Computer methods in applied mechanics and engineering

Comput. Methods Appl. Mech. Engrg. 191 (2002) 5595–5629

www.elsevier.com/locate/cma

On principles for the selection of shape functions for the Generalized Finite Element Method ☆

Ivo Babuška [a], Uday Banerjee [b,*], John E. Osborn [c]

[a] *Texas Institute for Computational and Applied Mathematics, ACE 6.412, University of Texas at Austin, Austin, TX 78712, USA*
[b] *Department of Mathematics, 215 Carnegie, Syracuse University, Syracuse, NY 13244, USA*
[c] *Department of Mathematics, University of Maryland, College Park, MD 20742, USA*

Received 26 November 2001; received in revised form 11 June 2002

Abstract

Effective shape functions for the Generalized Finite Element Method should reflect the available information on the solution. This information is partially fuzzy, because the solution is, of course, unknown, and typically the only available information is the solution's inclusion in a variety of function spaces. It is desirable to choose shape functions that perform robustly over a family of relevant situations. Quantitative notions of robustness are introduced and discussed. We show, in particular, that in one dimension polynomials are robust when the available information consists in inclusions in Sobolev-type spaces that are x-independent.

Keywords: Generalized finite element methods; Approximations; Galerkin methods; Algebraic polynomials; Trigonometric polynomials; n-widths; Sup–inf; Error estimates

1. Introduction

The h, p, and hp versions of the finite element method employ local polynomial approximating or shape functions. These approximating functions, and the finite element methods based on them, are effective in many situations. In certain other situations, non-polynomial approximating functions have been used, and have been shown to be effective. We mention pullback polynomials (see e.g., [7]), the "quarterpoint" elements [10,15], and the enrichment of the finite element spaces by special functions. Recently, the generalized finite element method (GFEM), and its various meshfree methods versions, have been developed and analyzed. These methods also make flexible use of non-polynomial shape functions. They are often very effective (see e.g., [1,3,6,11,13,14,19–21,23,26,27]).

☆ This research was supported by NSF Grant # DMS-98-02367 and ONR Grant # N00014-99-1-0724.
* Corresponding author.
E-mail addresses: banerjee@syr.edu (U. Banerjee), jeo@math.umd.edu (J.E. Osborn).
URLs: http://bhaskara.syr.edu, http://www.wam.umd.edu/~jeo.

0045-7825/02/$ - see front matter © 2002 Elsevier Science B.V. All rights reserved.
PII: S0045-7825(02)00467-X

Polynomials have been studied for many years, starting in the 19th century, and they have been shown to have mostly good approximation properties. Nevertheless, they are not "good for all seasons". In [8], it was shown that for differential equations with rough coefficients, the finite element method using polynomial shape functions can lead to arbitrarily "bad" results.

Hence the following questions arise: What kinds of approximating functions should be used in various situations? In which situations are polynomials preferable? It is clear that the answer to these questions depends on the available information on the function to be approximated. This information can be of a priori or a posteriori type. In both cases the information has a *fuzzy* character, because the functions to be approximated is typically the unknown solution of a boundary value problem. In this paper we will address these questions in the one dimensional setting. The higher dimensional case will be addressed in a forthcoming paper.

To illustrate the issues we have raised, consider the problem

$$\begin{cases} -w''(x) = f(x), & x \in I = (-1,1), \\ w(-1) = w(1) = 0 \end{cases} \tag{1.1}$$

which has the variational formulation: Find $w \in H_0^1(I)$ satisfying

$$B(w,v) = F(v), \quad \text{for all } v \in H_0^1(I), \tag{1.2}$$

where

$$B(w,v) = \int_{-1}^{1} w'(x)v'(x)\,\mathrm{d}x, \tag{1.3}$$

$$F(v) = \int_{-1}^{1} f(x)v(x)\,\mathrm{d}x, \tag{1.4}$$

and

$$H_0^1(I) = \left\{ u : |u|_1 = \int_{-1}^{1} |u'(x)|^2\,\mathrm{d}x < \infty,\ u(\pm 1) = 0 \right\}. \tag{1.5}$$

We then approximate the solution of (1.2) by the Galerkin method. Toward this end we assume we have a set of basis functions, $\eta_1, \eta_2, \ldots$, in $H_0^1(I)$, and define the approximate solution by:

Find $w_n = \sum_{k=1}^{n} a_k \eta_k$ satisfying

$$B(w_n, \eta_j) = F(\eta_j), \quad \text{for } j = 1, \ldots, n. \tag{1.6}$$

w_n is called the Galerkin approximation to w determined by the basis functions $\{\eta_k\}$. We now consider, as examples, two specific sets of basis functions $\{\eta_k\}$:

(A) The sine functions,

$$\eta_k^T = \sin\frac{\pi}{2}k(x+1), \quad k = 1,2,\ldots \tag{1.7}$$

(B) The polynomial functions,

$$\eta_k^P = (1-x^2)x^{k-1}, \quad k = 1,2,\ldots \tag{1.8}$$

Each of these families is linearly independent and complete in $H_0^1(I)$. We note that the span $\{\eta_j^P,\ j = 1,2,\ldots,n\} = \text{span}\{\int_{-1}^{x} L_j(x)\,\mathrm{d}x,\ j = 1,2,\ldots,n\}$, where $L_j(x)$ is the Legendre polynomial of degree j. We then denote by w_n^T and w_n^P the approximate solutions using n trigonometric and n polynomial basis functions, respectively.

Is it better to use the trigonometric function or the polynomials? To begin to answer this question, let

$$e_n^T = w - w_n^T, \qquad e_n^P = w - w_n^P, \tag{1.9}$$

$$E_n^T = \frac{|e_n^T|_1}{|w|_1}, \qquad E_n^P = \frac{|e_n^P|_1}{|w|_1}, \tag{1.10}$$

be the errors and relative errors in the approximate solutions.

Let us examine E_n^T and E_n^P for the following two-parameter family of functions (exact solution of (1.2)):

$$w = w_{\beta,\gamma} = \begin{cases} (1+x)^\beta (x-0.5)^\gamma, & -1 \leqslant x \leqslant 0.5, \\ (1-x)^\beta (x-0.5)^\gamma, & 0.5 \leqslant x \leqslant 1, \end{cases} \tag{1.11}$$

where $\beta \geqslant 1$ and $\gamma \geqslant 1$ are integers. Note that the function $w_{\beta,\gamma}$ given in (1.11) is analytic (in fact, is a polynomial) in each of the intervals $-1 \leqslant x \leqslant 0.5$ and $0.5 \leqslant x \leqslant 1$. Also, $\mathrm{d}^{\gamma-1}w/\mathrm{d}x^{\gamma-1}$ is continuous on $[-1, 1]$ and $\mathrm{d}^\gamma w/\mathrm{d}x^\gamma$ has a jump discontinuity at the point 0.5.

Fig. 1 shows E_n^T and E_n^P for $\gamma = 5$ and for $\beta = 1$ and 7. For numerical values, see [2]. Fig. 1 shows the following features:

(a) The asymptotic rate of convergence of E_n^P is $\mathrm{O}(n^{-(\gamma-1/2)})$, which is independent of β.
(b) The pre-asymptotic performance of polynomial approximation shows dependence on β. Although the asymptotic rate for polynomial approximation is the same for $\beta = 1$ and $\beta = 7$, the pre-asymptotic behavior clearly influences the magnitude of the errors for large n.
(c) For E_n^T, the rate of convergence depends on both β and γ, with

$$E_n^T = \mathrm{O}(n^{-\min(\gamma-1/2,\beta+1/2)}), \quad \text{for } \beta \text{ odd}. \tag{1.12}$$

(d) Trigonometric polynomials perform marginally better than algebraic polynomials for $(\beta, \gamma) = (7, 5)$, but perform much worse for $(\beta, \gamma) = (1, 5)$.

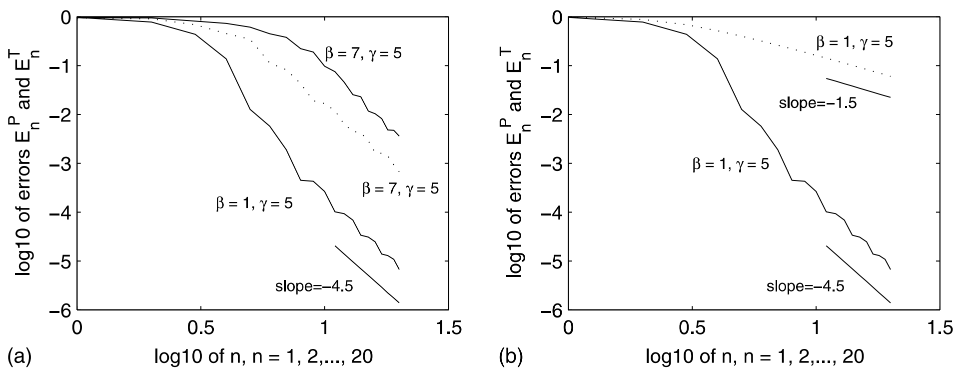

Fig. 1. The relative errors as a function of n when $w_{\beta,\gamma}$ is approximated by trigonometric and algebraic polynomials. E_n^P is represented by the solid line and E_n^T by the dotted line. Note the same convergence rate, namely $\mathrm{O}(n^{-4.5})$, for algebraic polynomial approximations (independent of β) in (a), and the different rates of convergence for algebraic and trigonometric polynomial approximations in (b). Also note in (a) that the error for trigonometric polynomial approximation is marginally smaller than that for algebraic polynomials in the case $\beta = 7$, irrespective of the same $\mathrm{O}(n^{-4.5})$ convergence.

The functions we wish to approximate may have a number of features. These include: smoothness, boundary behavior, boundary layers, oscillations, etc. Functions in higher dimension may have additional features. These features influence the approximation of these function in both the pre-asymptotic and asymptotic range. The family $w_{\beta,\gamma}$ has only two of these features; namely smoothness, parameterized by γ, and boundary behavior, parameterized by β. Nevertheless, we have chosen this two parameter family of functions because it illustrates, in a simple setting, the variety of features functions may have, and, in addition, the interplay of these features. In practice, many families of shape functions are used for approximation, especially in higher dimensions, in connection with the GFEM and with Meshfree Methods. Nevertheless, we consider only two families of shape functions—algebraic and trigonometric polynomials—because these families share many features with the larger class of shape functions that are used.

The available information on the approximated function, typically the unknown solution of a boundary value problem, is fuzzy. It is usually characterized by the function's inclusion in a family of function spaces. Effective shape functions should have good approximation properties in all the spaces of this family. In this paper, we develop quantitative concepts for assessing and comparing effectiveness of various families of shape functions. We focused on the principles that should govern the selection of shape functions, and elaborated in detail the one dimensional case. These results are arrived through a collection of theoretical results. These include a thorough study of the properties of the quantitative notions mentioned above. We have also presented detailed numerical computations, which illustrate these theoretical results.

Section 2 briefly describes the GFEM, introduced to further motivate our study. In Section 3 we describe a wide family of function spaces, and discuss their role in understanding approximation. We also discuss the notion of "smoothness", and explain some of the observations made about the example in the present section. Section 4 introduces the basic concepts in term of which we assess the effectiveness of specific families of approximating functions. In Section 5 we present the main results of the paper. In Section 6, we present computational results and their interpretation. These computations illustrate the results presented in Section 5. Finally, in Section 7 we summarize the conclusions of this paper.

