JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS 20, 110-123 (1967)

On the $\mathscr{L}_2$ n-Width of Certain Classes of Functions of Several Variables*

JOSEPH W. JEROME[1]

Mathematics Research Center, United States Army, The University of Wisconsin, Madison, Wisconsin 53706

Submitted by Ky Fan

1. INTRODUCTION

The notion of the *n-width* (or *n-dimensional diameter*) of a set $\mathscr{A}$ in a normed linear space $\mathscr{X}$ was introduced by Kolmogorov [1] as a means of characterizing the approximability of $\mathscr{A}$ by linear manifolds of finite dimension. Indeed, in order to obtain the n-width, $d_n(\mathscr{A})$, of $\mathscr{A}$ we take the maximal distance (or *deflection*)

$$E(\mathscr{A}, \mathscr{M}) = \sup_{f \in \mathscr{A}} \inf_{g \in \mathscr{M}} \| f - g \| \tag{1.1}$$

of $\mathscr{A}$ from an n-dimensional subspace $\mathscr{M}$ of $\mathscr{X}$ and then vary the subspace, and take the infimum

$$d_n(\mathscr{A}) = \inf \{E(\mathscr{A}, \mathscr{M}) : \mathscr{M} \subset \mathscr{X}, \dim \mathscr{M} = n\} \tag{1.2}$$

of the distances for all possible subspaces. The reader is referred to the survey article of Lorentz [2] and the paper of Tihomirov [3] for a summary of some important results concerning n-widths as well as for a bibliography of the subject. Some more recent results may be found in [4]-[7]. Auxiliary to the notion of n-width is the notion of an *optimal approximating n-subspace* (or *extremal subspace*), i.e., a linear manifold of dimension less than or equal to n whose deflection from $\mathscr{A}$ is equal to $d_n(\mathscr{A})$.

This paper will be concerned with the determination of n-widths and extremal subspaces for certain subsets of Hilbert space, especially of $\mathscr{L}_2(\Omega)$, Ω a bounded open connected set in m-dimensional Euclidean space. In particular, a now classical theorem of Kolmogorov [1] regarding the n-widths

* Sponsored by the Mathematics Research Center, United States Army, Madison, Wisconsin, under Contract No.: DA-31-124-ARO-D-462.

[1] The contents of this report formed part of the author's doctoral dissertation at Purdue University under Professor Michael Golomb.

of certain classes in $\mathscr{L}_2(0, 1)$ will be extended to the setting of $\mathscr{L}_2(\Omega)$. To describe Kolmogorov's result set, for $k \geqslant 1$,

$$\mathscr{A}_k(0, 1) = \{f \in \mathscr{L}_2(0, 1) : f^{(k-1)} \text{ exists and is absolutely continuous on } [0, 1] \text{ and } f^{(k)} \in \mathscr{L}_2(0, 1)\}. \tag{1.3}$$

Kolmogorov considered the classes

$$\mathscr{F}_k = \{f \in \mathscr{A}_k(0, 1) : \|f^{(k)}\| \leqslant 1\} \tag{1.4}$$

and proved

THEOREM (Kolmogorov). *Let λ_n, $n = 1, 2,...$ denote all nonzero eigenvalues of the boundary value problem*

$$(-1)^k y^{(2k)} = \lambda y$$

$$y^{(k)}(0) = y^{(k)}(1) = \cdots = y^{(2k-1)}(0) = y^{(2k-1)}(1) = 0 \tag{1.5}$$

and $\mathscr{M}_n$ the corresponding (one dimensional) eigenmanifolds. Then all λ_n are positive; if they are arranged in increasing order one has

$$d_n(\mathscr{F}_k) = \infty \quad \text{if} \quad n < k, \tag{1.6}$$

while

$$d_n(\mathscr{F}_k) = \lambda_{n-k+1}^{-1/2} \quad \text{if} \quad n \geqslant k. \tag{1.7}$$

Moreover, in the latter case, one has

$$d_n(\mathscr{F}_k) = E(\mathscr{F}_k, \mathscr{P}_k + \mathscr{M}_1 + \cdots + \mathscr{M}_{n-k}), \tag{1.8}$$

where $\mathscr{P}_k$ is the set of polynomials of degree less than k.

In the present paper it is shown that an analogous result holds for certain function classes in $\mathscr{L}_2(\Omega)$ if the boundary value problem is replaced by an appropriate (functional) variational problem over a Sobolev space. In addition, the analogy with the Kolmogorov results is made complete under assumptions of sufficient smoothness on the boundary of Ω. In this case the variational problem is equivalent to an ordinary boundary value problem for a partial differential equation. In a future paper, asymptotic estimates will be obtained for the n-widths and the exact asymptotic order exhibited in certain cases.

Golomb [8] has introduced the notion of the ellipsoid determined by a general positive self-adjoint linear operator on a Hilbert space and has obtained results concerning their optimal approximation by subspaces not necessarily of finite dimension. The importance of ellipsoids, aside from interest in them for their own sake, arises from the fact that many important function classes in $\mathscr{L}_2(\Omega)$ have the same topological closure as ellipsoids and thus the same n-widths and extremal subspaces, which are invariant under closure of the classes which are approximated. Golomb's major results, then, will form the basis for the abstract setting of this paper.

