtain upper bounds on the condition numbers for various interpolation matrices.

II. An Interpolation Problem

To provide motivation for our discussion of conditionally negative definite radial functions, to establish notation and terminology, and to give a precise statement of the problem that we want to solve, we wish to review the scattered data interpolation problem; this has been discussed in detail in several papers [6, 10–12].

Given a continuous function $h:\mathbb{R}^s\to\mathbb{C}$, vectors $\{x_j\}_1^N$ in $\mathbb{R}^s$, and scalars $\{y_j\}_1^N$, under what conditions on h can we always find a function f such that the system of equations,

$$f(x_j)=y_j, \qquad j=1, ..., N,$$

has a solution of the form

$$f(x)=\sum_{j=1}^{N} c_j h(x-x_j)+p_{m-1}(x),$$

where p_{m-1} is in π_{m-1}, the set of all polynomials in x with total degree $m-1$ or less, and where the c_j's are subject to the condition

$$\sum_{j=1}^{N} c_j q(x_j)=0, \qquad \forall q \in \pi_{m-1}. \tag{2.1}$$

A sufficient condition for this problem to be solved (see [10–12]) is the following: For every possible finite set $\{x_j\}_1^N$ in $\mathbb{R}^s$ and every set of complex number $\{c_j\}_1^N$ that satisfy (2.1), the function h satisfies

$$\sum_{j,k=1}^{N} \bar{c}_j c_k h(x_k-x_j)<0. \tag{2.2}$$

Such a function belongs to the well-known class defined below.

DEFINITION 2.1. Let h: $\mathbb{R}^s \to \mathbb{C}$ be continuous. We say that h is conditionally negative definite of order m if for every finite set $\{x_j\}_1^N$ of distinct points in $\mathbb{R}^s$ and for every set of complex numbers $\{c_j\}_1^N$ satisfying (2.1), we have

$$\sum_{j,k=1}^{N} \bar{c}_j c_k h(x_k-x_j) \leqslant 0. \tag{2.3}$$

We will denote this set by $\mathcal{N}_m^s$. In addition, if the inequality is strict—i.e., h satisfies (2.2)—then we will say that h is strictly conditionally negative definite of order m.

We remark that functions of order 0 are negatives of the usual functions of positive type, as defined by Bochner. We also point out that only the topological property of continuity $\mathbb{R}^s$ is used, and no particular norm on $\mathbb{R}^s$ is singled out as special; thus, one is free to work in a norm convenient for the case at hand. Finally, it should be noted that for $m=0$ and $m=1$, it is very easy to extend these definitions to cases in which $\mathbb{R}^s$ is replaced by a topological group.

There is another definition that we wish to make, one that clears up a semantical difficulty occurring in the current literature. Let us now suppose that $\mathbb{R}^s$ has a norm $\|\cdot\|$. We define the function v: $\mathbb{R}^s \to \mathbb{R}^+$ by $v(x)=\|x\|$.

DEFINITION 2.2. We will say that a continuous function F: $\mathbb{R}^+ \to \mathbb{R}$ is a conditionally negative definite radial function of order m if $F \circ v$ is in $\mathcal{N}_m^s$. We will denote the set of all such functions by $\mathcal{RN}_m^s(\|\cdot\|)$.

In the language just introduced above, a sufficient condition for the scattered data interpolation problem of order m to have a solution is that h be *strictly* conditionally negative definite of order m. In the $m=1$ case, one can show that if $F \in \mathscr{RN}_1^s(\|\cdot\|_2)$ is nonnegative and if $h(x)=F(\|x\|_2)$ is *strictly* conditionally negative definite, then the $N \times N$ matrix A, with $A_{j,k}=(x_j-x_k)$, is invertible. (See Lemma 2.3 below.) This in turn implies that the interpolating function $f(x)$ has the form

$$f(x)=\sum_{j=1}^{N} c_j h(x-x_j)+\alpha. \tag{2.4}$$

The constant α and the c_j's, which satisfy $\sum_{j=1}^{N} c_j=0$, are obtained as follows. Let $U=(1 \cdots 1)^T$, $Y=(y_1 \cdots y_N)^T$, and $C=(c_1 \cdots c_N)^T$. We then have

$$\alpha=\frac{\langle A^{-1}U, Y\rangle}{\langle A^{-1}U, U\rangle} \quad \text{and} \quad C=A^{-1}(Y-\alpha U). \tag{2.5}$$

We remark that $\langle A^{-1}U, U\rangle \neq 0$. If we did have $\langle A^{-1}U, U\rangle=0$, then with $Z=A^{-1}U$ the fact that h is *strictly* negative definite would imply that

$$\langle A^{-1}U, U\rangle=\langle AZ, Z\rangle=\sum_{j,k=1}^{N} \bar{z}_j z_k h(x_k-x_j)<0,$$

which is a contradiction. Also, note that if A^{-1} exists and $F \in \mathscr{RN}_1^s(\|\cdot\|_2)$, then the interpolant f in (2.4) can be taken to be a "pure" radial interpolant—i.e., $f(x)=\sum_{j=1}^{N} c_j F(\|x-x_j\|_2)$.

What has been said above illustrates the role that the matrix A^{-1} plays, and indicates the importance of estimating its norm. Indeed, estimating the norm of A^{-1}, when A is generated by $h(x)=F(\|x\|_2)$, with $F \in \mathscr{RN}_1^s(\|\cdot\|_2)$, is precisely the problem we are addressing in this paper.

We need to say a few words about notation. Our chief concern here is with order $m=1$ conditionally negative definite radial functions on $\mathbb{R}^s$ with the Hilbert-space norm, $\|\cdot\|_2$. To avoid carrying along notational baggage, we will set

$$\mathscr{RN}^s=\mathscr{RN}_1^s(\|\cdot\|_2).$$

Also, we will always taken $|x|=\|x\|_2$ when x is a vector, and for a matrix B we will take $\|B\|$ to be the matrix norm corresponding to $\|\cdot\|_2$. Having introduced the notation that we need, we can now state the lemma alluded to earlier.

LEMMA 2.3 (K. Ball). *Let $\{x_j\}_1^N$ be distinct points in $\mathbb{R}^s$ and let $F \in \mathscr{RN}^s$ be nonnegative and suppose that $h(x)=F(|x|)$ is a strictly conditionally* *negative definite function of order* 1. *Also, let* A *be the matrix with entries* $A_{j,k} = h(x_j - x_k)$. *If the inequality*

$$\sum_{j,k=1}^{N} A_{jk} \xi_j \bar{\xi}_k \leqslant -\theta \sum_{j=1}^{N} |\xi_j|^2 \tag{2.6}$$

is satisfied whenever the complex numbers ξ_j *satisfy* $\sum_{j=1}^{N} \xi_j = 0$, *then*

$$\|A^{-1}\| \leqslant \theta^{-1}. \tag{2.7}$$

We omit the details of the proof; see [1]. We do wish to point out that the proof involves only elementary matrix theory, and that a similar lemma was proved somewhat earlier by Schoenberg [13].

This result turns the task of getting estimates for $\|A^{-1}\|$ into one of estimating θ. The remainder of the paper is devoted to carrying out that task.