2. The Generalized Finite Element Method

In this section we briefly describe the GFEM. Let $\Omega \subset \mathbf{R}^d$ be a bounded domain with boundary Γ, and denote by $H^1(\Omega)$ the usual Sobolev space of functions with square integrable first derivatives. Let V_i, $i = 1, \dots, N$, be a collection of Lipschitz functions with supports σ_i, respectively, (we do not assume $\sigma_i \subset \Omega$), and suppose that

(a) $\sum_{i=1}^{N} V_i(x) = 1$ on Ω,
(b) $|V_i(x)| \leqslant C_\infty$, $i = 1, \dots, N$,
(c) $|\mathrm{grad} V_i(x)| \leqslant C_\mathrm{G}/\mathrm{diam}(\sigma_i)$, $i = 1, \dots, N$,
(d) there is an M such that any $x \in \Omega$ lies in at most M of the supports σ_i.

The supports σ_i are called patches. We then assume we are given functions $\eta_{i,j}$, $j = 1, \dots, m(i)$, $i = 1, \dots, N$, with $\eta_{i,j} \in H^1(\sigma_i)$ and $\eta_{i,1} = 1$. If $\sigma_{i,\mathrm{D}} = \sigma_i \cap \Gamma_\mathrm{D}$, where $\Gamma_\mathrm{D} \subset \Gamma$ is the portion of the boundary where (zero) Dirichlet boundary conditions are prescribed, and the measure of $\sigma_{i,\mathrm{D}} \neq 0$, then we assume the $\eta_{i,j} = 0$ on $\sigma_{i,\mathrm{D}}$ (and hence the condition $\eta_{i,1} = 1$ is not imposed).

For fixed i, let S_i be the span of $\eta_{i,j}$, $j = 1, 2, \dots$ Clearly

$$V_i \eta_{i,j}|_\Omega \in H^1(\Omega), \quad i = 1, \dots, N, \;\; j = 1, \dots, m(i), \tag{2.1}$$

where $V_i\eta_{i,j}|_\Omega$ denotes the restriction of $V_i\eta_{i,j}$ to Ω. We denote by S the span of all the $V_i\eta_{i,j}$, $j = 1, \dots, m(i)$, $i = 1, \dots, N$. For $w \in H^1(\Omega)$, it can be shown [3,6,23] that

$$\inf_{\chi\in S} |w - \chi|^2_{H^1(\Omega)} \leqslant C \sum_{i=1}^{N} \inf_{\rho\in S_i} |w - \rho|^2_{H^1(\sigma_i\cap\Omega)}, \tag{2.2}$$

where the constant C in (2.2) depends only on C_∞, C_G, and M.

The GFEM is the Galerkin method using S as approximating functions. Estimate (2.2), together with the usual quasi-optimal estimate for Galerkin methods, shows that the error in the GFEM is directly related to the approximation properties of the systems $\{\eta_{i,j}\}_{j=1}^{m(i)}$ on the patches σ_i. A special form of GFEM was introduced in [3] to address problems with rough coefficients. This method was extended in [6,23], and called the partition of unity finite element method (PUFEM). For further development of GFEM (see [26,27]). The GFEM is closely related to meshfree methods [1,11–14,19–22].

Let us now specialize the GFEM to one dimensional problems on $I = (-1, 1)$. Consider, for example, the specific problem

$$\begin{cases} -w''(x) = f(x), & \text{for } x \in I = (-1, 1), \\ w'(-1) = w'(1) = 0, \\ \int_{-1}^{1} w\,\mathrm{d}x = 0, \end{cases} \tag{2.3}$$

where f is a given function satisfying $\int_{-1}^{1} f(x)\,\mathrm{d}x = 0$. Then, if we wish to use the GFEM to approximate the solution w, we need to construct a system $\eta_{i,j}$, $i = 1, \dots, N$, $j = 1, \dots, m(i)$ that closely approximates the solution w on the patch σ_i in the H^1-seminorm. This is equivalent to the construction of a system $\phi_{i,j} = \eta'_{i,j}$ that closely approximates $w'(x)$ on σ_i in the L_2-norm. The situation is similar for other problems. For example, for the problem

$$\begin{cases} -w''(x) = f(x), & \text{for } x \in I = (-1, 1), \\ w(-1) = w(1) = 0 & \text{or } w' = w(1) = 0, \end{cases} \tag{2.4}$$

we are again interested in L_2-approximation of w' because we can easily construct the approximation in H^1 that satisfies the essential (Dirichlet) boundary conditions.

Hence we are led to the problem of selecting a system of basis functions, $\phi_{i,j} = \eta'_{i,j}$, on σ_i that closely approximate the derivative $u = w'$ of the unknown solution w in L_2. We note that generally it is not possible to view approximations in this way. In fact, for second order boundary value problem in higher dimension, approximations are naturally viewed in H^1. Nevertheless, in this paper we restrict our attention to approximation in L_2. Without loss of generality we can assume $\sigma_i = \sigma = (-1, 1)$. We are thus seeking approximating functions ϕ_i that closely approximate u on $\sigma = (-1, 1)$ in L_2. These functions are called basis or shape functions.

3. Function spaces and their role in approximation

In this section we define the Sobolev-type spaces we will use in our analysis. As pointed out in the Section 1, the assessment of the approximation properties of a specific family of basis functions depends on the available information on the function we are approximating. We typically know that the function u of interest is included in some function space. But it is also typical that the function is included in many spaces, and this information provides a more complete understanding of the function. It is thus important to have for our use a broad family of appropriate function spaces.

For $k = 0, 1, \ldots$ and $I = (-1, 1)$, $H^k(I)$ denotes the kth order Sobolev space, with the norm

$$\|u\|_k^2 = \sum_{j=0}^{k} |u|_j^2, \tag{3.1}$$

where

$$|u|_j^2 = \int_{-1}^{1} |u^{(j)}(x)|^2 \mathrm{d}x. \tag{3.2}$$

Let $\alpha = (\alpha_0, \alpha_1, \ldots)$ be a sequence of non-negative numbers, with $\alpha_0 = 1$, and let k be a non-negative integer or $+\infty$. We then define the Sobolev-type space $H^{\alpha,k}(I)$ by

$$H^{\alpha,k}(I) = \left\{ u \in L_2(I) : \|u\|_{H^{\alpha,k}}^2 \equiv \sum_{j=0}^{k} \alpha_j |u|_j^2 < \infty \right\}. \tag{3.3}$$

With the associated inner product, $H^{\alpha,k}(I)$ is a Hilbert space. We will often write $H^{\alpha,k}$ for $H^{\alpha,k}(I)$. We note that when $k = \infty$,

$$\|u\|_{H^{\alpha,\infty}}^2 = \sum_{j=0}^{\infty} \alpha_j |u|_j^2 = \lim_{k \to \infty} \|u\|_{H^{\alpha,k}}^2. \tag{3.4}$$

Remark 3.1. The sequence α introduced here does not depend on k. It will be useful, however, to also consider k-dependent α, by which we mean a family of finite sequences: $\alpha(k) = (\alpha_0(k), \alpha_1(k), \ldots, \alpha_k(k))$, $0 \leqslant k < \infty$. The norm on the associated Sobolev space is the natural extension of the norm defined in (3.3): $\|u\|_{H^{\alpha(k),k}}^2 = \sum_{j=0}^{k} \alpha_j(k) |u|_j^2$. We will speak of k-independent α and k-dependent $\alpha(k)$. For any k-independent $\alpha = (\alpha_0, \alpha_1, \ldots)$, we can define an associated k-dependent $\alpha = \alpha(k)$ by letting $\alpha_i(k) = \alpha_i$ for $0 \leqslant i \leqslant k$ and $\alpha_i(k) = 0$ for $i > k$. With this convention, $H^{\alpha,k} = H^{\alpha(k),k}$, and k-independent α can be also be considered as k-dependent.

Remark 3.2. For a k-independent α, it is easily seen that

$$H^{\alpha,k+1} \subset H^{\alpha,k} \quad \text{and} \quad \|u\|_{H^{\alpha,k}} \leqslant \|u\|_{H^{\alpha,k+1}}, \quad \text{for all } u \in H^{\alpha,k+1}.$$

This result does not hold in general for k-dependent $\alpha(k)$, but there are $\alpha = \alpha(k)$ for which the result does hold, i.e., $\alpha(k)$ for which

$$H^{\alpha(k+1),k+1} \subset H^{\alpha(k),k} \tag{3.5}$$

and

$$\|u\|_{H^{\alpha(k),k}} \leqslant \|u\|_{H^{\alpha(k+1),k+1}}, \quad \text{for all } u \in H^{\alpha(k+1),k+1}, \tag{3.6}$$

or at least (3.5), hold.

We consider the following choices for α:

$$\alpha_j = 1, \tag{3.7a}$$

$$\alpha_j = \binom{k}{j}, \quad \text{for } j \leqslant k, \text{ where } k < \infty, \tag{3.7b}$$

$$\alpha_j = \begin{cases} 1 & \text{if } j = 0, \\ 0 & \text{if } 1 \leqslant j \leqslant k-1, \quad \text{where } k < \infty, \\ 1 & \text{if } j = k, \end{cases} \tag{3.7c}$$

$$\alpha_j = j!, \tag{3.7d}$$

$$\alpha_j = 1/j!, \tag{3.7e}$$

$$\alpha_j = 10^{2j}, \tag{3.7f}$$

$$\alpha_j = 10^{-2j}. \tag{3.7g}$$

Remark 3.3. Choices (3.7a), (3.7d)–(3.7g) are k-independent, whereas choices (3.7b) and (3.7c) are k-dependent. (3.5) holds for both of these k-dependent choices. (3.6) holds for choice (3.7b), but not for choice (3.7c).

The inclusion $u \in H^{\alpha,k}$ provides information on u and its derivatives that depends on α. We now examine this information for the choices for α given above. With α as in (3.7a) and with $k < \infty$, we have the usual Sobolev space $H^k : H^{\alpha,k} = H^k$ and $\|u\|_{H^{\alpha,k}} = \|u\|_k$; if $k = 0$, we have L_2 and the usual L_2-norm. The norm $\|u\|_k$ gives equal weight to all of the seminorms $|u|_j$, $j = 0, 1, \ldots, k$. Suppose we know that $\|u\|_k \leqslant 1$. Then we know that $|u|_j \leqslant 1$, $j = 0, 1, \ldots, k$; but we have very little additional information on the $|u|_j$.

The choice of α in (3.7b) arises if we define the Sobolev norm on $(-\infty, \infty)$ via the Fourier transform. This choice gives higher weight to derivatives of orders approximately midway between 0 and k. From the information that $\|u\|_{H^{\alpha,k}} \leqslant 1$, we would know that

$$|u|_j \leqslant \binom{k}{j}^{-1/2}, \tag{3.8}$$

which is more restrictive for j approximately midway between 0 and k, and less restrictive for j near 0 or k.

The choice of α in (3.7c) gives equal weight to u and $u^{(k)}$, but 0 weight to $u^{(j)}$, for $j = 1, \ldots, k-1$. From the information $\|u\|_{H^{\alpha,k}} \leqslant 1$, we know that $|u|_0 \leqslant 1$ and $|u|_k \leqslant 1$; but the $|u|_j$, $1 \leqslant j \leqslant k-1$, are nearly unconstrained. There is, however, some information on $|u|_j$ because there is a constant C_k such that

$$|u|_j \leqslant C_k(|u|_0^2 + |u|_k^2)^{1/2}, \quad 1 \leqslant j \leqslant k, \tag{3.9}$$

but C_k is very large.