I. n-Widths in Hilbert Space

2. *Approximation of Ellipsoids*

Let $\mathscr{H}$ be an infinite dimensional complex Hilbert space with inner product (f, g) and norm $\|f\| = (f,f)^{1/2}$. If R is a linear, not necessarily bounded, operator with domain $\mathscr{D}_R$ dense in $\mathscr{H}$ then R is *positive* if $(Rf, f) \geqslant 0$ for all $f \in \mathscr{D}_R$, *strictly positive* if R is positive and one to one and *positive definite* if there exists a positive constant c such that $(Rf, f) \geqslant c(f, f)$ for all $f \in \mathscr{D}_R$. For a positive self-adjoint operator R we define the *ellipsoid* $\mathscr{R}$ determined by R:

$$\mathscr{R} = \{ f \in \mathscr{D}_R : (Rf, f) \leqslant 1 \}.$$

For such an operator, let $\lambda \to E_\lambda$ be the spectral family of R which is continuous from the left. We denote by $\mathscr{E}_\lambda$ the closed linear manifold which is the range of E_λ. We state now two results from [8]. $\mathscr{M}^\perp$ denotes the orthogonal complement of $\mathscr{M}$.

Theorem 1. *$E(\mathscr{R}, \mathscr{M}) \leqslant \delta$ if $\mathscr{M} = \mathscr{E}_{\delta-2}$ while $E(\mathscr{R}, \mathscr{M}) \geqslant \delta$ if $\mathscr{M}^\perp \cap \mathscr{E}_{\delta-2} \neq (0)$.*

Theorem 2. *$E(\mathscr{R}, \mathscr{E}_{\delta-2}) = \delta$ if and only if δ^{-2} is in the spectrum of R.*

These two results yield a particularly elegant expression for the n-widths of $\mathscr{R}$ in the special cases that R is *dispersive*, i.e., when the spectrum of R consists of isolated eigenvalues of finite multiplicity. Indeed, in this case, there exist nonnegative numbers λ_ν and finite rank projectors P_ν, $\sum P_\nu = I$, such that $R = \sum \lambda_\nu P_\nu$ with

$$\mathscr{D}_R = \left\{ f \in \mathscr{H} : \sum | \lambda_\nu |^2 \| P_\nu f \|^2 < \infty \right\}.$$

We agree, for convenience, to designate the first eigenvalue as $\lambda_0 = 0$ if 0 is an eigenvalue and as λ_1 otherwise. In the latter case we take $P_0 = 0$. If we denote by $\mathscr{M}_\nu$ the projection manifold of P_ν and set

$$N_\nu = \dim (\mathscr{M}_0 + \mathscr{M}_1 + \cdots + \mathscr{M}_\nu)$$

then we have

Theorem 2.1. *If R is dispersive then*

$$d_n(\mathscr{R}) = \infty \qquad \text{if} \qquad n < N_0, \tag{2.1}$$

while

$$d_n(\mathscr{R}) = \lambda_{\nu+1}^{-1/2} \qquad \text{if} \qquad N_\nu \leqslant n < N_{\nu+1} \qquad \nu \geqslant 0. \tag{2.2}$$

Furthermore, in the latter case,

$$d_n(\mathscr{R}) = E(\mathscr{R}, \mathscr{M}_0 + \mathscr{M}_1 + \cdots + \mathscr{M}_\nu), \tag{2.3}$$

i.e., $\mathscr{M}_0 + \mathscr{M}_1 + \cdots + \mathscr{M}_\nu$ *is an optimal approximating n-subspace.*

PROOF. Set

$$E_\lambda = \begin{cases} 0 & \text{if} \quad \lambda \leqslant 0 \\ \sum_{\lambda_\nu < \lambda} P_\nu & \text{if} \quad \lambda > 0. \end{cases}$$

Notice that $\{E_\lambda\}$ is the spectral family corresponding to R which is continuous from the left. Now from Theorem 2 it follows that

$$E(\mathscr{R}, \mathscr{M}_0 + \mathscr{M}_1 + \cdots + \mathscr{M}_\nu) = \lambda_{\nu+1}^{-1/2}, \qquad \nu \geqslant 0. \tag{2.4}$$

If $N_\nu \leqslant n < N_{\nu+1}$ then, for any linear manifold $\mathscr{M}$ of dimension n and $\epsilon > 0$, $\mathscr{M}^\perp \cap \mathscr{E}_{\lambda_{\nu+1}+\epsilon} \neq (0)$ since $\mathscr{E}_{\lambda_{\nu+1}+\epsilon}$ has dimension $\geqslant N_{\nu+1}$ and since $n < N_{\nu+1}$. By Theorem 1, however,

$$E(\mathscr{R}, \mathscr{M}) \geqslant (\lambda_{\nu+1} + \epsilon)^{-1/2} \qquad \text{for each} \qquad \epsilon > 0$$

and it follows that

$$E(\mathscr{R}, \mathscr{M}) \geqslant (\lambda_{\nu+1})^{-1/2}. \tag{2.5}$$

By (2.4) and (2.5) any n-dimensional linear manifold $\mathscr{M}$ which contains $\mathscr{M}_0 + \cdots + \mathscr{M}_\nu$ satisfies

$$E(\mathscr{R}, \mathscr{M}) = \lambda_{\nu+1}^{-1/2} \qquad \text{if} \qquad N_\nu \leqslant n < N_{\nu+1}.$$