III. Representations of Conditionally Negative Definite Radial Functions

In this section, we begin by recalling that functions in $\mathscr{RN}^s$ have an integral representation in terms of a measure, a representation that plays an important role when used in connection with Lemma 2.3. We then use this representation to get bounds on radial functions. Employing these bounds and a distribution theoretic argument, we give a simple method for calculating the measure that appears in the representation—at least in the important case when $s = 3$. The bounds themselves will prove useful in estimating condition numbers; see Section VII.

It is known [13, 16] that $F(r)$ is a conditionally negative definite radial function on $\mathbb{R}^s$, that is, $F \in \mathscr{RN}^s$, if and only if there is a positive measure $d\alpha$ on $\mathbb{R}^+$ such that the function F (cf. [16, p. 38]; their F^2 corresponds to our F) has the integral representation

$$F(r) = F(0) + \int_0^\infty \frac{1 - \Omega_s(ur)}{u^2}\, d\alpha(u), \tag{3.1}$$

where we have that $d\alpha$ satisfies the condition $\int_1^\infty u^{-2}\, d\alpha(u) < \infty$, and the function $\Omega_s(\cdot)$ is [16, p. 27]

$$\Omega_s(x) = \begin{cases} \cos x & \text{for} \quad s = 1, \\ \dfrac{\int_0^\pi e^{ix\cos\phi} \sin^{s-2}\phi\, d\phi}{\int_0^\pi \sin^{s-2}\phi\, d\phi} & \text{for} \quad s = 2, 3, \ldots. \end{cases} \tag{3.2}$$

We remark that there are several useful representations for Ω_s; see [16, pp. 26, 27]. In particular, we note that when $s=3$ we have

$$\Omega_3(x)=\frac{\sin x}{x}. \tag{3.3}$$

We will now use (3.1) to obtain bounds on $F(r)$ and, when $s=3$, develop a simple method for calculating the measure $d\alpha$ appearing in (3.1); we will then apply this method to calculate $d\alpha$ in several cases of interest.

We begin by getting the bounds we need. Using the integral on the right in (3.1), we may extend $F(r)$ to be an even function on $\mathbb{R}$. Decompose the right side of (3.1) into the sum

$$F(r)=F(0)+\int_0^1 \frac{1-\Omega_s(ur)}{u^2}\,d\alpha(u)+\int_1^\infty \frac{1-\Omega_s(ur)}{u^2}\,d\alpha(u). \tag{3.4}$$

To estimate the integral with $u\geqslant 1$, observe that from (3.2) we have $|\Omega_s(x)|\leqslant 1$, and so

$$\left|\int_1^\infty \frac{1-\Omega_s(ur)}{u^2}\,d\alpha(u)\right|\leqslant 2\int_1^\infty u^{-2}\,d\alpha(u).$$

For the integral with $0\leqslant u\leqslant 1$, we first use (3.2) and Taylor's Theorem to get $|1-\Omega_s(x)|\leqslant |x|^2/2$, and then we immediately arrive at

$$\left|\int_0^1 \frac{1-\Omega_s(ur)}{u^2}\,d\alpha(u)\right|\leqslant (r^2/2)\int_0^1 d\alpha(u).$$

Combining (3.4) with the last two inequalities then yields

$$|F(r)|\leqslant c_1 r^2+c_2. \tag{3.5}$$

Thus F is bounded by a quadratic polynomial. Since F is continuous, we also have that F may be regarded as being a tempered distribution; i.e., $F\in\mathscr{S}'(\mathbb{R})$.

Let us now turn to finding a method of calculating $d\alpha$ in the case in which $s=3$. Of course, we have that F is a tempered distribution and therefore so is $rF(r)$. If we take $G\in\mathscr{S}$ (we assume G is real-valued as well), then $D_r^3 G$ is too. Hence, it makes sense to form $\langle rF(r), D_r^3 G(r)\rangle$. We may compute this quantity using integration by parts in the distributional sense. The result is that

$$\langle r(F(r)-F(0)), D_r^3 G(r)\rangle = -\langle D_r^3(rF(r)), G(r)\rangle. \tag{3.6}$$

On the other hand, using (3.1) with $s=3$, (3.3), and Fubini's Theorem, we also have that

$$\langle r(F(r)-F(0)), D_r^3 G(r)\rangle = \int_0^\infty \int_{-\infty}^\infty \frac{r-\sin ur/u}{u^2} D_r^3 G(r)\, dr\, d\alpha(u).$$

Integrating by parts in the integral over r, we find that

$$\langle r(F(r)-F(0)), D_r^3 G(r)\rangle = -\int_0^\infty \int_{-\infty}^\infty \cos urG(r)\, dr\, d\alpha(u).$$

Observe that if G is an *odd* function, then $D_r^3 G(r)$ will be *even* and, because $r(F(r)-F(0))$ is *odd*, we would have $\langle r(F(r)-F(0)), D_r^3 G(r)\rangle = 0$. Thus, in addition to choosing G to be real, we lose nothing if we also require it to be *even*. With this added assumption on G, the last equation can be rewritten with the inner integral replaced by $\hat{G}(u)$, the Fourier transform of $G(r)$; the result is

$$\langle r(F(r)-F(0)), D_r^3 G(r)\rangle = -\int_0^\infty \hat{G}(u)\, d\alpha(u).$$

Because G is even, $\hat{G}$ is too. If we extend $d\alpha$ to be an even measure on $\mathbb{R}$, then by regarding $\alpha'(u) =: d\alpha(u)/du$ as a tempered distribution (possible, since $(u^2+1)^{-1}\, d\alpha(u)$ is a finite measure), we can transform our last equation into

$$\langle r(F(r)-F(0)), D_r^3 G(r)\rangle = -(1/2)\langle \alpha'(u), \hat{G}(u)\rangle.$$

Comparing this with (3.6) gives us

$$\langle D_r^3(rF(r)), G(r)\rangle = (1/2)\langle \alpha'(u), \hat{G}(u)\rangle.$$

By using the definition of the Fourier transform of a tempered distribution [14, Chap. 25], we finally arrive at the equation

$$(1/2\pi)\langle [D_r^3(rF(r))]\hat{\ }(u), \hat{G}(u)\rangle = (1/2)\langle \alpha'(u), \hat{G}(u)\rangle. \tag{3.7}$$

Because of the parity and reality of the tempered distributions involved, (3.7) actually holds for all $G \in \mathscr{S}$. We thus obtain the following result.

THEOREM 3.1. *Let $F(r)$ be in $\mathscr{R}\mathscr{N}^3$. If $F(r)$ also denotes the even extension of F, then F is a tempered distribution and, in a distributional sense*

$$\alpha'(u) = (1/\pi)[D_r^3(rF(r))]\hat{\ }(u). \tag{3.8}$$

Alternatively, $\alpha'(u)$ *may be expressed in the form*

$$\alpha'(u) = (2/\pi)[D_r^3(rF(r))]_c(u), \tag{3.9}$$

where $[\cdot]_c$ *denotes the Fourier cosine transform.*

Proof. We established (3.8) within the preceding discussion; (3.9) follows immediately from (3.8) and the evenness of $D_r^3(rF(r))$. ∎

We remark that higher dimensional analogues of this theorem exist, but they are not as readily applicable to cases of interest.