For choices of α in (3.7d) and (3.7f), $\|u\|_{H^{\alpha,k}}$ weights the higher derivatives much more than the lower derivatives. From the information $\|u\|_{H^{\alpha,k}} \leqslant 1$, we know that

$$|u|_j \leqslant (j!)^{1/2}, \quad \text{for choice (3.7d)}; \quad |u|_j \leqslant 10^{-j}, \quad \text{for choice (3.7f)}.$$

For choices of α in (3.7e) and (3.7g), $\|u\|_{H^{\alpha,k}}$ weights the lower derivatives much more than the higher derivatives. From the information $\|u\|_{H^{\alpha,k}} \leqslant 1$, we have

$$|u|_j \leqslant (j!)^{1/2}, \quad \text{for choice (3.7e)}; \quad |u|_j \leqslant 10^{j}, \quad \text{for choice (3.7g)}.$$

Remark 3.4. The reason for the choices (3.7a)–(3.7g) for α will become clear in Section 6. We note that $j! > 10^{2j}$ for large j, i.e., α_j in (3.7d) is greater than α_j in (3.7f) for large j. This fact will be relevant in explaining certain computational results in Section 6.

For all the k-independent choices (3.7a), (3.7d)–(3.7g), functions $u \in H^{\alpha,\infty}$ are entire functions, i.e., they have analytic extensions to the entire complex plane.

In addition to $H^{\alpha,k}$ we will use the following spaces:

$$H^{\alpha,k}_{\text{even}} \equiv \{u \in H^{\alpha,k} : u^{(j)}(-1) = u^{(j)}(1) = 0,\ \ 0 \leqslant j \leqslant k-1,\ \ j \text{ even}\} \tag{3.10}$$

and

$$H^{\alpha,k}_{\text{odd}} \equiv \{u \in H^{\alpha,k} : u^{(j)}(-1) = u^{(j)}(1) = 0,\ \ 0 \leqslant j \leqslant k-1,\ \ j \text{ odd}\}, \tag{3.11}$$

where we assume $k < \infty$ and $\alpha_k \neq 0$. On these spaces we use the $H^{\alpha,k}$-norm:

$$\|u\|_{H^{\alpha,k}_{\text{even}}} = \|u\|_{H^{\alpha,k}} \quad \text{and} \quad \|u\|_{H^{\alpha,k}_{\text{odd}}} = \|u\|_{H^{\alpha,k}}.$$

With α as in choice (3.7a) above ($\alpha_j = 1$), we often write k for the superscript α, k on these spaces. The spaces $H^{\alpha,k}_{\text{even}}$ and $H^{\alpha,k}_{\text{odd}}$ impose constraints at the end points of the interval I.

We introduce still additional spaces. Let

$$V^k(I) = \left\{ u \in L_2(I) : \|u\|^2_{V^k} = \sum_{j=0}^{k} |u|^2_{V^j_j} < \infty \right\}, \tag{3.12}$$

where

$$|u|^2_{V^j_j} = \int_{-1}^{1} (1-x^2)^j |u^{(j)}|^2 \, \mathrm{d}x. \tag{3.13}$$

The seminorms $|u|_{V^j_j}$ are called Jacobi seminorms. If, as above, $\alpha = (\alpha_0, \alpha_1, \ldots)$ is a sequence of non-negative numbers, with $\alpha_0 = 1$, and $0 \leqslant k \leqslant \infty$, we define

$$V^{\alpha,k}(I) = \left\{ u \in L_2(I) : \|u\|^2_{V^{\alpha,k}} = \sum_{j=0}^{k} \alpha_j |u|^2_{V^j_j} < \infty \right\}. \tag{3.14}$$

With the associated inner product, $V^{\alpha,k}(I)$ is a Hilbert space. We will often write $V^{\alpha,k}$ for $V^{\alpha,k}(I)$. With α as in choice (3.7a) above ($\alpha_j = 1$) and with $k < \infty$, we have $\|u\|_{V^{\alpha,k}} = \|u\|_{V^k}$ and $V^{\alpha,k}(I) = V^k(I)$.

We see that $H^{\alpha,k} \subset V^{\alpha,k}$. The major difference between the spaces $H^{\alpha,k}$ and $V^{\alpha,k}$ is that derivatives of a function u are suppressed in $\|u\|_{V^{\alpha,k}}$ by the multiplicative weight function $(1-x^2)^j$, and thus functions $u \in V^{\alpha,k}$ are permitted to have singular behavior at the endpoints ± 1 of the interval I. We note that the smoothness of the functions in $H^{\alpha,k}$ is characterized uniformly in x, i.e., the weight function in $|u|_j$ is 1 (cf. (3.2)). On the other hand, the smoothness of the functions in $V^{\alpha,k}$ is *not* characterized uniformly in x because of the presence of the weight function $(1-x^2)^j$ in $|u|_{V^j_j}$. In this paper, we will not consider spaces with norms incorporating other x-dependent weight factors.

In Section 1 we considered the Galerkin approximation w^T_n of the exact solution, w, of (1.1), by linear combinations of the functions $\eta^T_j(x) = \sin(\pi/2)j(x+1)$ introduced in (1.7), and measured the error in H^1_0 (cf. (1.9) and (1.10)). It can be easily seen that $[w^T_n]'$ is the best L_2-approximation of w' in span $\{\cos(\pi/2)(x+1), \ldots, \cos(\pi/2)n(x+1)\}$. Similarly, one can see that $[w^P_n]'$ is the best L_2-approximation to w' in span $\{x, \ldots, x^n\}$.

The quality of approximation by a specific basis depends on the properties of the approximated function, and these properties are partially known through the information we have on the function. Typically, this information consists of function space inclusions, which may include boundary conditions. Let us examine this information for the functions $w_{7,5}(x)$ and $w_{1,5}(x)$ introduced in Section 1, and show how it can be used, in conjunction with results from Section 5.

We easily see that $w'_{1,5} \in H^4(-1,1)$ and $w'_{7,5} \in H^4(-1,1)$, and hence both $w'_{1,5}$ and $w'_{7,5}$ belong to $V^4(-1,1)$. $w'_{1,5}$ and $w'_{7,5}$ are also in $H^{\alpha,4}$ and $V^{\alpha,4}$ for other values of α, but, for the sake of simplicity, we consider here only α as in (3.7a). We also see that $w'_{1,5} \in H^1_{\text{odd}}$ and $w'_{7,5} \in H^4_{\text{odd}}$.

Consider first the approximation of $w'_{1,5}$ and $w'_{7,5}$ by cosines. Using (5.48) in Section 5 with $k = 1$ and 4, together with the information that $w'_{1,5} \in H^1_{\text{odd}}$ and $w'_{7,5} \in H^4_{\text{odd}}$, we have

$$\frac{\|w' - [w_n^T]'\|_{L_2}}{\|w'\|_{L_2}} \leqslant \begin{cases} Cn^{-1}\|w'\|_{H^1_{\text{odd}}}/\|w'\|_{L_2}, & \text{if } w = w_{1,5}(x), \\ Cn^{-4}\|w'\|_{H^4_{\text{odd}}}/\|w'\|_{L_2}, & \text{if } w = w_{7,5}(x). \end{cases} \tag{3.15}$$

Next consider the approximation of $w'_{1,5}$ and $w'_{7,5}$ by algebraic polynomials. Using (5.47) in Section 5 with $k = 4$, together with the information that $w'_{1,5}, w'_{7,5} \in V^4(-1,1)$, we have

$$\frac{\|w' - [w_n^P]'\|_{L_2}}{\|w'\|_{L_2}} \leqslant Cn^{-4}\frac{\|w'\|_{V^4}}{\|w'\|_{L_2}}, \quad \text{for } w = w_{1,5} \text{ or } w_{7,5}. \tag{3.16}$$

Remark 3.5. For the sake of simplicity, we are using Sobolev-type spaces with integer orders. For sharper error estimates, we have to use Besov spaces with fractional order. Using these spaces, it is possible to establish the rates of convergence, $O(n^{-(\gamma-1/2)})$, for algebraic polynomial approximation, and $O(n^{-\min(\gamma-1/2,\beta+1/2)})$, for trigonometric approximation. These orders were seen in Fig. 1 and were stated in items (a) and (c) near the end of Section 1.

Smoothness of a function involves more than just function space inclusions; it also involves the sizes of the norms of the function. We have noted that both $w'_{1,5}$ and $w'_{7,5}$ belong to $H^4(-1,1)$, and hence to $V^4(-1,1)$. But from Fig. 1 we see that E_n^P for $w_{7,5}$ is much larger than the E_n^P for $w_{1,5}$ for the same value of n. To understand this difference, we report $\|w'\|_{H^k}/\|w'\|_0$ and $\|w'\|_{V^k}/\|w'\|_0$, for $k = 0, \ldots, 4$, in Table 1, with $w = w_{1,5}$ and $w_{7,5}$.

We see that the V^k-norms (normalized) of $w'_{7,5}$ are larger than the corresponding V^k-norms of $w'_{1,5}$ for $k \geqslant 1$; specifically, $\|w'_{1,5}\|_{V^k}/\|w'\|_0 < \|w'_{7,5}\|_{V^k}/\|w'\|_0$. Therefore we say that $w'_{1,5}$ is "smoother" than $w'_{7,5}$, even though they belong to the same space $V^4(-1,1)$. Hence, $w'_{1,5}$ can be better approximated by algebraic polynomials than can $w'_{7,5}$. For a further qualitative understanding of this issue, see [2].

We also note that in classical finite element analysis, function inclusions are usually considered in H^k, and the H^k-norm does not allow singularities at the endpoints of the interval of definition. Therefore, the error estimate for polynomial approximation of a function with end-point singularities, considered in H^k, can be large, as compared to an error estimate with the function considered in V^k. In fact, we see from Table 1 that the ratio of $\|w'\|_{H^k}/\|w'\|_{L_2}$ to $\|w'\|_{V^k}/\|w'\|_{L_2}$ for $w = w'_{1,5}$, which has an end point singularity, is greater than

Table 1
The seminorms $\|w'\|_{H^k}/\|w'\|_0$ and $\|w'\|_{V^k}/\|w'\|_0$ of the function defined in (1.11) for $\beta = 1$, 7 and $\gamma = 5$

k	$\beta = 1, \gamma = 5$		$\beta = 7, \gamma = 5$	
	$\|w'\|_{H^k}/\|w'\|_0$	$\|w'\|_{V^k}/\|w'\|_0$	$\|w'\|_{H^k}/\|w'\|_0$	$\|w'\|_{V^k}/\|w'\|_0$
0	1.0000	1.0000	1.0000	1.0000
1	7.9912	3.4847	6.0699	5.8676
2	39.237	12.172	46.027	43.048
3	132.13	41.266	421.78	375.68
4	313.39	126.21	5372.7	4113.5

Note that $\|w'\|_{V^k}/\|w'\|_0$ for $(\beta,\gamma) = (1,5)$ is much smaller than for $(\beta,\gamma) = (7,5)$. Also the ratio of $\|w'\|_{H^k}/\|w'\|_0$ to $\|w'\|_{V^k}/\|w'\|_0$ for $(\beta,\gamma) = (1,5)$ is larger than for $(\beta,\gamma) = (7,5)$.

for $w = w'_{7,5}$, which does not. This suggests that it is better to consider a function in V^k when assessing approximation by algebraic polynomials.