Thus, (2.4) and (2.5) together imply (2.2) and (2.3). Finally, if $n < N_0$ and $\mathscr{M}$ is any linear manifold of dimension n then $\mathscr{M}^\perp \cap \mathscr{M}_0 \neq (0)$ and if $0 < \delta \leqslant \lambda_1$ we have, by Theorem 1,

$$E(\mathscr{R}, \mathscr{M}) \geqslant \delta^{-2}. \tag{2.6}$$

Since δ can be chosen arbitrarily small, (2.1) follows from (2.6).

REMARK 2.1. It follows from the standard characterization of the λ_ν for a positive dispersive operator R and Theorem 2.1 that the variational problem of the n-width of the ellipsoid determined by R is really the same as the well-known variational problem for eigenvalues. Indeed, we have if $N_\nu \leqslant n < N_{\nu+1}$,

$$d_n^2(\mathscr{R}) = \sup \{(Rf, f)^{-1} : f \in \mathscr{D}_R, \|f\| = 1, (f, \varphi) = 0 \text{ if } \varphi \in \mathscr{M}_0 + \cdots + \mathscr{M}_\nu\}.$$

3. *Sets Determined by Quadratic Forms*

Let $\mathscr{H}_0$ be a linear manifold which is dense in $\mathscr{H}$ and which is a Hilbert space with inner product

$$(f, g)_0 = [f, g] + (f, g), \tag{3.1}$$

where $[\cdot]$ is a positive Hermitian form on $\mathscr{H}_0$. To say that $[\cdot]$ is *positive* means that $[f, f] \geqslant 0$ for all $f \in \mathscr{H}_0$; if $[f, f]^{1/2}$ is a norm then $[\cdot]$ is said to be *strictly positive* and if $[f, f] \geqslant c(f, f)$ for some positive constant c then $[\cdot]$ is *positive definite.* Now we set

$$\mathscr{R}_0 = \{f \in \mathscr{H}_0 : [f, f] \leqslant 1\}$$

and say that $\mathscr{R}_0$ is determined by the quadratic form $[f, f]$. The main theorem of this section is Theorem 3.3 which characterizes the n-widths of $\mathscr{R}_0$ if $\mathscr{H}_0$ is imbedded compactly in $\mathscr{H}$. We need two preliminary theorems, the first of which we merely state. Its proof is routine. In what follows, $C\ell_0$ will denote closure in $\mathscr{H}_0$ and $C\ell$ will denote closure in $\mathscr{H}$.

THEOREM 3.1. *Let $\mathscr{H}_0^-$, be a linear manifold dense in $\mathscr{H}_0$, i.e., $C\ell_0\mathscr{H}_0^- = \mathscr{H}_0$. Let*

$$\mathscr{R}_0^- = \{f \in \mathscr{H}_0^- : [f, f] \leqslant 1\}.$$

Then

$$C\ell_0\mathscr{R}_0^- = C\ell_0\mathscr{R}_0. \tag{3.2}$$

It follows from this that

$$C\ell\mathscr{R}_0 = C\ell\mathscr{R}_0. \tag{3.3}$$

We relate the notion of a set determined by a quadratic form with that of an ellipsoid in

THEOREM 3.2. *There exists a positive self-adjoint operator R on $\mathscr{H}_0$ satisfying*

$$[f, g] = (Rf, g) \qquad \textit{for all} \qquad f \in \mathscr{D}_R, g \in \mathscr{H}_0 \tag{3.4}$$

such that $C\ell\mathscr{R} = C\ell\mathscr{R}_0$, where $\mathscr{R}$ is the ellipsoid determined by R and $\mathscr{D}_R$ is characterized by (3.4). R is strictly positive (positive definite) if $[\cdot]$ is. Moreover, if the injection $I : \mathscr{H}_0 \to \mathscr{H}$ is compact then R is dispersive and the equation

$$[\varphi, g] = \lambda(\varphi, g) \qquad \textit{for all} \qquad g \in \mathscr{H}_0 \tag{3.5}$$

has, for a sequence of nonnegative values λ_ν, $\lambda_\nu \to \infty$, corresponding finite *dimensional eigenmanifolds $\mathscr{M}_\nu$ which are orthogonal with dense linear span in both $\mathscr{H}_0$ and $\mathscr{H}$. The λ_ν are eigenvalues of R.*

PROOF. By a standard construction [9, pp. 332-333] there exists a self-adjoint operator $B : \mathscr{H} \to \mathscr{H}_0$ defined by

$$(g, Bf)_0 = (g, f) \qquad \text{for all} \qquad g \in \mathscr{H}_0,$$

which is bounded as a transformation from $\mathscr{H}$ to $\mathscr{H}_0$ and from $\mathscr{H}$ to $\mathscr{H}$:

$$\| Bf \| \leqslant (Bf, Bf)_0^{1/2} \leqslant \| f \| .$$

$A = B^{-1}$ exists, is self-adjoint and satisfies

$$(f, g)_0 = (Af, g) \qquad \text{for all} \qquad f \in \mathscr{D}_A, \, g \in \mathscr{H}_0. \tag{3.6}$$