Let us now compute the $d\alpha$'s for the functions mentioned in Section I; these are

$$F_j(r) = \begin{cases} r, & \text{if } j=1, \\ \sqrt{1+r^2}, & \text{if } j=2, \\ \log(1+r^2), & \text{if } j=3. \end{cases} \tag{3.10}$$

For F_1, the even extension is $F_1(r) = |r|$. A standard distributional calculation (see [9, p. 25]) gives us that

$$D_r^3(r\,|r|) = 4\delta(r), \qquad \text{where} \quad \delta = \text{the Dirac } \delta\text{-function.}$$

From this and (3.8), we see that $\alpha_1'(u) = 4/\pi$, and so

$$d\alpha_1 = (4/\pi)\,du. \tag{3.11}$$

The functions F_2 and F_3 are both even as analytical expressions. Using *Mathematica* to do symbolic differentiation, we found that

$$D_r^3(rF_2(r)) = 3(1+r^2)^{-5/2} \qquad \text{and}$$

$$D_r^3(rF_3(r)) = 16(1+r^2)^{-3} - 8(1+r^2)^{-2} - 2(1+r^2)^{-1}.$$

These are both smooth, integrable functions; no distribution theoretic calculations are necessary to deal with them. Indeed, using (3.9) and the table of cosine transforms in [2, p. 11, No. 7], we find that

$$d\alpha_2 = (2u^2/\pi)K_2(u)\,du \qquad \text{and} \qquad d\alpha_3 = 2u(1+u)e^{-u}\,du. \tag{3.12}$$

Here, the function $K_2(u)$ is a modified Bessel function of the second kind.

We close this section by pointing out that it should be possible to use Theorem 3.1 in conjunction with a table of cosine transforms to produce a wide variety of conditionally negative definite functions, all potentially useful in applications.

IV. The Method

We are now ready to describe our method for estimating the l_2-norm of the inverse interpolation matrix, A^{-1}, described in Section II. Examples illustrating how our method works will be given later; see Sections V and VI.

Recall that Lemma 2.3 reduces the problem of estimating $\|A^{-1}\|$ to one of estimating $-\theta$, the maximum of the quadratic form (2.6), the estimate being that given in (2.7): $\|A^{-1}\| \leqslant \theta^{-1}$. As we shall see, the estimate we arrive at turns out to be independent of the number of data points; indeed, *all* that it depends on is the minimal separation distance between data points.

To begin, assume that F is in $\mathscr{RN}^s$; thus it has the representation given in (3.1). In (2.3), replace $A_{jk} = F(|x_j - x_k|)$ by the representation from (3.1). Noting that $\sum_{j,k=1}^N \xi_j \bar{\xi}_k = |\sum_{j=1}^N \xi_j|^2 = 0$ and letting

$$Q := -\sum_{j,k=1}^{N} A_{j,k}\xi_j\bar{\xi}_k \geqslant \theta \sum_{j=1}^{N} |\xi_j|^2,$$

we see that the quadratic form Q is given by

$$\begin{aligned} Q &= -\int_0^\infty \left[\sum (1 - \Omega_s(u\,|x_j - x_k|))\xi_j\bar{\xi}_k\right]\frac{d\alpha(u)}{u^2} \\ &= \int_0^\infty \left[\sum \Omega_s(u\,|x_j - x_k|)\,\xi_j\bar{\xi}_k\right]\frac{d\alpha(u)}{u^2}. \end{aligned}$$

In our last expression for Q, let us use this form [16, p. 26] for Ω_s:

$$\Omega_s(u\,|x|) = \omega_{s-1}^{-1}\int_{S_{s-1}} e^{iu\langle x,\eta\rangle}\,d\sigma_{s-1}(\eta),$$

where S_{s-1}, ω_{s-1}, and $d\sigma_{s-1}$ are, respectively, the unit sphere in $\mathbb{R}^s$, its volume, and the usual measure on it. This results in

$$\begin{aligned} Q &= \omega_{s-1}^{-1}\int_0^\infty \int_{S_{s-1}} \left[\sum_{j,k=1}^{N} e^{iu\langle x_j - x_k,\eta\rangle}\xi_j\bar{\xi}_k\right] d\sigma_{s-1}(\eta)\,\frac{d\alpha(u)}{u^2} \\ &= \omega_{s-1}^{-1}\int_0^\infty \int_{S_{s-1}} \left|\sum_{j=1}^{N} e^{iu\langle x_j,\eta\rangle}\xi_j\right|^2 d\sigma_{s-1}(\eta)\,\frac{d\alpha(u)}{u^2}. \end{aligned}$$

To complete our argument (which was inspired by a theorem in Zygmund [17, pp. 222–224], we make the following assumption: namely, that we can find a function χ defined on $\mathbb{R}^s$ having Fourier transform $\hat{\chi}$ that satisfies

(i) $\hat{\chi} \geqslant 0$

(ii) $\hat{\chi}$ is a radial function and

(iii) $d\alpha(u)/u^2 \geqslant \hat{\chi}(u)\, u^{s-1}\, du$.

From our last expression for Q, it then follows that

$$Q \geqslant \omega_{s-1}^{-1} \int_0^\infty \int_{S_{s-1}} \left| \sum_{j=1}^N e^{iu\langle x_j, \eta\rangle} \xi_j \right|^2 d\sigma_{s-1}(\eta) \hat{\chi}(u) u^{s-1}\, du. \tag{4.1}$$

By expanding the right side of (4.1) and using the usual expression for the inverse Fourier transform of $\hat{\chi}$, we arrive at

$$Q \geqslant \frac{(2\pi)^s}{\omega_{s-1}} \sum_{j,k=1}^N \xi_j \bar{\xi}_k \chi(|x_j - x_k|). \tag{4.2}$$

In (4.2), we break the right side into a sum over $j=k$ and $j \neq k$. Using a standard inequality, we obtain

$$Q \geqslant \frac{(2\pi)^s}{\omega_{s-1}} \left[\chi(0) \left(\sum_{j=1}^N |\xi_j|^2 \right) - \sum_{j \neq k} \frac{1}{2} (|\xi_j|^2 + |\xi_k|^2) |\chi(|x_j - x_k|)| \right]. \tag{4.3}$$

If we let

$$\gamma_N = \max_k \sum_{j=1}^{N}{}' |\chi(|x_j - x_k|)|, \tag{4.4}$$

where the prime indicates that $k \neq j$ in the summation, then we see from (4.3) that

$$Q \geqslant \frac{(2\pi)^s}{\omega_{s-1}} (\chi(0) - \gamma_N) \left[\sum_{j=1}^N |\xi_j|^2 \right]. \tag{4.5}$$

Assuming that $\chi(0) - \gamma_N > 0$, we come to our first lower bound on θ:

$$\theta \geqslant \frac{(2\pi)^s}{\omega_{s-1}} (\chi(0) - \gamma_N). \tag{4.6}$$

The bound for θ given in (4.6) depends on the details of the distribution of the x_j's. We will now derive a bound that depends only on the smallest distance separating points in this distribution, and not on either the relative positions of the x_j's or on their number, N.