We end this section with a comment that (5.48), which was used to establish (3.15), can also be understood in terms of periodic extensions and Fourier series. This idea has been discussed in [2].

4. Assessment of the effectiveness of approximating functions

In this section we define the concepts in terms of which we will assess the effectivity of specific families of basis functions.

Suppose the function we wish to approximate is included in a Hilbert space V, with norm $\|\cdot\|_V$, and we wish to measure the approximate error in the norm $\|\cdot\|_H$, where H is a Hilbert space that contains V. We will consider sequences $\Phi = (\phi_1, \phi_2, \ldots)$ of approximating shape functions in H, and assume Φ is a basis, i.e., is linearly independent ($\phi_1, \ldots, \phi_n$ is linearly independent for each n) and complete (ϕ is linearly dense in H).

We let

$$S(\Phi, n) \equiv \operatorname{span}(\phi_1, \ldots, \phi_{\mathrm{n}}), \tag{4.1}$$

and define

$$\Psi(\Phi, n, V, H) = \sup_{\substack{u \in V \\ \|u\|V \leqslant 1}} \inf_{\phi \in S(\Phi,n)} \|u - \phi\|_H. \tag{4.2}$$

We will refer to $\Psi(\Phi, n, V, H)$ as the sup–inf for the basis Φ. It is immediate that

$$\inf_{\phi \in S(\Phi,n)} \|u - \phi\|_H \leqslant \Psi(\Phi, n, V, H)\|u\|_V, \quad \text{for all } u \in V. \tag{4.3}$$

This is the fundamental estimate for the best H-approximation of an arbitrary $u \in V$ by $\Phi \in S(\Phi, n)$. Note that $\Psi(\Phi, n, V, H)$ measures the approximability of the function in the V-unit ball that is most difficult to approximate.

Given two families Φ^1 and Φ^2 of basis functions, we consider the ratio

$$\kappa(\Phi^1, \Phi^2, n, V, H) = \frac{\Psi(\Phi^1, n, V, H)}{\Psi(\Phi^2, n, V, H)} \tag{4.4}$$

in terms of which we can assess the effectiveness of Φ^1 as compared to Φ^2, or vise versa. Roughly speaking, if κ is near 1, we have no reason to prefer one of Φ^1 or Φ^2 over the other. On the other hand, if κ is much smaller than 1, then we prefer Φ^1; while if κ is much larger than 1, then we prefer Φ^2.

The above criteria compare the bases Φ^1 and Φ^2. It is also informative to compare $\Phi = \Phi^1$ with all other bases Φ^2, and thereby obtain a more absolute assessment the approximability properties of Φ. This can be done with notion of n-width, introduced by Kolmogorov [18]. Let $d_n(V, H)$ denote the n-width of the V-unit ball in H:

$$d_n(V, H) = \inf_{\substack{X_n \subset H \\ \dim X_n = n}} \sup_{\substack{u \in V \\ \|u\|_V \leqslant 1}} \inf_{\phi \in X_n} \|u - \phi\|_H. \tag{4.5}$$

We can also express $d_n(V, H)$ as follows:

$$d_n(V, H) = \inf_{\substack{X_n \subset H \\ \dim X_n = n}} \Psi(X_n, n, V, H) = \inf_{\substack{\phi_1, \ldots, \phi_n \in H \\ \text{linearly independent}}} \Psi(\Phi, n, V, H). \tag{4.6}$$

An n-dimensional subspace $\widetilde{X}_n = \widetilde{X}_n(V,H)$ of H is said to be optimal if

$$d_n(V,H) = \sup_{\substack{u\in V\\ \|u\|_V \leqslant 1}} \inf_{\phi\in\widetilde{X}_n} \|u-\phi\|_H = \Psi(\widetilde{X}_n, n, V, H). \tag{4.7}$$

Given a specific basis Φ, we will assess its effectiveness by considering the ratio

$$\Lambda(\Phi,n,V,H) \equiv \frac{\Psi(\Phi,n,V,H)}{d_n(V,H)} = \frac{\Psi(\Phi,n,V,H)}{\Psi(\widetilde{X}_n,n,V,H)}. \tag{4.8}$$

We see that the ratio Λ compares Φ with the optimal subspace, and therefore $\Lambda \geqslant 1$. If Λ is bounded in n, then, as above, Φ and $\widetilde{X}_n$ have similar approximation properties. In this situation, we would use Φ since $\widetilde{X}_n$ is usually not explicitly known. But if $\Lambda(\Phi,n,V,H) \to \infty$ as $n \to \infty$, the choice of Φ will not yield efficient approximations, as compared with $\widetilde{X}_n$.

We now discuss a situation we will return to in Section 5. Let $\mathscr{V} = \{V\}$ be a family of spaces $V \subset H$. It follows immediately from the definition of n-widths that $\Lambda(\Phi,n,V,H) \geqslant 1$ for any $V \subset \mathscr{V}$, and it is likely that $\Lambda(\Phi,n,V,H) > 1$ for any particular V because the functions $\phi_i \in \Phi$ are not likely to be optimal in V. Suppose, however, that $\Lambda(\Phi,n,V,H)$ is almost bounded in the sense that for each $0 < \sigma \leqslant 1$, there is a constant $C(\sigma)$ such that

$$\Lambda_\sigma(\Phi,n,V,H) \equiv \frac{\Psi(\Phi,n,V,H)}{d_n(V,H)^{1-\sigma}} \leqslant C(\sigma) \quad \text{for all } n = 1,2,\ldots \text{ and for all } V \in \mathscr{V}. \tag{4.9}$$

In this situation we expect the $\phi_i \in \Phi$ to be effective for approximating any $u \in V$ for any $V \in \mathscr{V}$. We say that the basis Φ is *almost uniformly optimal* with respect to $\mathscr{V}$.

In the rest of this paper, we will use $H = L_2(-1,1)$, and will suppress H in $d_n(V,H)$ and $\Psi(\Phi,n,V,H)$, i.e., we will write $d_n(V)$ and $\Psi(\Phi,n,V)$ instead of $d_n(V,H)$ and $\Psi(\Phi,n,V,H)$, respectively. Similarly, we will write $\Lambda(\Phi,n,V)$ and $\Lambda_\sigma(\Phi,n,V)$.

For the space V we will use $H^{\alpha,k}(-1,1)$, $V^{\alpha,k}(-1,1)$, $H^{\alpha,k}_{\text{odd}}(-1,1)$, or $H^{\alpha,k}_{\text{even}}(-1,1)$, i.e., one of the spaces defined in Section 3. We will consider several specific bases Φ:

$$\mathscr{P} = \{1, x, x^2, \ldots\},$$

$$\mathscr{C} = \left\{1, \cos\frac{\pi}{2}(x+1), \cos\pi(x+1), \ldots\right\},$$

$$\mathscr{S} = \left\{\sin\frac{\pi}{2}(x+1), \sin\pi(x+1), \ldots\right\}.$$

We note that $S(\mathscr{P}, n+1) = \mathscr{P}_n$, the space of polynomials of degree n.

Typically, the effectiveness of finite element methods is assessed by comparing their performance on a small set of benchmark problems. This could lead to misleading conclusions, unless these benchmark problems represent a more-or-less clearly defined class of problems. The families of spaces we have considered, i.e., $\mathscr{V} = \{V\}$, where $V = H^{\alpha,k}$, $V^{\alpha,k}$, $H^{\alpha,k}_{\text{odd}}$, or $H^{\alpha,k}_{\text{even}}$, represent well-defined large classes of solutions of boundary value problems. The various notions—κ, Λ, Λ_σ, etc.—introduced in this section will be used to understand and compare the performance of various bases Φ for the family $\mathscr{V}$. We finally note that the concepts defined in this section are not only relevant for one dimensional problems; they will also be used to understand higher dimensional problems in a forthcoming paper.

5. Results on n-widths and sup–infs

In Section 4 we introduced the n-width, $d_n(V,H)$, and the sup–inf, $\Psi(\Phi,n,V,H)$; these quantities are of fundamental importance in assessing the approximation properties of bases Φ in H. The motivation for our study is the development of criteria for the selection of effective shape functions for use in the GFEM in higher dimensions. Nevertheless, for the one dimensional problems we are elaborating on, as pointed out in the last paragraph of Section 2, we can view the approximation in L_2—using the notation of that paragraph, we approximate $u = w'$ in L_2. Thus, throughout this section, we will let $H = L_2(-1, 1)$.

In this section we will present several results on d_n and Ψ. They include: (a) characterization of d_n and Ψ in terms of eigenvalues of certain variational eigenvalue problems; (b) dependence of $d_n(V)$ on n and k for $V = H^{\alpha,k}_{\mathrm{odd}}(-1, 1)$, $H^{\alpha,k}(-1, 1)$, or $V^{\alpha,k}(-1, 1)$; and (c) dependence of $\Psi(\Phi, n, V)$ on n and k for $V = H^{\alpha,k}(-1, 1)$ or $V^{\alpha,k}(-1, 1)$ and for $\Phi = \mathscr{P}$ or $\mathscr{C}$. Other main results of this section are Theorems 5.23, 5.24, and 5.25. Theorem 5.23 shows that algebraic polynomials are effective in approximating functions in a wide family of Sobolev-type spaces. Theorems 5.24 and 5.25 show the limitations of algebraic polynomials in approximating functions in certain other Sobolev-type spaces with restrictive boundary conditions. These results explain the computational results presented in Section 6. We suggest this section be read together with Section 6. For the proofs of the results in this section (see [2]).

Some of the theorems presented in this section are classical (e.g., Theorems 5.1 and 5.2), and some (e.g., Theorems 5.6, 5.7, 5.15–5.18) are observations that follow directly from relevant definitions or other results in this paper. Nevertheless, we have highlighted them as theorems for completeness, and for the ease of discussion of certain other results presented in this section.

We begin with the properties of $d_n(V, H)$; we will later discuss properties of $\Psi(\Phi, n, V, H)$.

5.1. Characterization of $d_n(V, H)$

Suppose H and V are two Hilbert spaces with inner products and norms $(u, v)_V$ and $\|u\|_V$, and $(u, v)_H$ and $\|u\|_H$, respectively, and assume

$$V \subset H, \text{ compactly,} \tag{5.1}$$

$$\|u\|_H \leqslant C\|u\|_V, \quad \text{for all } u \in V. \tag{5.2}$$

Then the n-width $d_n(V, H)$ and the corresponding optimal subspace $\widetilde{X}_n$ can be characterized in terms of the eigenpairs of the following eigenvalue problem:

$$\begin{aligned} &\lambda \in R, \quad u \in V, \ u \neq 0, \\ &(u, v)_V = \lambda(u, v)_H, \quad \text{for all } v \in V. \end{aligned} \tag{5.3}$$

Since V is compact in H, problem (5.3) has eigenvalues and eigenvectors

$$0 < \lambda_1 \leqslant \lambda_2 \leqslant \cdots \nearrow +\infty, \quad u_1, u_2, \ldots,$$

and the eigenvectors (can be chosen to) satisfy

$$(u_i, u_j)_V = \lambda_i(u_i, u_j)_H = \lambda_i\delta_{i,j}. \tag{5.4}$$

As a consequence of (5.2), we have $C^{-1/2} \leqslant \lambda_1$.

We first state a fundamental theorem on n-widths.