Indeed, $\mathscr{D}_A$ is characterized as the set of all $f \in \mathscr{H}_0$ for which (3.6) holds. $\mathscr{D}_A$ is dense in $\mathscr{H}_0$ and A is positive definite:

$$(Af, f) \geqslant (f, f) \qquad \text{for all} \qquad f \in \mathscr{D}_A .$$

If we set $\mathscr{D}_R = \mathscr{D}_A$, $R = A - I$ then R is a positive self-adjoint operator satisfying (3.4). Moreover, it follows from Theorem 3.1 and relation 3.4, taking $\mathscr{H}_0^- = \mathscr{D}_A$, that $C\ell\mathscr{R} = C\ell\mathscr{R}_0$.

If $I : \mathscr{H}_0 \to \mathscr{H}$ is compact then sets bounded in $\mathscr{H}_0$ have compact closure in $\mathscr{H}$ and the strictly positive operator B is compact, with spectrum consisting of 0 which is not an eigenvalue and a sequence $\{\mu_\nu\}$ of eigenvalues converging to 0, in fact, $1 \geqslant \mu_1 > \mu_2 > \cdots > 0$. The eigenmanifolds corresponding to the μ_ν are known to be finite dimensional and mutually orthogonal in $\mathscr{H}$ and their closed span is all of $\mathscr{H}$. The spectrum of $A = B^{-1}$ consists precisely of the reciprocals $\lambda_\nu' = \mu_\nu^{-1}$ of the μ_ν with the same eigenmanifolds. It follows trivially that R is a positive dispersive operator whose spectrum is the spectrum of A shifted one unit to the left and whose eigenmanifolds are the $\mathscr{M}_\nu$. Thus (3.5) follows from (3.4). The orthogonality and denseness properties of the $\mathscr{M}_\nu$ in $\mathscr{H}_0$ follow directly from the corresponding properties in $\mathscr{H}$. This concludes the proof.

REMARK 3.1. By the denseness of $\mathscr{D}_R$ in $\mathscr{H}_0$ we can readily establish the following characterization of the λ_ν:

$$\lambda_\nu = \inf \{[f, f] : f \in \mathscr{H}_0, \, \| f \| = 1 \text{ and } (f, \varphi) = 0 \text{ if } \varphi \in \mathscr{M}_0 + \cdots + \mathscr{M}_{\nu-1}\}.$$

We are now ready to state the main result of this section. It is a direct consequence of Theorems 3.1 and 3.2, Remark 3.1 and the observation that n-widths and extremal subspaces are invariant under closure of the set being approximated.

THEOREM 3.3. *If the injection $I : \mathscr{H}_0 \to \mathscr{H}$ is compact then, for any subset $\mathscr{R}_0^-$ of $\mathscr{H}$ which satisfies $Cl\mathscr{R}_0^- = Cl\mathscr{R}_0$, we have the following characterization of $d_n(\mathscr{R}_0^-)$:*

$$d_n(\mathscr{R}_0^-) = \infty \qquad \textit{if} \qquad n < N_0$$

$$d_n(\mathscr{R}_0^-) = \lambda_{\nu+1}^{-1/2} \qquad \textit{if} \qquad N_\nu \leqslant n < N_{\nu+1},$$

where the λ_ν satisfy (3.5) with corresponding eigenmanifolds $\mathscr{M}_\nu$ and

$$N_\nu = \dim(\mathscr{M}_0 + \cdots + \mathscr{M}_\nu).$$

Moreover, if $N_\nu \leqslant n < N_{\nu+1}$ then

$$d_n(\mathscr{R}_0^-) = E(\mathscr{R}_0^-, \mathscr{M}_0 + \cdots + \mathscr{M}_\nu),$$

and the numbers $d_n(\mathscr{R}_0^-)$ may be characterized by

$$d_n^2(\mathscr{R}_0^-) = \sup\{[f,f]^{-1} : f \in \mathscr{H}_0, \|f\| = 1 \textit{ and } (f, \varphi) = 0 \text{ if } \varphi \in \mathscr{M}_0 + \mathscr{M}_1 + \cdots + \mathscr{M}_\nu\}.$$

4. *Application to Kolmogorov's Result*

We will outline the proof of Kolmogorov's theorem in this section using the framework of results which we have developed. The first step in the proof is to show that the eigenfunctions of the boundary value problem 1.5 form a complete orthogonal system in $\mathscr{L}_2(0, 1)$ and that the eigenvalues consist of 0, which is an eigenvalue of multiplicity k, and a sequence λ_n of simple positive eigenvalues with $\lambda_n \to \infty$. The reader is referred to [2] for the proof.