By inspecting the expression for γ_N given in (4.4), we see that there will be some k_0 for which the sum on the right will actually equal γ_N, and that because χ is radial the sum itself depends only on the distances between the other points in the distribution and x_{k_0}. We may thus translate the whole distribution so that x_{k_0} is at the origin. Also, we lose nothing if we renumber our points so that $k_0 \to 1$ and the others are arranged so that $0 < |x_2| \leqslant |x_3| \leqslant \cdots \leqslant |x_N|$. With these conventions, we have that

$$\gamma_N = \sum_{j=2}^{N} |\chi(|x_j|)|. \tag{4.7}$$

We now define a quantity q to be one-half of the smallest distance between any two points in our distribution of x_j's; that is, $2q$ is the *separation distance* for the distribution. We will call q the *separation radius*, because it represents the radius of the largest ball that can be placed around every point of the distribution in such a way that no two balls penetrate one another. Finally, we define

$$\mathscr{E}_n := \{x_j : nq \leqslant |x_j| \leqslant (n+1)q\}. \tag{4.8}$$

One can easily get an upper bound for the cardinality of $\mathscr{E}_n$. Suppose that the distribution of points is contained in a plane with dimension $d \leqslant s$. If we put a d-dimensional ball of radius q about each $x_j \in \mathscr{E}_n$, then each ball occupies a volume of $(\omega_{d-1}/d)q^d$. Moreover, the union of these non-penetrating balls is contained in the shell that is centered at the origin and that has smaller radius $(n-1)q$ and larger radius $(n+2)q$. Obviously, the cardinality of $\mathscr{E}_n$ is bounded above by the ratio of the volume of the shell to the volume of a single ball. Hence, for a d-dimensional distribution we have

$$\operatorname{card}(\mathscr{E}_n) \leqslant (n+2)^d - (n-1)^d \leqslant 3^d n^{d-1}. \tag{4.9}$$

If we now define the quantity

$$\kappa_n := \sup\{|\chi(|x|)| : nq \leqslant |x| \leqslant (n+1)q\}, \tag{4.10}$$

we see that γ_N satisfies

$$\gamma_N \leqslant \sum_{n=1}^{\infty} \operatorname{card}(\mathscr{E}_n)\kappa_n \leqslant 3^d \sum_{n=1}^{\infty} n^{d-1}\kappa_n. \tag{4.11}$$

Assuming that the distribution of points is confined to a d-dimensional plane in $\mathbb{R}^s$, we can combine (4.6) and (4.11) to obtain this bound for θ:

$$\theta \geqslant \frac{(2\pi)^s}{\omega_{s-1}}(\chi(0) - 3^d \Sigma(q)), \qquad \text{where} \quad \Sigma(q) = \sum_{n=1}^{\infty} n^{d-1}\kappa_n. \tag{4.12}$$

Provided that we can find a χ for which the right side of the inequality in (4.12) is positive, we have an upper bound for θ^{-1} and, hence, for the norm of A^{-1}. We also note that although θ depends on the details of the distribution of the x_j's, its lower bound in (4.12) does so only to the extent that it is a function of q, the separation radius.

We now turn to the task of using the method described above to get norm estimates for A^{-1} in several examples.

V. Examples with Measures Having Polynomial Decay

In this section, we will illustrate the method sketched in Section IV by obtaining norm estimates for the matrix A^{-1} arising in connection with functions generated by measures bounded below by certain polynomially decaying measures. Specifically, we will prove that for functions in $\mathscr{R}\mathscr{N}^s$ arising from measures bounded below by polynomially decaying measures, there exist norm estimates that depend *only* upon the separation distance. The proof that we give is based both on the method sketched in Section IV and on the proof that Ball [1] used in getting the bound when the function in $\mathscr{R}\mathscr{N}^s$ is taken to be $F(r)=r$. (See (5.5) below.) Cases in which measures decay exponentially cannot be treated so simply, and a different method must be employed. For purposes of comparison, we use such a method to derive norm estimates for A^{-1} in the case where $F(r)=r$ and $s=3$. In Section VI, we will treat specific, "standard" examples in which the measures have exponential decay.

We need to be somewhat more precise about the measures that we will study in this section. We will say that a measure $d\alpha$ is *polynomially bounded below* if there exists a polynomial $P(u)$ that is positive on the interval $[0, \infty)$ and that satisfies

$$\frac{d\alpha(u)}{u^2} \geqslant \frac{u^{s-1}}{P(u)}\, du. \tag{5.1}$$

There are many examples of such measures. In particular, the one generating $F_1(r)=r$, $d\alpha_1(u)=(4/\pi)\, du$, is polynomially bounded below.

The approach used in [1] to find bounds on the norm of the matrix A^{-1} associated with $F_1(r)$ can be easily described in terms of our method. Essentially, a *specific* function χ that satisfies the conditions of Section IV and also has *compact support* is given. It is possible to generate a whole class of functions satisfying these criteria, as we see from the result below.

PROPOSITION 5.1. *Fix $a>0$. Let $\psi(x) \not\equiv 0$ be a real-valued, infinitely differentiable function defined on $\mathbb{R}^s$. Suppose that ψ depends only on $|x|$ and that it has support contained in the closed ball $|x| \leqslant a/2$. The function*

$$\chi(x) := \psi * \psi(x),$$

where $*$ denotes the usual convolution product of two functions, has the following properties:

(i) *χ is an infinitely differentiable, radial function;*

(ii) *χ has support contained in the closed ball $|x| \leqslant a$; in addition, $\chi(0)>0$;*

(iii) *the Fourier transform of χ, $\hat{\chi}$, is a nonnegative Schwartz function, and therefore falls to zero at infinity faster than any polynomial.*

Proof. The proof follows from standard Fourier analytic techniques, and so we omit the details. ▮

Suppose that $F \in \mathscr{RN}^s$ has the representation (3.1), with $F(0) \geqslant 0$ and $d\alpha$ polynomially bounded below. If the distribution of points has separation distance $2q$, where q is the separation radius that we defined in Section IV, then we may choose the χ required by our method in the following way. Let χ be one of the functions whose existence is established in Proposition 5.1, and choose the parameter $a<q$. The function χ then has its support contained in $|x|<2q$. In addition, because $\hat{\chi}$ is a Schwartz function and decays rapidly as $|x| \to \infty$, we can find a constant $c>0$ such that

$$\frac{1}{P(u)} \geqslant c\hat{\chi}(u) \geqslant 0 \tag{5.2}$$

for all $u \geqslant 0$ and for every fixed polynomial $P(u)>0$ on $u \geqslant 0$. (The constant c depends on P, of course.) Combining (5.1) and (5.2), we have

$$\frac{d\alpha(u)}{u^2} \geqslant \frac{u^{s-1}}{P(u)}\, du \geqslant u^{s-1} c\hat{\chi}(u)\, du.$$

If we absorb the constant c into the function χ, then

$$\frac{d\alpha(u)}{u^2} \geqslant u^{s-1}\hat{\chi}(u)\, du. \tag{5.3}$$

Thus, χ satisfies the conditions imposed on it in the previous section, including tacitly assumed continuity requirements. Moreover, from (4.4) and the fact that the support of χ is contained in $|x|<2q$, we see that the quantity $\gamma_N = 0$, because $|x_j - x_k| \geqslant 2q$, the separation distance. Next, from this equation, (4.6), and $\chi(0) > 0$, we get

$$\theta \geqslant \frac{(2\pi)^s}{\omega_{s-1}} \chi(0) > 0.$$

Finally, from this inequality and Lemma 2.3, we arrive at

$$\|A^{-1}\| \leqslant \frac{\omega_{s-1}}{(2\pi)^s \chi(0)}. \tag{5.4}$$

Thus we have shown the interpolation matrix for F, A, is invertible, and that the norm of its inverse is bounded above by a quantity that depends only on F (through $d\alpha$) and on the separation radius q. We collect these results below.