Theorem 5.1. *Suppose V and H satisfy* (5.1) *and* (5.2). *Then, for $n = 1, 2, \ldots,$*

$$d_n(V, H) = \lambda_{n+1}^{-1/2} \tag{5.5}$$

and

$$\widetilde{X}_n = \mathrm{span}\{u_1, \ldots, u_n\} \tag{5.6}$$

is an optimal subspace for $d_n(V, H)$.

Remark 5.1. Kolmogorov [18] introduced the notion of n-width and proved Theorem 5.1 (see also [25]). It is known [16,17] that there are optimal subspaces other than the ones given in (5.6); optimal subspaces can, in fact, be selected to be spline subspaces [24]. Optimal subspaces are, however, generally not known explicitly.

The eigenvalue problem (5.3) is written in variational or weak form. In certain situations it is possible to write it in strong form. For basic results on eigenvalue problem we refer to [9].

Theorem 5.1 can be used to determine $d_n(V,H)$ numerically. The numerical values of $d_n(V,H)$ presented in Section 7 were computed via the eigenvalues of Problem (5.3).

We next state a general property of n-widths. This result is given in Theorem 1.1 in [25]; we include it here for the sake of completeness.

Theorem 5.2. *Suppose V and H satisfy* (5.1) *and* (5.2). *Then*

$$d_{n+1}(V,H) \leqslant d_n(V,H). \tag{5.7}$$

In this subsection, we characterized n-widths and optimal subspaces in terms of the eigenvalues of a certain eigenvalue problem, and stated a result showing that spans of the eigenvectors are associated optimal subspaces.

We now state certain properties of $d_n(H^{\alpha,k}_{\text{odd}})$, $d_n(V^{\alpha,k})$, and $d_n(H^{\alpha,k})$. We first present the results on $d_n(H^{\alpha,k}_{\text{odd}})$. We then consider $d_n(V^{\alpha,k})$ and $d_n(H^{\alpha,k})$ together, as they have many common properties.

5.2. Properties of $d_n(H^{\alpha,k}_{odd})$

The eigenpairs of (5.3) can be found analytically for $V = H^{\alpha,k}_{\text{odd}}$; using these formulae we have:

Theorem 5.3. *Suppose $V = H^{\alpha k}_{\text{odd}}(-1,1)$, as defined in* (3.11), *where $k \geqslant 1$ and $\alpha_k \neq 0$. Then, for $n = 1,2,\ldots,$*

$$d_n(H^{\alpha,k}_{\text{odd}}) = \left[\sum_{j=0}^{k} \alpha_j \frac{n^{2j}\pi^{2j}}{2^{2j}}\right]^{-1/2} \tag{5.8}$$

and

$$\widetilde{X}_n = \text{span}\left\{\cos\left(\frac{j\pi}{2}(x+1)\right)\right\}_{j=0}^{n-1} = S(\mathscr{C},n) \tag{5.9}$$

is an optimal subspace for $d_n(H^{\alpha,k}_{\text{odd}})$.

We now consider the dependence of $d_n(H^{\alpha,k}_{\text{odd}})$ on n. We note that $d_n(V,H)$ decreases as n increases, as stated in Theorem 5.2. The next result is an asymptotic estimate for $d_n(H^{\alpha,k}_{\text{odd}})$, which shows that $d_n(H^{\alpha,k}_{\text{odd}}) \to 0$ as $n \to \infty$.

Theorem 5.4. *Suppose $V = H^{\alpha,k}_{\text{odd}}$, where $k \geqslant 1$ and $\alpha_k \neq 0$. Then, for fixed k,*

$$d_n(H^{\alpha,k}_{\text{odd}}) \leqslant C(k)n^{-k}, \tag{5.10}$$

where $C(k) = \alpha_k^{-1/2}(2^k/\pi^k)$.

We next consider the dependence of $d_n(H^{\alpha,k}_{\text{odd}})$ on k.

Theorem 5.5. *Let α be k-independent. Then, for fixed n,*

$$d_n(H^{\alpha,k+1}_{\text{odd}}) \leqslant d_n(H^{\alpha,k}_{\text{odd}}). \tag{5.11}$$

Remark 5.2. Recalling that $\alpha = \alpha(k)$ may depend on k, if we are more precise, (5.8) would read

$$d_n(H^{\alpha,k}_{\text{odd}}) = \left[\sum_{j=0}^{k} \alpha_j(k) \frac{n^{2j}\pi^{2j}}{2^{2j}}\right]^{-1/2}. \tag{5.12}$$

From (5.12) we see that for a fixed n, $d_n(H^{\alpha,k}_{\text{odd}})$ is decreasing in k provided $\alpha_j(i) \leqslant \alpha_j(i+1)$, which is true for α as defined in (3.7b). But this hypothesis is not satisfied by the α defined in (3.7c); nevertheless, examination of (5.12) shows that $d_n(H^{\alpha,k}_{\text{odd}})$ decreases in k in this case.

Theorem 5.6. *Let α be k-independent. Then, for fixed n,*

$$\lim_{k\to\infty} d_n(H^{\alpha,k}_{\text{odd}}) = 0 \tag{5.13}$$

if and only if

$$\sum_{j=0}^{\infty} \alpha_j \frac{n^{2j}\pi^{2j}}{2^{2j}} \text{ is divergent.} \tag{5.14}$$

Remark 5.3. We note that (5.14) is satisfied for all n for the choices of α given in (3.7a), (3.7d) and (3.7f). Thus $d_n(H^{\alpha,k}_{\text{odd}}) \to 0$ as $k \to \infty$ for all n for these choices of α. We further note that if the series in (5.14) converges, then $d_n(H^{\alpha,k}_{\text{odd}}) \geqslant C(n) > 0$ for all k. For α as in (3.7e), the series in (5.14) converges for all n, and hence $d_n(H^{\alpha,k}_{\text{odd}}) \geqslant C(n) > 0$ for all k and n for this α. For α as in (3.7g), the infinite series in (5.14) converges for $n \leqslant 6$ and diverges otherwise. Hence, for this choice of α, we have $d_n(H^{\alpha,k}_{\text{odd}}) \geqslant C(n) > 0$ for all k when $n \leqslant 6$, and $d_n(H^{\alpha,k}_{\text{odd}}) \to 0$ as $k \to \infty$ when $n \geqslant 7$.

Remark 5.4. When α is k-dependent, we see from (5.12) that

$$\lim_{k\to\infty} d_n(H^{\alpha,k}_{\text{odd}}) = 0$$

if and only if

$$\lim_{k\to\infty} \left(\sum_{j=0}^{k} \alpha_j(k) \frac{n^{2j}\pi^{2j}}{2^{2j}}\right) = \infty. \tag{5.15}$$

The condition (5.15) is satisfied for all n by both of our k-dependent choices of α given in (3.7b) and (3.7c). Hence $d_n(H^{\alpha,k}_{\text{odd}}) \to 0$ as $k \to \infty$ for both these choices of α. We further note that if $\sum_{j=0}^{k} \alpha_j(k)(n^{2j}\pi^{2j}/2^{2j})$ is bounded in k for k-dependent α, then $d_n(H^{\alpha,k}_{\text{odd}}) \geqslant C(n) > 0$ for all k. We note however that neither of our k-dependent α satisfy this hypothesis.

Remark 5.5. Let $V = H^{\alpha,k}_{\text{even}}(-1,1)$, where $k \geqslant 1$ and $\alpha_k \neq 0$. Then, similar to Theorem 5.3,

$$d_n(H^{\alpha,k}_{\text{even}}) = \left[\sum_{j=1}^{k} \alpha_j \frac{n^{2j}\pi^{2j}}{2^{2j}}\right]^{-1/2},$$

and

$$\widetilde{X}_n = \text{span}\left\{ \sin\left(\frac{j\pi}{2}(x+1)\right)\right\}_{j=1}^{n} = S(S,n)$$

is an optimal subspace for $d_n(H^{\alpha,k}_{\text{even}})$.

We found explicit formulae for n-widths, $d_n(H^{\alpha,k}_{\text{odd}})$ and $d_n(H^{\alpha,k}_{\text{even}})$. We also indicated that the cosines are optimal shape functions with respect to $H^{\alpha,k}_{\text{odd}}$, and the sines are optimal shapes functions with respect to $H^{\alpha,k}_{\text{even}}$, for any choice of α and k. Thus we see that the trigonometric polynomials are "ideal" shape functions for any α and k; but it should be stressed that they are ideal only for functions satisfying the boundary conditions (restraints) that inclusion in the spaces $H^{\alpha,k}_{\text{odd}}$ of $H^{\alpha,k}_{\text{even}}$ entail.

We showed that for fixed n, $d_n(H^{\alpha,k}_{\text{odd}})$ decreases in k. Also, as expected, $d_n(H^{\alpha,k}_{\text{odd}})$ decreases in n for fixed k.

5.3. Properties of $d_n(V^{\alpha,k})$ and $d_n(H^{\alpha,k})$

We first state a result comparing $d_n(V^{\alpha,k})$ and $d_n(H^{\alpha,k})$.

Theorem 5.7. *For any α, k, and n we have*

$$d_n(H^{\alpha,k}) \leqslant d_n(V^{\alpha,k}). \tag{5.16}$$

Next, the eigenpairs of (5.3) can be found analytically for $V = V^{\alpha,k}$; using these formulae we have:

Theorem 5.8. *Suppose $V = V^{\alpha,k}(-1,1)$, as defined in* (3.14), *where $1 \leqslant k \leqslant \infty$. Then*

$$d_n(V^{\alpha,k}) = \left[\sum_{j=0}^{\min(k,n)} \alpha_j \frac{(n+j)!}{(n-j)!}\right]^{-1/2} \tag{5.17}$$

and

$$\widetilde{X}_n = \text{span}\{L_0(x),\ldots,L_{n-1}(x)\} = S(\mathscr{P},n) \tag{5.18}$$

is an optimal subspace for $d_n(V^{\alpha,k})$ ($L_j(x)$ is the jth degree Legendre polynomial).

Remark 5.6. Theorem 5.8 shows that $V^{\alpha,k}$ are natural spaces for understanding approximation by polynomials. Similar results hold also in higher dimensions when the domain is a cube. They were used in the analysis of p-version of FEM (see [4,5]).

Remark 5.7. We note that for k-dependent α, the formula (5.17) in Theorem 5.8 should be written as

$$d_n(V^{\alpha,k}) = \left[\sum_{j=0}^{\min(k,n)} \alpha_j(k) \frac{(n+j)!}{(n-j)!}\right]^{-1/2}. \tag{5.19}$$

Remark 5.8. With V as in Theorems 5.3 and 5.8, we see that the optimal subspaces are independent of V, i.e., they are independent of α and k in $H^{\alpha,k}_{\text{odd}}$ and $V^{\alpha,k}$. Specifically, from Theorem 5.3, we see that $S(\mathscr{C},n)$ is optimal for $d_n(H^{\alpha,k}_{\text{odd}})$ for all α and k, and from Theorem 5.8, $S(\mathscr{P},n)$ is optimal for $d_n(V^{\alpha,k})$ for all α and k. We note that these are very exceptional situations. Usually optimal subspaces depend very strongly on the space V.

We now consider the dependence of $d_n(H^{\alpha,k})$ and $d_n(V^{\alpha,k})$ on n. We again note that $d_n(V,H)$ decreases as n increases, as stated in Theorem 5.2. The next theorem gives asymptotic estimates for $d_n(V^{\alpha,k})$ and $d_n(H^{\alpha,k})$, which show that they approach 0 as $n \to \infty$.