The second step of the proof is to notice that the class $\mathscr{A}_k(0, 1)$ defined in 1.3 is actually the same set as the set of functions f such that the distribution derivatives $D^j f$, $0 \leqslant j \leqslant k$, are in $\mathscr{L}_2(0, 1)$ (see [17]). Thus $\mathscr{A}_k(0, 1)$ with the inner product

$$(f, g)_k = (f, g) + [f, g]_k = \int_{(0,1)} f\bar{g} + \int_{(0,1)} f^{(k)}\overline{g^{(k)}}$$

yields a Hilbert space which we designate by $\mathscr{W}_k(0, 1)$. $\mathscr{F}_k$ is simply the set $\{f \in \mathscr{W}_k(0, 1) : [f, f]_k \leqslant 1\}$.

In step three we proceed, as in Theorem 3.2, to construct a positive self-adjoint operator R on $\mathscr{W}_k(0, 1)$ satisfying

$$[f, g]_k = (Rf, g) \qquad \text{for all} \qquad f \in \mathscr{D}_R, g \in \mathscr{W}_k(0, 1)$$

such that $C\ell\mathscr{R} = C\ell\mathscr{F}_k$, where $\mathscr{R}$ is the ellipsoid determined by R. We then show that R is a dispersive operator with eigenvalues precisely those of (1.5). This is easily accomplished through the formula, obtained by integration by parts,

$$\int_{(0,1)} z\overline{f^{(k)}} = \lambda \int_{(0,1)} y\bar{f} \qquad z = y^{(k)},$$

which is valid for $f \in \mathscr{A}_k(0, 1)$ and y a solution of (1.5), and through the characterization of $\mathscr{D}_R$.

The last step of the proof is to observe that $\mathscr{R}$ and $\mathscr{F}_k$ have the same n-widths and extremal subspaces. Thus, we can apply Theorem 2.1 with $N_\nu = k + \nu$ and $\mathscr{M}_0 = \mathscr{P}_k$ to obtain (1.6), (1.7), and (1.8). This completes the outline of the proof.

REMARK 4.1. The eigenfunctions of (1.5) are (real) analytic functions on (0, 1); in fact they can be extended to entire functions. In extending the theory to $\mathscr{L}_2(\Omega)$ we will see that the real analyticity of the eigenfunctions is preserved. This is a valuable property since the eigenfunctions span extremal subspaces.

II. n-WIDTHS IN $\mathscr{L}_2(\Omega)$

5. *The General Problem*

Let Ω be a bounded open connected set in m-dimensional Euclidean space and let k be an integer satisfying $k \geqslant 1$. Let $\mathscr{L}_2(\Omega)$ be the Hilbert space of complex-valued functions f square integrable on Ω with inner product

$$(f, g)_{\mathscr{L}_2} = \int_\Omega f\bar{g}$$

and norm

$$\| f \| = \left(\int_\Omega | f |^2\right)^{1/2},$$

and let $\mathscr{A}_k(\Omega)$ be the set of functions f in $\mathscr{L}_2(\Omega)$ such that f has distribution derivatives of order $\leqslant k$ in $\mathscr{L}_2(\Omega)$. We define for nonnegative integers α_i,

$$\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_m), \qquad | \alpha | = \sum_{i=1}^m \alpha_i, \qquad C_\alpha = \frac{| \alpha |!}{\alpha_1!\, \alpha_2! \cdots \alpha_m!}$$

and any $f \in \mathscr{L}_2(\Omega)$ the distribution derivative

$$D^\alpha f = \frac{\partial^{|\alpha|} f}{(\partial x_1)^{\alpha_1} (\partial x_2)^{\alpha_2} \cdots (\partial x_m)^{\alpha_m}}.$$

A natural extension of Kolmogorov's result is the determination of the n-widths and extremal subspaces for the class

$$\left\{ f \in \mathscr{A}_k(\Omega) : \int_\Omega \sum_{|\alpha|=k} C_\alpha \mid D^\alpha f \mid^2 \leqslant 1 \right\},$$

where the constants C_α enable one to consider all possible permutations in the order of differentiation. $\mathscr{A}_k(\Omega)$ with the inner product

$$(f, g)_{\mathscr{W}_k} = \int_\Omega \sum_{|\alpha| \leqslant k} C_\alpha D^\alpha f D^\alpha \bar{g}$$

is a well-known Hilbert space which we designate by $\mathscr{W}_k(\Omega)$. If we designate by $\mathring{\mathscr{H}}_k(\Omega)$ the completion in $\mathscr{W}_k(\Omega)$ of the infinitely differentiable functions $C_c^\infty(\Omega)$ with compact support in Ω then we can consider a somewhat more general approximation problem by considering any Hilbert space $\mathscr{V}_k$ satisfying $\mathring{\mathscr{H}}_k(\Omega) \subset \mathscr{V}_k \subset \mathscr{W}_k(\Omega)$ and determining the n-widths and extremal subspaces for the class

$$\mathscr{R}_k = \{ f \in \mathscr{V}_k : [f, f]_k \leqslant 1 \},$$

where

$$[f, f]_k = \int_\Omega \sum_{|\alpha|=k} C_\alpha \mid D^\alpha f \mid^2.$$

We will do this in this section under the assumptions

(i) the norm $(f, f)_k^{1/2} = \{(f, f)_{\mathscr{L}_2} + [f, f]_k\}^{1/2}$ is equivalent to the standard norm on $\mathscr{V}_k$ induced by $\mathscr{W}_k(\Omega)$, and,

(ii) the injection $I : \mathscr{V}_k \to \mathscr{L}_2(\Omega)$ is compact.