THEOREM 5.2. *Let $F \in \mathscr{R}\mathscr{N}^s$ and let $F(0) \geqslant 0$. If in the representation* (3.1) *the measure $d\alpha$ is polynomially bounded below, then the interpolation matrix A, which has j, k-entry equal to $F(|x_j - x_k|)$, is invertible. Moreover, $\|A^{-1}\|$ is bounded above by a quantity that depends only upon F and the separation radius q for the distribution of data points, $\{x_j\}_{j=1}^N$.*

Rather than give an example that utilizes the construction upon which Theorem 5.2 is based, we will refer the reader to [1], where it was shown that, for s odd, an s-dimensional distribution of points in $\mathbb{R}^s$ leads to the estimate

$$\|A^{-1}\| \leqslant \frac{3.55s}{2^s q} \binom{s-1}{(1/2)(s-1)}, \tag{5.5}$$

when the function $F \in \mathscr{R}\mathscr{N}^s$ is chosen to be $F_1(r) = r$. For purposes of comparison, we note that, when $s = 3$, (5.5) becomes

$$\|A^{-1}\| \leqslant 2.66 q^{-1}. \tag{5.5'}$$

The construction leading up to Theorem 5.2 will not work for measures that decay exponentially fast, for the simple reason that one cannot obtain a bound like (5.3) using a χ that is compactly supported. The reason is that $\hat{\chi}(u)$ decaying exponentially implies that its inverse Fourier transform χ is an *analytic* function in a region of $\mathbb{C}^s$ containing $\mathbb{R}^s$. Thus χ could have compact support if and only if $\chi(r) \equiv 0$.

We will deal with F's generated by exponentially decaying measures in the next section. Since their treatment is somewhat involved, we will now illustrate the method used in the case of exponential decay, but with the function $F_1(r)=r$. We will also take $F_1 \in \mathscr{R}\mathscr{N}^3$; the measure generating this function was found in Section III, and is $d\alpha(u)=(4/\pi)\,du$. (See (3.11).)

According to the method described in Section III, we need to find a radial function $\hat{\chi}(u)$ that decays rapidly enough to satisfy the inequality

$$0 \leqslant \hat{\chi}(u) \leqslant \frac{4}{\pi u^4}.$$

We can, in fact, find a class of rational functions that meet the criteria, namely $\hat{\chi}_\beta(u)=4\pi^{-1}(\beta^4+u^4)^{-1}$, $\beta>0$. The parameter that we choose for a particular set of data will depend on the separation radius, q.

We begin by finding the inverse Fourier transform of $\hat{\chi}_\beta$, χ_β. This is given by

$$\chi_\beta(|x|)=(2\pi)^{-3}\int_{\mathbb{R}^3} e^{-i\langle x,\xi\rangle}\hat{\chi}_\beta(|\xi|)\,d^3\xi.$$

Doing the radial part of the integral and manipulating the result, we obtain

$$\begin{aligned}\chi_\beta(r)&=\frac{1}{2\pi^2}\int_0^\infty \frac{\sin(ru)}{ru}\,\hat{\chi}_\beta(u)u^2\,du\\ &=\frac{1}{2\pi^2 r}\int_0^\infty u\hat{\chi}_\beta\sin(ru)\,du\\ &=\frac{2}{\pi^3 r}\int_0^\infty \frac{u\sin(ru)}{\beta^4+u^4}\,du\\ &=\frac{1}{\pi^3 r}\,\mathrm{Im}\left\{\int_{-\infty}^{\infty}\frac{ue^{iru}}{\beta^4+u^4}\,du\right\}.\end{aligned}$$

Using standard residue arguments, we arrive at the representation

$$\chi_\beta(r)=\frac{1}{\pi^2 r\beta^2}\,e^{-(\sqrt{2}/2)\beta r}\sin\left(r\frac{\sqrt{2}}{2}\beta\right). \tag{5.6}$$

Employing the notation of Section IV and combining (5.6), (4.10), and (4.12), we easily see that θ has the lower bound

$$\theta \geqslant \beta^{-1}\left(\sqrt{2}-2\frac{3^d}{\beta q}\sum_{n=1}^{\infty} n^{d-2}e^{-(\sqrt{2}/2)\beta nq}\right), \tag{5.7}$$

where we can take $d=1, 2$, or 3. We can simplify this expression by letting $w=(\sqrt{2}/2)\beta q$ in (5.7). The result is the inequality

$$\theta \geqslant \frac{q}{w}\left(1-\frac{3^d}{w}\sum_{n=1}^{\infty} n^{d-2}e^{-nw}\right). \tag{5.8}$$

TABLE I

$d = \cdots$	$w = \cdots$	$\theta \geqslant \cdots$	$\|A^{-1}\| \leqslant \cdots$
1	1.93	$0.392q$	$2.55q^{-1}$
2	2.83	$0.283q$	$3.53q^{-1}$
3	3.79	$0.219q$	$4.56q^{-1}$

Of course, we can sum the series in (5.8). Doing so, we obtain

$$\theta \geqslant \frac{q}{w}(1 - g_d(w)), \tag{5.9}$$

where the function g_d is given by

$$g_d(w) = \begin{cases} -(3/w)\log(1 - e^{-w}) & \text{for } d = 1, \\ (9/w))e^{-w}(1 - e^{-w})^{-1} & \text{for } d = 2, \\ (27/w)e^{-w}(1 - e^{-w})^{-2} & \text{for } d = 3. \end{cases} \tag{5.10}$$

Because β was an arbitrary positive number, w is also arbitrary. We now choose w so that the quantity on the right in (5.9) is a maximum. Choosing w in this way results in Table I.

As Table I shows, our method produces an upper bound for $\|A^{-1}\|$ that is comparable to (5.5′). In higher dimensions, similar calculations can be used to estimate $\|A^{-1}\|$. Although we have not carried them out, we believe that they too would yield results comparable to (5.5), but with somewhat larger numerical factors. The increased size of these factors is expected because of the crudeness of the estimate (4.9) for the "packing" constant $\operatorname{card}(\mathscr{E}_n)$ used in (4.11).

VI. Examples with Measures Having Exponential Decay

There are two functions in $\mathscr{R}\mathscr{N}^3$ representative of those commonly used in interpolation problems; these are $F_2(r) = \sqrt{1 + r^2}$ and $F_3(r) = \log(1 + r^2)$. We wish to apply the method developed in earlier sections to estimate $\|A^{-1}\|$ for these two functions. From the results of Section III, we see that both of these are generated by measures with exponential decay. As we noted in Section V, exponentially decaying measures cannot be bounded below by a measure using a compactly supported function χ, so the method used to prove Theorem 5.2 is not available. To get around this, we will derive upper bounds for $\|A^{-1}\|$ using the technique employed to get the estimates given in Table I. As before, we will adopt the notation used in Section IV. Our aim is to obtain a lower bound for θ, which itself is the infimum of the quadratic form Q.