Theorem 5.9. *Suppose $V = H^{\alpha,k}(-1,1)$ or $V^{\alpha,k}(-1,1)$, where $\alpha_k \neq 0$. Then, for fixed k,*

$$d_n(H^{\alpha,k}) \leqslant d_n(V^{\alpha,k}) \leqslant C(k)n^{-k}, \quad \text{for } n \geqslant k, \tag{5.20}$$

where $C(k) = \alpha_k^{-1/2}(e/2)^k$.

Theorem 5.10. *Suppose $V = H^{\alpha,k}(-1,1)$ or $V^{\alpha,k}(-1,1)$, where α is as defined in* (3.7b), (3.7d), (3.7e), (3.7f), *or* (3.7g). *Then, for $n \leqslant k$, we have*

$$d_n(H^{\alpha,k}) \leqslant d_n(V^{\alpha,k}) \leqslant \begin{cases} ((2n)!)^{-1/2}, & \text{for } \alpha \text{ as in (3.7b)}, \\ (n!(2n)!)^{-1/2}, & \text{for } \alpha \text{ as in (3.7d)}, \\ ((2n)!/n!)^{-1/2}, & \text{for } \alpha \text{ as in (3.7e)}, \\ (10^{2n}(2n)!)^{-1/2}, & \text{for } \alpha \text{ as in (3.7f)}, \\ (10^{-2n}(2n)!)^{-1/2}, & \text{for } \alpha \text{ as in (3.7g)} \end{cases} \tag{5.21}$$

showing that the rate of decrease of $d_n(H^{\alpha,k})$ and $d_n(V^{\alpha,k})$ with respect to n is higher than exponential, for large k.

Remark 5.9. The result of Theorem 5.10 shows the pre-asymptotic dependence of $d_n(H^{\alpha,k})$ and $d_n(V^{\alpha,k})$ on n for fixed k. This result is valid only for $n \leqslant k$, and not for all n. But, when k is large, the rate of decrease of $d_n(H^{\alpha,k})$ and $d_n(V^{\alpha,k})$ is $o(C^{-n})$ for any $C > 1$, i.e., higher than exponential rate as long as $n \leqslant k$.

We next consider $d_n(V^{\alpha,k})$ and $d_n(H^{\alpha,k})$ in their dependence in k.

Theorem 5.11. *Suppose $V = H^{\alpha,k}(-1,1)$ or $V^{\alpha,k}(-1,1)$, where α is k-independent. Then, for fixed n,*

$$d_n(H^{\alpha,k+1}) \leqslant d_n(H^{\alpha,k}) \tag{5.22}$$

and

$$d_n(V^{\alpha,k+1}) \leqslant d_n(V^{\alpha,k}). \tag{5.23}$$

Remark 5.10. (5.22) holds also for k-dependent α, where $\alpha = \alpha(k)$ is such that

$$H^{\alpha(k+1),k+1} \subset H^{\alpha(k),k} \quad \text{and} \quad \|u\|_{H^{\alpha(k),k}} \leqslant \|u\|_{H^{\alpha(k+1),k+1}}. \tag{5.24}$$

It is easily seen that α given in (3.7b) satisfies (5.24), and hence (5.22) holds for this α. Likewise (5.23) holds for any k-dependent $\alpha = \alpha(k)$ satisfying

$$V^{\alpha(k+1),k+1} \subset V^{\alpha(k),k} \quad \text{and} \quad \|u\|_{V^{\alpha(k),k}} \leqslant \|u\|_{V^{\alpha(k+1),k+1}}. \tag{5.25}$$

It is easily seen that (5.25) holds for α given in (3.7b), and hence (5.23) holds for this α. We further note that the hypothesis imposed on k-dependent α in Remark 5.2 implies conditions (5.24) and (5.25).

Remark 5.11. Conditions (5.24) and (5.25) do not hold for $\alpha = \alpha(k)$ defined in (3.7c). A careful examination of (5.19) for this α shows that $d_n(V^{\alpha,k})$ decreases in k and is less than 1 as long as $k \leqslant n$, and $d_n(V^{\alpha,k}) = 1$ for $k > n$. The dependence of $d_n(H^{\alpha,k})$ on k for this α is similar. From computational results in Section 6 (cf. Table 5(c)), we see that $d_n(H^{\alpha,k})$ decreases in k and is less than 1 for $k \leqslant n$. The result for $k > n$ is given in the next theorem.

Eigenpairs of (5.3) can be found analytically for $V = H^{\alpha,k}$ for α given in (3.7c) when $k > n$; using these formulae we have:

Theorem 5.12. *Suppose $V = H^{\alpha,k}(-1,1)$, with α as in* (3.7c) *and $k \geqslant 2$. Then, for $k > n$,*

$$d_n(H^{\alpha,k}) = 1 \tag{5.26}$$

and

$$\widetilde{X}_n = \operatorname{span}\{L_0(x), \ldots, L_{n-1}(x)\} = S(\mathscr{P}, n) \tag{5.27}$$

is an optimal subspace for $d_n(H^{\alpha,k})$.

Theorem 5.13. *Suppose $V = H^{\alpha,k}(-1,1)$ or $V^{\alpha,k}(-1,1)$, where α is k-independent. Then, for fixed n,*

$$\lim_{k\to\infty} d_n(H^{\alpha,k}) = d_n(H^{\alpha,\infty}) > 0 \tag{5.28}$$

and

$$\lim_{k\to\infty} d_n(V^{\alpha,k}) = d_n(V^{\alpha,\infty}) > 0. \tag{5.29}$$

Remark 5.12. The results of Theorem 5.13 is not generally true for k-dependent α. For one of our choices of k-dependent α, namely for α as in (3.7b), one can easily show that, for fixed n,

$$\lim_{k\to\infty} d_n(V^{\alpha,k}) = 0 \tag{5.30}$$

and from Theorem 5.7 and (5.30),

$$\lim_{k\to\infty} d_n(H^{\alpha,k}) = 0. \tag{5.31}$$

But, for α defined in (3.7c), we have

$$\lim_{k\to\infty} d_n(V) = 1, \tag{5.32}$$

where $V = H^{\alpha,k}$ or $V^{\alpha,k}$.

We derived explicit formulae for $d_n(V^{\alpha,k})$, and indicated that the Legendre polynomials are optimal shape functions with respect to $V^{\alpha,k}$. Thus the relation of the Legendre polynomials to the spaces $V^{\alpha,k}$ is similar to relation of the trigonometric polynomials to the spaces $H^{\alpha,k}_{\mathrm{odd}}$ and $H^{\alpha,k}_{\mathrm{even}}$. Nevertheless, the general features of the Legendre polynomials and the trigonometric polynomials are very different. For the spaces $H^{\alpha,k}$ such ideal shape functions—ideal in the sense that one family of shape function is simultaneously optimal for all α and k—do not exist. However, as we will see later, algebraic polynomial are "almost ideal" with respect to these spaces.

We have shown that for fixed n, $d_n(V)$, with $V = V^{\alpha,k}$ or $H^{\alpha,k}$, generally decreases in k. But it can also increase in k for some α, and thus k dependence of $d_n(V)$ is different from that of $d_n(H^{\alpha,k}_{\mathrm{odd}})$. Also, the rate of decrease of $d_n(V)$ in n, for fixed k, is faster than exponential in the pre-asymptotic range, which is not true for $d_n(H^{\alpha,k}_{\mathrm{odd}})$.

So far we have discussed several properties of $d_n(V,H)$. We now consider $\Psi(\Phi, n, V, H)$, defined in (4.2), where V and H satisfy (5.1) and (5.2), and $S(\Phi, n)$ is defined in (4.1). We will consider the spaces $V = V^{\alpha,k}(-1,1)$ and $H^{\alpha,k}(-1,1)$. Some of our results will be for the specific bases $\Phi = \mathscr{P}$ and $\mathscr{C}$, which were defined in Section 4.

5.4. Characterization of $\Psi(\Phi, n, V, H)$

We will characterize $\Psi(\Phi, n, V, H)$ in terms of the eigenvalues of a certain eigenvalue problem. Let π_n be the H-orthogonal projection onto $S(\Phi, n) = \text{span}\{\phi_1, \ldots, \phi_n\}$. Consider the eigenvalue problem

$$\begin{cases} u \in V, \\ (u, v)_V = \tilde{\lambda}((I - \pi_n)u, (I - \pi_n)v)_H, \quad \text{for all } v \in V. \end{cases} \tag{5.33}$$

(5.33) is a well-posed eigenvalue problem with eigenvalues

$$0 < \tilde{\lambda}_{n+1} \leqslant \tilde{\lambda}_{n+2} \leqslant \cdots$$

The smallest eigenvalue, $\tilde{\lambda}_{n+1}$, of this problem is characterized by

$$\tilde{\lambda}_{n+1} = \inf_{u \in V} \frac{\|u\|_V^2}{\|u - \pi_n u\|_H^2}. \tag{5.34}$$

For another characterization of $\tilde{\lambda}_{n+1}$, see [2].

We now present theorems relating $\Psi(\Phi, n, V, H)$, $\tilde{\lambda}_{n+1}$, and λ_{n+1}.

Theorem 5.14. *If $\tilde{\lambda}_{n+1}$ is the smallest eigenvalue of* (5.33), *then*

$$\Psi(\Phi, n, V, H) = \tilde{\lambda}_{n+1}^{-1/2}. \tag{5.35}$$

Theorem 5.15. *If λ_{n+1} is the $(n+1)$th eigenvalue of* (5.3) *and $\tilde{\lambda}_{n+1}$ is the smallest eigenvalue of* (5.33), *then*

$$\lambda_{n+1} \geqslant \tilde{\lambda}_{n+1} \tag{5.36}$$

which is equivalent to

$$d_n(V, H) \leqslant \Psi(\Phi, n, V, H). \tag{5.37}$$

The next result characterizes $\Psi(\Phi, n, V, H)$ as an optimal error bound.

Theorem 5.16. *Suppose each $u \in V$ can be approximated by $\phi_u \in S(\Phi, n)$ so that*

$$\|u - \phi_u\|_H \leqslant \Omega(\Phi, n, V, H)\|u\|_V. \tag{5.38}$$

Then

$$\Psi(\Phi, n, V, H) \leqslant \Omega(\Phi, n, V, H). \tag{5.39}$$

Remark 5.13. In classical approximation theory, e.g., in the finite element method, we typically construct a ϕ_u (e.g. by interpolation) and estimate the error with $\Omega(\Phi, n, V, H)$. We note that d_n, Ψ, and $\Omega(\Phi, n, V, H)$ are related by

$$d_n(V, H) \leqslant \Psi(\Phi, n, V, H) \leqslant \Omega(\Phi, n, V, H) \tag{5.40}$$

thus $\Psi(\Phi, n, V, H)$ is an optimal error bound for estimates of type (5.38).

Our next result concerns the dependence of $\Psi(\Phi, n, V, H)$ on n in a general setting.

Theorem 5.17. *For any V and H satisfying* (5.1) *and* (5.2), *and for any basis Φ in H, we have*

$$\Psi(\Phi, n+1, V, H) \leqslant \Psi(\Phi, n, V, H). \tag{5.41}$$

The sup–inf, $\Psi(\Phi, v, V, H)$, was characterized in terms of the eigenvalues of a certain eigenvalue problem. This characterization was used to establish the relation between $d_n(V, H)$ and $\Psi(\Phi, v, V, H)$. The relation follows immediately, of course, from the definitions of the n-width and the sup–inf.