Now these assumptions hold for $\mathring{\mathscr{H}}_k(\Omega)$ without any further conditions on Ω but this is not true for $\mathscr{W}_k(\Omega)$. In the next section we will apply the results of this section to the special cases of $\mathring{\mathscr{H}}_k(\Omega)$ and $\mathscr{W}_k(\Omega)$ and also give conditions on Ω which will guarantee that assumptions (i) and (ii) hold for $\mathscr{W}_k(\Omega)$ and a fortiori for $\mathscr{V}_k$. We now state the main result of this section. We preserve the notation of the previous sections. The Hermitian form $[f, g]_k$ will have the obvious meaning.

THEOREM 5.1. *The n-widths of* $\mathscr{R}_k$ *satisfy*

$$d_n(\mathscr{R}_k) = \infty \qquad \text{if} \qquad n < N_0 \tag{5.1}$$

$$d_n(\mathscr{R}_k) = \lambda_{\nu+1}^{-1/2} \qquad \text{if} \qquad N_\nu \leqslant n < N_{\nu+1} \qquad \nu \geqslant 0, \tag{5.2}$$

where the λ_ν *are nonnegative numbers satisfying the functional equation*

$$[\varphi, f]_k = \lambda(\varphi, f)_{\mathscr{L}_2} \qquad \textit{for all} \qquad f \in \mathscr{V}_k \tag{5.3}$$

with corresponding finite dimensional eigenmanifolds $\mathscr{M}_\nu \subset \mathscr{V}_k$. $(0) \subset \mathscr{M}_0 \subset \mathscr{P}_k$, *where* $\mathscr{P}_k$ *is the set of polynomials of degree less than k. If* $N_\nu \leqslant n < N_{\nu+1}$ *then an optimal approximating n-subspace is* $\mathscr{M}_0 + \mathscr{M}_1 + \cdots + \mathscr{M}_\nu$. *Finally, the* $\mathscr{M}_\nu$ *are orthogonal and dense in both* $\mathscr{L}_2(\Omega)$ *and* $\mathscr{V}_k$ *and the solutions* φ *of the functional Eq. (5.3) are (real) analytic on* Ω *and satisfy*

$$(-1)^k \Delta^k = \lambda\varphi \tag{5.4}$$

where

$$\Delta = \frac{\partial^2}{\partial x_1{}^2} + \cdots + \frac{\partial^2}{\partial x_m{}^2}.$$

PROOF. Because of assumptions (i) and (ii), Theorem 3.3 is applicable and from that theorem follow (5.1), (5.2), and the fact that the λ_ν satisfy (5.3) and that $\mathscr{M}_0 + \cdots + \mathscr{M}_\nu$ is an extremal subspace. To see that $(0) \subset \mathscr{M}_0 \subset \mathscr{P}_k$ it suffices to show that if $D^\alpha f = 0$ for all $|\alpha| = k$ then f is a polynomial of degree less than k. This fact may be found in [10, p. 46]. To prove (5.4) we proceed as follows. We notice that the dispersive operator R constructed as in Section 3 satisfies the relation

$$[f, g]_k = (Rf, g)_{\mathscr{L}_2} \qquad \text{for all} \qquad f \in \mathscr{D}_R, \quad g \in \mathscr{V}_k \tag{5.5}$$

and in fact that $\mathscr{D}_R$ is characterized by (5.5). This implies that

$$C_c^\infty(\Omega) \subset \mathscr{D}_R,$$

$$R\psi = (-1)^k \Delta^k \psi \qquad \text{if} \qquad \psi \in C_c^\infty(\Omega).$$

Indeed, setting $u = (-1)^k \Delta^k \psi$ for $\psi \in C_c^\infty(\Omega)$ and choosing $g \in \mathscr{V}_k$ we have, by the definition of distribution derivative,

$$[\psi, g]_k = (u, g)_{\mathscr{L}_2},$$

which by (5.5) implies that $\psi \in \mathscr{D}_R$ and $R\psi = u$. Now if φ is an eigenfunction of R corresponding to eigenvalue λ and ψ is an arbitrary element of $C_c^\infty(\Omega)$ then

$$(\varphi, \lambda\psi)_{\mathscr{L}_2} = (\lambda\varphi, \psi)_{\mathscr{L}_2} = (R\varphi, \psi)_{\mathscr{L}_2} = (\varphi, (-1)^k \Delta^k \psi)_{\mathscr{L}_2}.$$

Thus, φ is a distribution solution of

$$[(-1)^k \Delta^k - \lambda I]\, \mu = 0$$

and, since $(-1)^k \Delta^k - \lambda I$ is elliptic, it follows that φ is (real) analytic on Ω and $(-1)^k \Delta^k \varphi = \lambda\varphi$ (see [11, Corollary (4.4.1), p. 114]). This concludes the proof of the theorem.