We will start with the function $F_2(r)$. In Section III, we found that for F_2 the measure appearing in (3.1) is

$$d\alpha_2(u) = (2u^2/\pi) K_2(u)\, du. \tag{6.1}$$

We must find a function $\hat{\chi}(u)$ such that $u^{-2}\, d\alpha_2(u) \geqslant \hat{\chi}(u) u^2\, du$. To this end, consider the function $\hat{\chi}_2$ defined by

$$\hat{\chi}_2(u) \equiv (2/\pi)(u^2+\beta^2)^{-1} K_2(\sqrt{u^2+\beta^2}), \qquad \beta > 0. \tag{6.2}$$

Since K_2 is a decreasing function, as is u^{-2}, we have that $d\alpha_2(u)/u^2 \geqslant \hat{\chi}_2(u)\, u^2\, du$. As in Section V, the parameter β will be chosen later; it will depend on the separation radius, q.

We must now find the inverse $\mathbb{R}^3$-Fourier transform of $\hat{\chi}_2$. We begin by noting that this function is *radial* and, by a calculation similar to that used to derive (5.6), we see that its Fourier transform will be the radial function

$$\chi_2(r) = (2r\pi^2)^{-1} \int_0^\infty u\hat{\chi}_2(u) \sin ru\, du. \tag{6.3}$$

Inserting (6.2) in (6.3), we find that

$$\chi_2(r) = (r\pi^3)^{-1} \int_0^\infty u(u^2+\beta^2)^{-1} K_2(\sqrt{u^2+\beta^2}) \sin ru\, du. \tag{6.4}$$

Using the inverse sine transform formula and the sine transform pair in [2, p. 75, No. 35], we see that

$$\chi_2(r) = (2\beta\pi^2)^{-1} \exp(-\beta\sqrt{1+r^2}). \tag{6.5}$$

If we assume that the x_j's in the data set are confined to a 2-dimensional plane in $\mathbb{R}^3$, then using (4.12) and the observation that $\chi_2(r)$ is a decreasing function, we find that

$$\theta_2 \geqslant 2\pi^2 \left(\chi_2(0) - 9 \sum_{n=1}^{\infty} n\chi_2(nq) \right), \tag{6.6}$$

where q is the separation radius for the given distribution of points. We must estimate the sum

$$\Sigma_2(q) \equiv \sum_{n=1}^{\infty} n\chi_2(nq) \tag{6.7}$$

appearing in (6.6). To do so, we will need this lemma, which we will also use later.

LEMMA 6.1. *Fix $\beta>0$ and $q>0$, and let*

$$p=q^2(1+\sqrt{1+q^2})^{-1}. \tag{6.8}$$

For $k=0, 1, 2, \ldots,$ we have

$$\sum_{n=1}^{\infty} n^k e^{-\beta\sqrt{1+n^2q^2}} \leqslant e^{-\beta}(-\beta)^{-k} D_p^k\left(\frac{1}{e^{p\beta}-1}\right). \tag{6.9}$$

In particular, when $k=1$ we have

$$\sum_{n=1}^{\infty} n e^{-\beta\sqrt{1+n^2q^2}} \leqslant (1/4)e^{-\beta}\operatorname{csch}^2(p\beta/2). \tag{6.10}$$

Proof. We begin with the observation that

$$\sqrt{1+n^2q^2}=1+nq^2\left(\frac{n}{1+\sqrt{1+n^2q^2}}\right)>1+np,$$

where p is given in (6.8). We thus obtain this bound for the right side of (6.9):

$$\sum_{n=1}^{\infty} n^k e^{-\beta\sqrt{1+n^2q^2}} \leqslant e^{-\beta}\sum_{n=1}^{\infty} n^k e^{-np\beta}=e^{-\beta}(-\beta)^{-k}D_p^k\left[\sum_{n=1}^{\infty} n^k e^{-np\beta}\right].$$

Summing the bracketed series above, we arrive at the inequality

$$\sum_{n=1}^{\infty} n^k e^{-\beta\sqrt{1+n^2q^2}} \leqslant e^{-\beta}(-\beta)^{-k}D_p^k\left(\frac{e^{-p\beta}}{1-e^{-p\beta}}\right),$$

from which (6.9) follows immediately. To obtain (6.10), merely do the required differentiations in (6.9) and then simplify. ∎

From (6.5), (6.7), and (6.10), we see that

$$\Sigma_2(q)\leqslant(8\beta\pi)^{-2}e^{-\beta}\operatorname{csch}^2(p\beta/2).$$

Combining this inequality with (6.5) and (6.6) yields

$$\theta_2\geqslant\beta^{-1}e^{-\beta}(1-(9/4)\operatorname{csch}^2(p\beta/2)). \tag{6.11}$$

Up to now β has been a free, positive parameter. The best choice for β would be that value which makes the right side of (6.11) a maximum. This requires solving a rather messy maximization problem, however, and we will take the simpler path of choosing β so that

$$\beta = 3/p, \tag{6.12}$$

which gives us that

$$1-(9/4)\operatorname{csch}^2(p\beta/2) = 1-(9/4)\operatorname{csch}^2(1.5) \approx 0.504.$$

Using this and (6.12) in (6.11) results in the following lower estimate for θ_2,

$$\theta_2 \geqslant (0.168p)e^{-3/p}. \tag{6.13}$$

Here, p is the function of q given in (6.8). We point out that our lower bound depends only on the separation radius. We close our discussion of $F_2(r)$ by noting that for *large* separation radius q, $p \approx q$ and the lower bound for θ_2 behaves like $0.2q$. For *small* q, $p \approx q^2/2$, and the lower bound in (6.13) behaves like $(0.1q^2)e^{-6/q^2}$. For large q, we thus expect good interpolation properties for the scattered data problem when F_2 is used as the radial function. When q is small, our lower bound indicates that we can expect very poor interpolation properties for F_2. These results suggest that, for q small, one would obtain much better interpolation properties using the function $\sqrt{4q^2+r^2}$, which is easily seen to have $\|A^{-1}\| \leqslant 6/q$.

Let us now turn our attention to the function $F_3(r) = \log(1+r^2)$. Again we will discuss the case of data confined to a 2-dimensional set in $\mathbb{R}^3$. In Section III, we showed that for F_3 the measure appearing in (3.1) was

$$d\alpha_3 = 2u(1+u)e^{-u}\,du.$$

For purposes of estimating θ_3 in this case, we begin by noting that

$$\begin{aligned} d\alpha_3(u) &\geqslant 2u^4[(u^2+\beta^2)^{-1} + (u^2+\beta^2)^{-3/2}]e^{-\sqrt{u^2+\beta^2}}\,du \\ &\equiv u^4\hat{\chi}_3(u)\,du, \end{aligned} \tag{6.14}$$

where β is again an arbitrary positive number. Using the analogue of (6.3) and the Fourier sine transform pair found in [2, p. 112, No. 42], one finds $\hat{\chi}_3$'s inverse $\mathbb{R}^3$-Fourier transform, χ_3, is

$$\chi_3(r) = (1/\pi)^2 K_0(\beta\sqrt{1+r^2}). \tag{6.15}$$

Here K_0 is the order 0 modified Bessel function of the second kind.