5.5. Properties of $\Psi(\Phi, n, V)$ for $V = V^{\alpha,k}$ or $H^{\alpha,k}$ with specific bases

We will present results for $\Psi(\Phi, n, V)$ that are analogous to the results in Theorems 5.9, 5.11 and 5.13. We first note from the definition of $\Psi(\Phi, n, V)$ and (5.18) that

$$\Psi(\mathscr{P}, n, V^{\alpha,k}) = d_n(V^{\alpha,k}). \tag{5.42}$$

The next result compares $\Psi(\Phi, n, V^{\alpha,k})$ and $\Psi(\Phi, n, H^{\alpha,k})$ for any basis Φ.

Theorem 5.18. *For any α, k, and n, and for any basis Φ in L_2 we have,*

$$\Psi(\Phi, n, H^{\alpha,k}) \leqslant \Psi(\Phi, n, V^{\alpha,k}). \tag{5.43}$$

We now consider the dependence of $\Psi(\mathscr{P}, n, V)$ on n for $V = V^{\alpha,k}$ or $H^{\alpha,k}$. From Theorem 5.17, we know that $\Psi(\Phi, n, V)$ decreases as n increases. The next result gives an asymptotic estimate for Ψ with $\Phi = \mathscr{P}$, which shows that $\Psi(\mathscr{P}, n, V) \to 0$ as $n \to \infty$ for $V = V^{\alpha,k}$ or $H^{\alpha,k}$.

Theorem 5.19. *Suppose $V = H^{\alpha,k}(-1,1)$ or $V^{\alpha,k}(-1,1)$, where $\alpha_k \neq 0$. Then, for fixed k, we have*

$$\Psi(\mathscr{P}, n, H^{\alpha,k}) \leqslant \Psi(\mathscr{P}, n, V^{\alpha,k}) \leqslant C(k)n^{-k}, \quad n > k, \tag{5.44}$$

where $C(k) = \alpha_k^{-1/2}(e/2)^k$.

Remark 5.14. From (5.42) and (5.43) with $\Phi = \mathscr{P}$, it is clear that for $n \leqslant k$, the estimate (5.21) in Theorem 5.10 is also valid for $\Psi(\mathscr{P}, n, H^{\alpha,k})$ and $\Psi(\mathscr{P}, n, V^{\alpha,k})$. In other words, we can replace $d_n(H^{\alpha,k})$ and $d_n(V^{\alpha,k})$ in (5.21) by $\Psi(\mathscr{P}, n, H^{\alpha,k})$ and $\Psi(\mathscr{P}, n, V^{\alpha,k})$, respectively.

We now consider the dependence of $\Psi(\mathscr{C}, n, H^{\alpha,k})$ on n.

Theorem 5.20. *Suppose $V = H^{\alpha,k}(-1,1)$, with $\alpha_k \neq 0$, and let $\Phi = \mathscr{C}$. Then, for $k = 1$, there are positive constants $C_1(\alpha_1)$ and $C_2(\alpha_1)$, which depend on α_1 but are independent of n, such that*

$$\frac{C_1(\alpha_1)}{n} \leqslant \Psi(\mathscr{C}, n, H^{\alpha,k}) \leqslant \frac{C_2(\alpha_1)}{n}. \tag{5.45}$$

For $k \geqslant 2$, there are positive constants $\widetilde{C}_1(\alpha, k)$ and $\widetilde{C}_2(\alpha, k)$, which depend on α and k but are independent of n, such that

$$\frac{\widetilde{C}_1(\alpha, k)}{n^{3/2}} \leqslant \Psi(\mathscr{C}, n, H^{\alpha,k}) \leqslant \frac{\widetilde{C}_2(\alpha, k)}{n^{3/2}}. \tag{5.46}$$

Remark 5.15. From Theorems 5.18 and 5.20 it is clear that

$$\Psi(\mathscr{C}, n, V^{\alpha,k}) \geqslant \frac{C_1(\alpha_1)}{n}, \quad \text{for } k = 1$$

and

$$\Psi(\mathscr{C}, n, V^{\alpha,k}) \geqslant \frac{\tilde{C}_1(k,\alpha)}{n^{3/2}}, \quad \text{for } k \geqslant 2.$$

Remark 5.16. Using (4.3) and Theorems 5.8 and 5.9, we obtain an estimate for the error in the best L_2-approximation by $\Phi = \mathscr{P}$ of functions in $V^{\alpha,k}$:

$$\inf_{\chi \in S(\mathscr{P},n)} \|u-\chi\|_{L_2} \leqslant C(k) n^{-k} \|u\|_{V^{\alpha,k}}, \quad \text{for all } u \in V^{\alpha,k}, \tag{5.47}$$

where $C(k) = \alpha_k^{-1/2}(e/2)^k$. Using (4.3) and Theorems 5.3 and 5.4, we obtain an estimate for the error in the best L_2-approximation by $\Phi = \mathscr{C}$ for functions in $H^{\alpha,k}_{\text{odd}}$:

$$\inf_{\chi \in S(\mathscr{C},n)} \|u-\chi\|_{L_2} \leqslant C(k) n^{-k} \|u\|_{H^{\alpha,k}_{\text{odd}}}, \quad \text{for all } u \in H^{\alpha,k}_{\text{odd}}, \tag{5.48}$$

where $C(k) = \alpha_k^{-1/2}(2/\pi)^k$. Estimates (5.47) and (5.48) explain certain computational results in Section 1; cf. (3.15) and (3.16).

We now consider $\Psi(\Phi, n, H^{\alpha,k})$ and $\Psi(\Phi, n, V^{\alpha,k})$ in their dependence in k.

Theorem 5.21. *Suppose $V = H^{\alpha,k}(-1,1)$ or $V^{\alpha,k}(-1,1)$, where α is k-independent. Let Φ be a basis in L_2. Then, for fixed n,*

$$\Psi(\Phi, n, H^{\alpha,k+1}) \leqslant \Psi(\Phi, n, H^{\alpha,k}) \tag{5.49}$$

and

$$\Psi(\Phi, n, V^{\alpha,k+1}) \leqslant \Psi(\Phi, n, V^{\alpha,k}). \tag{5.50}$$

Remark 5.17. This theorem is also valid for any k-dependent α provided $\alpha = \alpha(k)$ satisfies (5.24) and (5.25). These conditions are true for the k-dependent α defined in (3.7b), as mentioned in Remark 5.10. Therefore (5.49) and (5.50) hold for α defined in (3.7b).

Remark 5.18. It was mentioned in Remark 5.11 that conditions (5.24) and (5.25) are not valid for $\alpha = \alpha(k)$ as defined in (3.7c). For this α and for $\Phi = \mathscr{P}$, it is clear from (5.42) and Remark 5.11 that $\Psi(\mathscr{P}, n, V^{\alpha,k})$ decreases in k and is less than 1 as long as $k \leqslant n$, and $\Psi(\mathscr{P}, n, V^{\alpha,k}) = 1$ for $k > n$. Also, from (5.37) with $V = H^{\alpha,k}$ and $H = L_2$, and Theorem 5.12, it is clear that $\Psi(\mathscr{P}, n, H^{\alpha,k}) \geqslant 1$ for $k > n$, for this α and $\Phi = \mathscr{P}$. In fact, $\Psi(\mathscr{P}, n, H^{\alpha,k}) = 1$ for $k > n$. We have observed from our computations [2] that, for this value of α, $\Psi(\mathscr{P}, n, H^{\alpha,k})$ decreases in k and is less than 1, for $k \leqslant n$.

Theorem 5.22. *Suppose $V = H^{\alpha,k}(-1,1)$ or $V^{\alpha,k}(-1,1)$, where α is k-independent. Let Φ be a basis in L_2. Then, for fixed n,*

$$\lim_{k\to\infty} \Psi(\Phi, n, H^{\alpha,k}) \geqslant \Psi(\Phi, n, H^{\alpha,\infty}) > 0 \tag{5.51}$$

and

$$\lim_{k\to\infty} \Psi(\Phi, n, V^{\alpha,k}) \geqslant \Psi(\Phi, n, V^{\alpha,\infty}) > 0. \tag{5.52}$$

Remark 5.19. The results of Theorem 5.22 are generally not true for k-dependent α. For α as defined in (3.7b), it is clear from (5.42) and Remark 5.12 that $\Psi(\mathscr{P}, n, V^{\alpha,k}) \to 0$ as $k \to \infty$ for a fixed n. Using this fact in (5.43) with $\Phi = \mathscr{P}$, we also see (for the same α) that $\Psi(\mathscr{P}, n, H^{\alpha,k}) \to 0$ as $k \to \infty$. For α as defined in (3.7c), it is clear from Remark 5.18 that $\lim_{k\to\infty} \Psi(\mathscr{P}, n, V) = 1$, for $V = V^{\alpha,k}$ or $H^{\alpha,k}$.

We have shown that k dependence of $\Psi(\mathscr{P}, n, H^{\alpha,k})$ for fixed n, and n dependence of $\Psi(\mathscr{P}, n, H^{\alpha,k})$ for fixed k, are similar to the corresponding k and n dependence of $d_n(H^{\alpha,k})$. We have noted before that there does not exist an ideal basis, independent of k and α, with respect to $H^{\alpha,k}$; but we see that polynomials behave almost like an ideal basis with respect to these spaces. We have also shown that cosines perform much worse than algebraic polynomials with respect to $H^{\alpha,k}$.

We also stated the dependence of $\Psi(\Phi, n, H^{\alpha,k})$ on k for fixed n for any basis Φ; in particular, we stated the limiting behavior as $k \to \infty$. This limit statement shows that for fixed n the difference between the approximation errors for "reasonably smooth" functions (k of moderate size) and for "very smooth" function ($k = \infty$) is not large.

We now present one of the main results of the paper.

5.6. Near-optimal performance of $\Phi = \mathscr{P}$ in $H^{\alpha,k}$

We know from Theorem 5.8 that the performance of the basis $\Phi = \mathscr{P}$ is optimal in the space $V = V^{\alpha,k}$, with $H = L_2$. We now show the effectiveness of $\Phi = \mathscr{P}$ for $V = H^{\alpha,k}$, with $H = L_2$. To assess this effectiveness, we use the ratio $\Lambda(\Phi, n, V, H)$ introduced in (4.8), with $\Phi = \mathscr{P}$, $V = H^{\alpha,k}$, and $H = L_2$:

$$\Lambda(\mathscr{P}, n, H^{\alpha,k}) = \frac{\Psi(\mathscr{P}, n, H^{\alpha,k})}{d_n(H^{\alpha,k})}.$$

It is clear from the definition of n-widths that $\Lambda(\mathscr{P}, n, H^{\alpha,k}) \geqslant 1$, and it is likely that $\Lambda(P, n, H^{\alpha,k}) > 1$, since it is unlikely that the polynomials are the optimal shape functions in any of the space $H^{\alpha,k}$. For the rest of the results of this section, we consider a class of k-independent α satisfying $\alpha_j \leqslant \alpha_{j+1}$. In the following theorem we show that $\Lambda(\mathscr{P}, n, H^{\alpha,k})$ is almost uniformly bounded in the sense defined at the end of Section 4.