REMARK 5.1. It is proved in [12] that the class $C^\infty(\Omega)$ is dense in $\mathscr{W}_k(\Omega)$. Hence, the class $C^\infty(\Omega) \cap \mathscr{R}_k$ has the same $\mathscr{L}_2(\Omega)$ closure as $\mathscr{R}_k$ and Theorem 5.1 may be viewed as a result on the approximation of classes of smooth functions.

REMARK 5.2. The theory presented in Section 5 can be readily extended to the case of coercive Hermitian forms $B[f, g]$ over $\mathscr{V}_k$ of the form

$$B[f, g] = \sum_{\substack{|\alpha| \leqslant k \\ |\beta| \leqslant k}} \int_\Omega D^\alpha f \cdot \bar{a}_{\alpha\beta}\, D^\beta \bar{g}$$

$$B[f, f] \geqslant c_0 (f, f)_{\mathscr{W}_k} - \mu_0 (f, f)_{\mathscr{L}_2}, \qquad B[f, f] \geqslant 0,$$

where $a_{\alpha\beta} \in C^\infty(\bar{\Omega})$ and $c_0 > 0$, $\mu_0 \geqslant 0$. In this case the widths of the classes

$$\mathscr{S}_k = \{f \in \mathscr{V}_k : B[f, f] \leqslant 1\}$$

may be characterized, with the aid of assumption (ii), in terms of the eigenvalues of the functional equation

$$B[\varphi, g] = \lambda(\varphi, g)_{\mathscr{L}_2}. \tag{5.6}$$

The eigenfunctions φ of (5.6) are of class C^∞ in Ω and satisfy

$$\sum_{\substack{|\alpha| \leqslant k \\ |\beta| \leqslant k}} (-1)^{|\alpha|} D^\alpha [a_{\alpha\beta}\, D^\beta \varphi] = \lambda\varphi. \tag{5.7}$$

The reader is referred to [13, pp. 141, 142] for a more complete description of coercive forms. Also, one observes that $a_{\alpha\beta} \in C^k(\bar{\Omega}) \cap C^\infty(\Omega)$ is a sufficient hypothesis above.

6. *The Special Cases of $\mathring{\mathscr{H}}_k(\Omega)$ and $\mathscr{W}_k(\Omega)$.*

As remarked in Section 5, assumptions (i) and (ii) hold for $\mathring{\mathscr{H}}_k(\Omega)$ without any further assumptions on Ω. They follow from Poincare's inequality and a version of Rellich's lemma. In this case $\lambda = 0$ is not an eigenvalue of the functional equation (5.3) and $N_0 = 0$. If Ω satisfies a condition known as the *restricted cone condition* then the assumptions hold for the case $\mathscr{W}_k(\Omega)$. This condition does not depend for its statement upon the boundedness of Ω and we will state it for an open connected set. It must be understood, however, that the boundedness of Ω is crucial for the validity of assumptions (i) and (ii) (see [13, Theorems 3.3 and 3.8]).

DEFINITION 6.1. *An open connected set Ω in m-dimensional Euclidean space satisfies the restricted cone condition if the boundary $\partial\Omega$ of Ω has a locally finite open covering V_i and corresponding cones K_i with vertices at the origin and the property that $x + K_i \subset \Omega$ for all $x \in \Omega \cap V_i$. By the cones K_i we mean sets of the form*

$$K_i = \{x = t(y_i + z), z \cdot y_i = 0, z \cdot z \leqslant r_i^2, 0 \leqslant t \leqslant 1\}$$

where $y_i \neq 0$ and $r_i > 0$.

REMARK 6.1. A bounded region Ω satisfies the restricted cone condition if it is of class C^1. Also, if Ω satisfies the restricted cone condition then the infinitely differentiable functions $C^\infty(\bar{\Omega})$ with derivatives of all orders uniformly continuous in Ω are dense in $\mathscr{W}_k(\Omega)$ (see [13, Theorem 2.1, p. 11]).

Now in the case of $\mathscr{W}_k(\Omega)$, $N_0 = \dim \mathscr{P}_k$ and $\mathscr{M}_0 = \mathscr{P}_k$. We thus have

THEOREM 6.1. *If the bounded region Ω satisfies the restricted cone condition then assumptions* (i) *and* (ii) *of Section 5 are satisfied and Theorem 5.1 holds. $\mathscr{R}_k$ may be replaced by any class $\mathscr{S}_k$ satisfying $C^\infty(\bar{\Omega}) \cap \mathscr{R}_k \subset \mathscr{S}_k \subset \mathscr{R}_k$. $\mathscr{M}_0 = \mathscr{P}_k$ if $\mathscr{V}_k = \mathscr{W}_k(\Omega)$.*

7. *The Boundary Value Problem*

In this section we point out the equivalence, under smoothness assumptions on $\partial\Omega$, of the variational problems in the cases $\mathscr{V}_k = \mathring{\mathscr{H}}_k(\Omega)$ and $\mathscr{V}_k = \mathscr{W}_k(\Omega)$ with boundary value problems. Thus the approximation problem can be solved by solving a boundary value problem when $\partial\Omega$ is smooth.