As in the previous case, this function is a decreasing one, and the sum we need to estimate is

$$\Sigma_3(q) \equiv \sum_{n=1}^{\infty} n\chi_3(nq) = \pi^{-2} \sum_{n=1}^{\infty} nK_0(\beta\sqrt{1+n^2q^2}). \tag{6.16}$$

To do this, first replace K_0 with the integral representation for it found in [15, p. 185], and then, in the formula obtained, interchange sum and integral to finally arrive at

$$\Sigma_3(q) = (1/\pi)^2 \int_1^{\infty} \left(\sum_{n=1}^{\infty} ne^{-t\beta\sqrt{1+n^2q^2}} \right) (t^2-1)^{-1/2}\, dt. \tag{6.17}$$

We can estimate the sum (6.16) using (6.17) and the inequality (6.10), with $\beta \to t\beta$. The inequality we get is

$$\Sigma_3(q) \leqslant (1/2\pi)^2 \int_1^{\infty} e^{-t\beta} \operatorname{csch}^2(tp\beta/2)(t^2-1)^{-1/2}\, dt, \tag{6.18}$$

where p is given in terms of q in (6.8). Because $\operatorname{csch}^2(tp\beta/2)$ is a decreasing function of t when $t>0$, we have that for $t \geqslant 1$, $\operatorname{csch}^2(tp\beta/2) \leqslant \operatorname{csch}^2(p\beta/2)$. Using this in (6.18) gives us

$$\Sigma_3(q) \leqslant (1/2\pi)^2 \operatorname{csch}^2(p\beta/2) \int_1^{\infty} e^{-t\beta}(t^2-1)^{-1/2}\, dt. \tag{6.19}$$

Recognizing the integral as $K_0(\beta)$ (see [15, p. 185]), we come to the following estimate for Σ_3:

$$\Sigma_3(q) \leqslant (1/2\pi)^2 \operatorname{csch}^2(p\beta/2) K_0(\beta). \tag{6.20}$$

We can now estimate θ_3, the lower bound for Q when $F = F_3$. From (4.12), the estimate is

$$\theta_3 \geqslant 2\pi^2(\chi_3(0) - 9\Sigma_3(q)). \tag{6.21}$$

From (6.15), (6.20), and (6.21), we see that

$$\theta_3 \geqslant 2K_0(\beta)(1-(7/4)\operatorname{csch}^2(p\beta)). \tag{6.22}$$

Again we are faced with a situation where $\beta > 0$ is arbitrary. Obviously, the best choice for β is that value which for fixed p results in a maximum for the left side of (6.22). As before, this results in a very messy maximization problem. Rather than solving this problem, we shall simply choose $\beta = 3/p$, as we did in our previous problem. Doing so yields

$$\theta_3 \geqslant 1.01K_0(3/p) \tag{6.23}$$

as a lower estimate for θ_3. Here again p is given by (6.8). As before, the lower bound depends only on the separation distance q. We also again have that for small q the behavior of p is $p \approx q^2/2$, and that for large q, $p \approx q$. Consequently, when q is small, $3/p$ is large, and we can evaluate $K_0(3/p)$ by using K_0's large-argument asymptotic formula [15, p. 202] to obtain

$$K_0(3/p) \approx (p\pi/3)^{1/2} e^{-3/p}.$$

Using this, (6.23), and $p \approx q^2/2$ yields

$$\theta_3(q) \geqslant 0.729 q e^{-6/q^2} \tag{6.24}$$

as our lower estimate for q small. When $p \approx q$ is large, we have [15, p. 80]

$$K_0(3/p) \approx -\log[3/(2q)],$$

which implies that for q large

$$\theta_3(q) \geqslant 1.01 \log(2q/3). \tag{6.25}$$

The lower bounds in (6.24) and (6.25) indicate that the remarks made concerning the interpolation properties of F_2 apply to F_3 as well, except that when q is small we expect slightly better behavior for F_3 and that when q is large the reverse should be true.

All that we have said so far applies to two-dimensional distributions of points. Distributions that are one or three dimensional can be dealt with in a similar way. For example, in the three-dimensional case the chief difference is that, from (4.12), the lower bound on θ is $\theta \geqslant 2\pi^2(\chi(0) - \sum_{n=1}^{\infty} 27n^2\chi(nq))$ instead of (6.6) or (6.21), and so one must use the $k=2$ case in Lemma 6.1. Bearing in mind these differences, one can easily show that for three-dimensional distributions

$$\theta_2 \geqslant 0.117 p e^{-4/p} \quad \text{and} \quad \theta_3 \geqslant 0.467 K_0(4/p). \tag{6.26}$$

TABLE II

F_2

$d = \cdots$	Restrictions	$\theta \geqslant \cdots$	$\|A^{-1}\| \leqslant \cdots$
2	$p = q^2(1+\sqrt{1+q^2})^{-1}$	$(0.168p)e^{-3/p}$	$(5.95/p)e^{3/p}$
2	$q \to \infty$	$0.2q$	$6/q$
2	$q \approx 0$	$(0.1q^2)e^{-6/q^2}$	$(12/q^2)e^{6/q^2}$
3	$p = q^2(1+\sqrt{1+q^2})^{-1}$	$0.117pe^{-4/p}$	$(8.55/p)e^{4/p}$
3	$q \to \infty$	$0.1q$	$9/q$
3	$q \approx 0$	$0.059q^2e^{-8/q^2}$	$(17/q^2)e^{8/q^2}$

TABLE III

F_3

$d = \cdots$	Restrictions	$\theta \geqslant \cdots$	$\|A^{-1}\| \leqslant \cdots$
2	$p = q^2(1+\sqrt{1+q^2})^{-1}$	$1.01 K_0(3/p)$	$0.99(K_0(3/p))^{-1}$
2	$q \to \infty$	$1.01 \log(2q/3)$	$0.99(\log(2q/3))^{-1}$
2	$q \approx 0$	$0.729 q e^{-6/q^2}$	$(1.37/q) e^{6/q^2}$
3	$p = q^2(1+\sqrt{1+q^2})^{-1}$	$0.467 K_0(4/p)$	$2.14(K_0(4/p))^{-1}$
3	$q \to \infty$	$0.467 \log(q/2)$	$2.14(\log(q/2))^{-1}$
3	$q \approx 0$	$0.886 q e^{-8/q^2}$	$(1.13/q) e^{8/q^2}$

Here we chose the parameter $\beta = 4/p$, where p is as in (6.8). As expected, the behavior is somewhat worse than the two-dimensional case. We did not carry out calculations for one-dimensional distributions. Doing them introduces nothing new. We summarize our results in Tables II and III.

VII. Bounds on Condition Numbers

In this section, we first obtain a bound on the norm for the interpolation matrix A corresponding to a function F in $\mathscr{R}\mathscr{N}^s$. This bound is then used in conjunction with our earlier estimates for $\|A^{-1}\|$ to get a bound on the condition number for A corresponding to one of the functions $F_1(r) = r$, $F_2(r) = \sqrt{1+r^2}$, and $F_3(r) = \log(1+r^2)$. For $F_1(r) = r$, these bounds have already been obtained in [1].

Throughout the section, we will let $S = \{x_j\}_{j=1}^N \subset \mathbb{R}^s$, and we will set $D = \max_{j \neq k} |x_j - x_k|$, $x_j \in S$, the diameter of S. We can now give our bound for $\|A\|$.