The functional equation

$$[\varphi, f]_k = \lambda(\varphi, f)_{\mathscr{L}_2} \qquad \text{for all} \qquad f \in \mathring{\mathscr{H}}_k(\Omega)$$

is known to be equivalent to the Dirichlet problem over $\mathring{\mathscr{H}}_k(\Omega)$

$$(-1)^k \Delta^k \varphi = \lambda\varphi$$

$$D^\alpha \varphi = 0 \qquad \text{on} \qquad \partial\Omega \qquad \text{if} \qquad |\alpha| < k$$

if Ω is of class C^r for sufficiently large r. Certainly $r \geqslant 3k + [m/2]$ will suffice (see [14, Theorem 4, p. 359] and [15, p. 304]).

The functional equation

$$[\varphi, f]_k = \lambda(\varphi, f)_{\mathscr{L}_2} \qquad \text{for all} \qquad f \in \mathscr{W}_k(\Omega)$$

is known to be equivalent to the boundary value problem

$$(-1)^k \Delta^k \varphi = \lambda\varphi$$

$$N_{2k-1-j}\varphi = 0 \qquad \text{on} \qquad \partial\Omega \qquad \text{for} \qquad j = 0, 1, \ldots, k-1$$

if Ω is of class C^r for sufficiently large r where $N_{2k-1-j}(x)$, $x \in \partial\Omega$, is a differential operator of order $2k - 1 - j$ such that the surface $\partial\Omega$ is nowhere charcteristic for each N_{2k-1-j}. Certainly $r \geqslant 2k + [m/2] + 1$ will suffice (see [13, p. 143] and [14, Theorem 4, p. 359].

REMARK 7.1. N_{2k-1-j} contains the term $(\partial/\partial\nu)^{2k-1-j}$ with nonvanishing coefficient where $\partial/\partial\nu$ denotes differentiation along the normal to $\partial\Omega$. In the case $k = 1$ there is one natural boundary operator, $N_1 = (\partial/\partial\nu)$. In the case $k = 2$ and $m = 2$ Aronszajn [16, p. 376] has computed the operators N_2 and N_3 for a simple closed curve C in terms of normal $(\partial/\partial\nu)$ and tangential $(\partial/\partial s)$ derivatives:

$$N_2 = \frac{\partial^2}{\partial\nu^2}, \qquad N_3 = -\frac{\partial}{\partial\nu}\Delta + \frac{\partial^2}{\partial s^2\,\partial\nu} - \frac{\partial}{\partial s}\left(\frac{1}{\rho}\frac{\partial}{\partial s}\right),$$

where ρ is the radius of curvature on C.

REFERENCES

1. A. N. KOLMOGOROV. Über die beste Annähergung von Funktionen einer gegebenen Funktionklasse. *Ann. Math.* (2) **37** (1936), 107-111.
2. G. G. LORENTZ. Metric entropy, widths and superpositions of functions. *Amer. Math. Monthly* **69** (1962), 469-485.
3. V. M. TIHOMIROV. The widths of sets in functional spaces and the theory of best approximations. *Uspehi Mat. Nauk.* **15** (1960), No. 3 (93) 81-120, M. R. 22, 8268.
4. A. BROWN. Best n-dimensional approximation to sets of functions. *Proc. London Math. Soc.* (3) **14** (1964), 577-594.
5. N. P. KORNEICUK. The exact value of best approximations and widths of certain classes of functions. *Dokl. Akad. Nauk. SSR* **150** (1963), 1218-1220, M. R. 27, 3984.

6. G. G. LORENTZ. "Approximation of Functions." Holt, Rinehart and Winston, New York, 1966.
7. B. S. MITJAGIN. Approximation of functions in L^p-and C-spaces on the torus. *Mat. Sb.* **58** (1962), 397-414, M. R. 27, 2772.
8. M. Golomb. Optimal approximating manifolds in L_2-spaces. *J. Math. Anal. Appl.* **12** (1965), 505-512.
9. F. RIESZ AND B. SZ.-NAGY. "Functional Analysis." Ungar, New York, 1955.
10. S. L. SOBOLEV. "Applications of Functional Analysis in Mathematical Physics." Trans. Math. Mono., Vol. 7, Amer. Math. Soc., Providence, R. I., 1963.
11. L. HORMANDER. "Linear Partial Differential Operators." Academic Press, New York, 1963.
12. N. MEYERS AND J. SERRIN. H = W. *Proc. Natl. Acad. Sci. U.S.A.* **51** (1964), 1055-1056.
13. S. AGMON. "Lectures on Elliptic Boundary Value Problems." Van Nostrand, Princeton, N. J., 1965.
14. F. BROWDER. On the regularity properties of solutions of elliptic differential equations. *Comm. Pure Appl. Math.* **9** (1956), 351-361.
15. A. FRIEDMAN. "Partial Differential Equations of Parabolic Type." Prentice-Hall, Englewood Cliffs, N. J., 1964.
16. N. ARONSZAJN. Green's functions and reproducing kernels. *Proc. Symp. Spectral Theory Diff. Probl.*, Stillwater, Oklahoma 1951.
17. S. GOLDBERG. "Unbounded Linear Operators." McGraw-Hill, New York, 1966.