Proposition 7.1. *Let $S \subset \mathbb{R}^s$ be a data set of diameter D, separation radius q, and cardinality N. If $F \in \mathscr{R}\mathscr{N}^s$, with $F(0) \geqslant 0$, then*

$$\|A\| = \|[F(|x_j - x_k|)]\| \leqslant NM, \qquad \text{where} \quad M = \max_{j \neq k} F(|x_j - x_k|). \tag{7.1}$$

In addition, $\|A\|$ also satisfies the bound

$$\|A\| \leqslant M \left(\frac{D+2q}{2q}\right)^s \tag{7.2}$$

where M is as in (7.1). *Finally, if F is increasing, then*

$$\|A\| \leqslant F(D) \left(\frac{D+2q}{2q}\right)^s. \tag{7.3}$$

Proof. The proof here is a generalization of a similar one used in [1]. We begin by noting that the operator norm of a matrix is dominated by its Hilbert–Schmidt norm, and so we have that

$$\|A\| \leqslant \left(\sum_{j,k=1}^{N} F^2(|x_j - x_k|) \right)^{1/2} \leqslant NM,$$

which establishes (7.1).

Since the separation radius of S is q, each point in S may be placed at the center of a ball of radius q. Moreover, these balls will not penetrate each other, and they will occupy a total volume of $Ns^{-1}\omega_{s-1}q^s$. On the other hand, since the diameter of S is D, the region occupied by them will also be contained in a large ball of diameter $D+2q$ and volume $s^{-1}\omega_{s-1}((D+2q)/2)^s$. Hence,

$$Ns^{-1}\omega_{s-1}q^s \leqslant s^{-1}\omega_{s-1}\left(\frac{D+2q}{2}\right)^s,$$

from which we see that the number of points in S, N, satisfies

$$N \leqslant \left(\frac{D+2q}{2q}\right)^s. \tag{7.4}$$

Combining (7.1) and (7.4) yields (7.2). For F increasing, which wil be the case whenever $F'(\sqrt{r})$ is completely monotonic (see [12]), the inequality (7.3) is an immediate consequence of (7.2). ∎

Several remarks are in order. First, if $F(0)=0$, the factor of N in (7.1) may be replaced by $\sqrt{N(N-1)}$, because all the N entries of the diagonal of A vanish and so do not contribute to the Hilbert–Schmidt norm of A. Second, when $F(r)=F_1(r)=r$, the inequality in (7.1), with $N \to \sqrt{N(N-1)}$ and $M=D$, becomes

$$\|A\| \leqslant D\sqrt{N(N-1)}. \tag{7.5}$$

This result agrees with that given in [1], where it is also shown that the inequality (7.5) is sharp when points in the data set are uiformly distributed.

Recall that the *condition number* of A is defined to be the product $\|A\|\,\|A^{-1}\|$. By combining the bounds for $\|A\|$ with the ones derived in Sections V and VI for $\|A^{-1}\|$, we can easily obtain upper estimates on the condition numbers for the three functions $F_1(r)=r$, $F_2(r)=\sqrt{1+r^2}$, and $F_3(r)=\log(1+r^2)$, with the data set S being either planar or some unrestricted subset of $\mathbb{R}^3$. Our results are given in Table IV. In that table, $s=3$ because all three functions are regarded as being in $\mathscr{RN}^3$. As usual,

TABLE IV

Upper Bounds on Condition Numbers

$d = \cdots$	$F_1(r) = r$	$F_2(r) = \sqrt{1+r^2}$	$F_3(r) = \log(1+r^2)$
2	$3.53Dq^{-1}\left(\frac{D+2q}{2q}\right)^3$	$5.95\frac{\sqrt{1+D^2}}{p}e^{3/p}\left(\frac{D+2q}{2q}\right)^3$	$0.99\frac{\log(1+D^2)}{K_0(3/p)}\left(\frac{D+2q}{2q}\right)^3$
3	$4.56Dq^{-1}\left(\frac{D+2q}{2q}\right)^3$	$8.55\frac{\sqrt{1+D^2}}{p}e^{4/p}\left(\frac{D+2q}{2q}\right)^3$	$2.14\frac{\log(1+D^2)}{K_0(4/p)}\left(\frac{D+2q}{2q}\right)^3$

q denotes the separation radius, which is half the minimal separation distance, and $p = q^2(1+\sqrt{1+q^2})^{-1}$. Finally, d is the dimension of the smallest affine subspace containing the data set S.

ACKNOWLEDGMENTS

The authors gratefully acknowledge helpful conversations with Keith Ball, Richard Beatson, Philip Smith, and Jay Walton. In addition, we thank the referee for several helpful suggestions.

REFERENCES

1. K. BALL, The conditioning number for radial basis interpolation, preprint.
2. A. ERDÉLYI, W. MAGNUS, F. OBERHETTINGER, AND F. G. TRICOMI, "Tables of Integral Transforms," McGraw–Hill, New York, 1954.
3. J. DUCHON, Interpolation des fonctions de deux variables suivant le principe de la flexion des plaques minces, *RAIRO Anal. Numér.* **10** (1976), 5–12.
4. J. DUCHON, Splines minimizing rotation invariant semi-norms in Sobolov spaces, *in* "Constructive Theory of Functions of Several Variables," Oberwolfach 1976 (W. Schempp and K. Zeller, Eds.), pp. 85–100, Springer-Verlag, Berlin, 1977.
5. N. DYN, Interpolation and approximation by radial and related functions, *in* "Approximation Theory IV" (C. K. Chui, L. L. Schumaker, and J. D. Ward, Eds.), Academic Press, Orlando, FL, in press.
6. N. DYN, W. A. LIGHT, AND E. W. CHENEY, Interpolation by piece-wise linear radial basis functions, I, *J. Approx. Theory*, in press.
7. R. FRANKE, Scattered data interpolation: Tests of some methods, *Math. Comp.* **38** (1982), 181–199.
8. R. L. HARDY, Multiquadric equations of topography and other irregular surfaces, *J. Geophys. Res.* **76** (1971), 1905–1915.
9. M. J. LIGHTHILL, "Fourier Analysis and Generalised Functions," Cambridge Univ. Press, London, 1958.
10. W. R. MADYCH AND S. A. NELSON, Multivariate interpolation: A variational theory, manuscript, 1983.
11. W. R. MADYCH AND S. A. NELSON, Multivariate interpolation and conditionally positive definite functions, *Approx. Theory Appl.* **4** (1988), 77–79.

12. C. A. MICCHELLI, Interpolation of scattered data: Distances, matrices, and conditionally positive definite functions, *Constr. Approx.* **2** (1986), 11–22.
13. I. J. SCHOENBERG, Metric spaces and completely monotone functions, *Ann. of Math.* **39** (1938), 811–841.
14. F. TREVES, "Topological Vector Spaces, Distributions and Kernels," Academic Press, New York, 1967.
15. G. N. WATSON, "Theory of Bessel Functions," Cambridge Univ. Press, London, 1966.
16. J. H. WELLS AND R. L. WILLIAMS, "Embeddings and Extensions in Analysis," Ergebnisse, Vol. 84, Springer-Verlag, Berlin, 1975.
17. A. ZYGMUND, "Trigonometric Series," Vols. I and II, 2nd ed., Cambridge Univ. Press, London, 1959.