Theorem 5.23. *Suppose $V = H^{\alpha,k}(-1,1)$, where α is k-independent and satisfies $\alpha_j \leqslant \alpha_{j+1}$, and $H = L_2(-1,1)$. Then, for each $0 < \sigma \leqslant 1$, there is a constant $C(\sigma)$, which depends on σ but is independent of n and k, such that*

$$\Lambda_\sigma(\mathscr{P}, n, H^{\alpha,k}) = \frac{\Psi(\mathscr{P}, n, H^{\alpha,k})}{d_n(H^{\alpha,k})^{1-\sigma}} \leqslant C(\sigma), \quad \textit{for all n and k}. \tag{5.53}$$

Remark 5.20. Because of (5.53), we say that algebraic polynomials are *almost uniformly optimal* with respect to the family $H^{\alpha,k}$. We thus expect the algebraic polynomials to be effective for approximating any $u \in H^{\alpha,k}$ for any k.

We presented one of the main results of the paper, namely, that algebraic polynomials are almost uniformly optimal with respect to the family $H^{\alpha,k}$. This uniformity is with respect to $k = 1, 2, \ldots$ and α satisfying $\alpha_j \leqslant \alpha_{j+1}$. Thus, polynomials are robust; they are effective in the GFEM for a larger class of problems. This feature is also valid for higher dimensional problems. It explains why polynomial shape functions perform well in the FEM.

We now present the other main results of this section.

5.7. Limited performance of $\Phi = \mathscr{P}$ in $H^{\alpha,k}_{odd}$

The following results show that polynomials do not perform as well as the optimal shape functions for the space $H^{\alpha,k}_{\rm odd}$. We first recall from Theorem 5.3 that $\Phi = \mathscr{C}$ is optimal when $V = H^{\alpha,k}_{\rm odd}$. To assess the effectiveness of $\Phi = \mathscr{P}$ in $H^{\alpha,k}_{\rm odd}$, we again consider the ratio Λ:

$$\Lambda(\mathscr{P}, n, H^{\alpha,k}_{\rm odd}) = \frac{\Psi(\mathscr{P}, n, H^{\alpha,k}_{\rm odd})}{d_n(H^{\alpha,k}_{\rm odd})}.$$

Our next result establishes a lower bound for $\Lambda(\mathscr{P}, n, H^{\alpha,k}_{\rm odd})$.

Theorem 5.24. *Suppose $V = H^{\alpha,k}_{\rm odd}(-1,1)$, where α is k-independent and satisfies $\alpha_j \leqslant \alpha_{j+1}$. Then, for each n, there is a positive constant $C(n)$, which depends on n but is independent of k, such that*

$$\frac{\Psi(\mathscr{P}, n, H^{\alpha,k}_{\rm odd})}{d_n(H^{\alpha,k}_{\rm odd})} \geqslant C(n)n^k. \tag{5.54}$$

Remark 5.21. We note that, for a fixed n, the ratio $\Psi(\mathscr{P}, n, H^{\alpha,k}_{\rm odd})/d_n(H^{\alpha,k}_{\rm odd})^{1-\sigma}$ is also unbounded for $0 < \sigma \leqslant 1$.

Theorem 5.25. *Suppose $V = H^{\alpha,k}_{\rm odd}(-1,1)$, where α is k-independent and satisfies $\alpha_j \leqslant \alpha_{j+1}$. Then, for each $0 < \sigma \leqslant 1$, there is a positive constant $C(\sigma)$, which depends on σ but is independent of n and k, such that*

$$\frac{\Psi(\mathscr{P}, n, H^{\alpha,k}_{\rm odd})}{d_n(H^{\alpha,k}_{\rm odd})^{1-\sigma}} \leqslant C(\sigma), \quad \textit{for } k \leqslant n. \tag{5.55}$$

Remark 5.22. Theorem 5.25 shows that for a fixed k, the polynomials are nearly optimal in $H^{\alpha,k}_{\rm odd}$ in the asymptotic range, i.e., when $n > k$. On the other hand, Theorem 5.24 shows that for large k, the polynomials are far from optimal in $H^{\alpha,k}_{\rm odd}$ in the pre-asymptotic range, i.e., for $n \leqslant k$.

We presented two results, one showing that algebraic polynomials are not recommended for pre-asymptotic approximation of functions in the spaces $H^{\alpha,k}_{\rm odd}$, the other showing that polynomials are nearly optimal in the asymptotic range.

6. Numerical results and their interpretations

In the previous sections we have addressed the major theoretical issues involved in the selection of shape functions for the GFEM, where the approximation was considered in L_2, as pointed out in Sections 2 and 5. In this section, we present detailed numerical results. These results illustrate the theoretical results in Section 5, and the results in Section 5 explain the numerical results presented here.

In Table 2, we report values of the n-width, $d_n(H^{\alpha,k})$, for $\alpha_j = 1$, as given in (3.7a). These values were computed using Theorem 5.1 (see [2] for computational details).

We see that for fixed n, $d_n(H^{\alpha,k})$ decreases as k increases, i.e., the n-width decreases as the smoothness of the functions increase. But the decrease of $d_n(H^{\alpha,k})$ with respect to k is visible only up to about $k = n+1$, after which $d_n(H^{\alpha,k})$ appears to reach a positive limiting value: $\lim_{k\to\infty} d_n(H^{\alpha,k}) > 0$. This behavior illustrates the result in Theorem 5.13.

Table 2
The values of the n-width, $d_n(H^{\alpha,k})$, for $\alpha_j = 1$

k	n							
	1	2	3	4	5	6	7	8
1	0.537	0.303	0.208	0.157	0.126	0.106	0.906e−1	0.793e−1
2	0.503	0.150	0.613e−1	0.323e−1	0.197e−1	0.133e−1	0.953e−2	0.717e−2
3	0.501	0.130	0.272e−1	0.898e−2	0.391e−2	0.203e−2	0.118e−2	0.747e−3
4	0.501	0.129	0.222e−1	0.363e−2	0.103e−2	0.387e−3	0.176e−3	0.909e−4
5	0.501	0.128	0.217e−1	0.282e−2	0.383e−3	0.958e−4	0.323e−4	0.132e−4
6	0.501	0.128	0.217e−1	0.274e−2	0.285e−3	0.334e−4	0.757e−5	0.232e−5
7	0.501	0.128	0.217e−1	0.273e−2	0.275e−3	0.239e−4	0.249e−5	0.517e−6
8	0.501	0.128	0.217e−1	0.273e−2	0.274e−3	0.230e−4	0.172e−5	0.161e−6

We make several observations: (a) For fixed n, $d_n(H^{\alpha,k})$ decreases in k, i.e., decreases as the smoothness of the functions increase, and converges to a positive limit. For $k > n+1$, $d_n(H^{\alpha,k})$ is essentially constant in k. (b) For fixed k, $d_n(H^{\alpha,k})$ decreases in n, and the higher the k, the more rapid the decrease in n. This is the effect of the smoothness of the approximated functions.

The values in Table 2 also reveal that for fixed k, $d_n(H^{\alpha,k})$ decreases as n increases, and the larger the k, the more rapid the decrease in n. This is the effect of the smoothness of the approximated functions. That $d_n(H^{\alpha,k})$ is decreasing in n, for fixed k, follows from Theorem 5.2. Fig. 2a and b are log–log plots of $d_n(H^{\alpha,k})$ versus n, for $k = 2$ and 7. For $k = 2$ we see that the slope is -2 for $n \geqslant 2$. Thus for $k = 2$ the graph is in the asymptotic range when $n \geqslant 2$, and the order of convergence is $O(n^{-2})$, which is algebraic. This observed rate matches the rate in Theorem 5.9.

Fig. 2b reveals additional features of $d_n(H^{\alpha,k})$. The graph first decreased sharply and is concave down, then at about the point $n = k = 7$ (marked with an $*$), the sense of concavity changes and the magnitude of the slope begins to decrease. This means that $d_n(H^{\alpha,k})$ has not reached the asymptotic range at $n = 8$, but the change in concavity suggests that it will reach the asymptotic range for higher values of n, as was established in Theorem 5.9. Because of this change in the concavity, the curve is referred to as S-shaped. That the graph is sharply decreasing and concave down in the pre-asymptotic range is partially explained in Theorem 5.10 and Remark 5.9. To see this we need only to note that any of the expressions on the right side of estimate (5.21) in Theorem 5.10 are sharply decreasing and concave down in n. Based on Remark 5.9, we say that $d_n(H^{\alpha,k})$ is decreasing faster than exponentially. Fig. 2a shows similar, but much less striking, behavior; in this graph, the $*$ marks the point $n = k = 2$. Thus, the convergence of $d_n(H^{\alpha,k})$ has two phases.

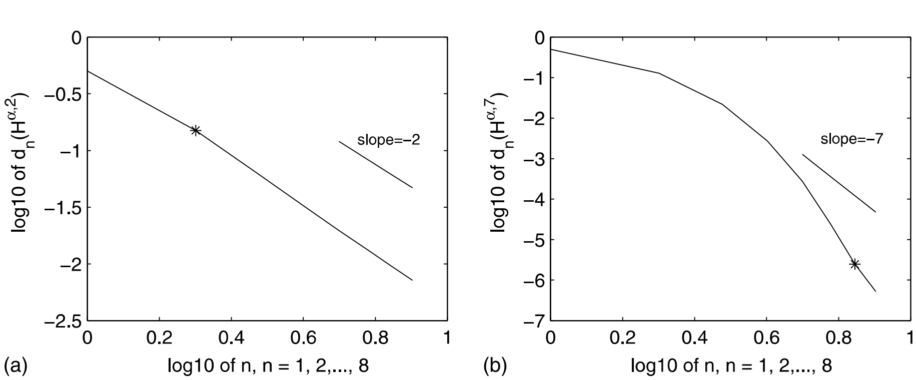

Fig. 2. The graphs of the n-width, $d_n(H^{\alpha,k})$, for $\alpha_j = 1$, and $k = 2$ and 7. The $*$ indicates the point (n, d_n) for $n = k$. (a) The graph is in the asymptotic range when $n \geqslant 2$, and the rate of convergence is 2 (algebraic); in (b), the graph is eventually in the asymptotic range, and the rate of convergence is 7 (algebraic). (b) The decrease of $d_n(H^{\alpha,7})$ is higher than exponential in the pre-asymptotic range ($n < 8$).

In the pre-asymptotic phase ($n < k$), the decrease is faster than exponential; whereas in the asymptotic phase ($n \gg k$), the rate is algebraic. Such S-shaped curves also occur in the convergence of the error in the p-version of the FEM [28].

In Table 3 we present the n-width, $d_n(V^{\alpha,k})$, for $\alpha_j = 1$. These values were computed using Theorem 5.8.

We see from Table 3 that n-widths for $V^{\alpha,k}$ have the same features as those for $H^{\alpha,k}$, although the functions in $V^{\alpha,k}$ are less smooth near the boundary. This is explained by the fact that Theorems 5.9–5.11 and 5.13 are valid for both $H^{\alpha,k}$ and $V^{\alpha,k}$. Also from Theorem 5.7, we have $d_n(H^{\alpha,k}) \leqslant d_n(V^{\alpha,k})$, but we see that the differences between $d_n(H^{\alpha,k})$ and $d_n(V^{\alpha,k})$ are not large.

Next, in Table 4, we present the values of $d_n(H^{\alpha,k}_{\text{odd}})$, which were computed using Theorem 5.3. We note that inclusion in $H^{\alpha,k}_{\text{odd}}$, in contrast to inclusion in $V^{\alpha,k}$, constrains boundary behavior.

Table 4 shows that the values of $d_n(H^{\alpha,k}_{\text{odd}})$ are qualitatively different than the values of $d_n(H^{\alpha,k})$ and $d_n(V^{\alpha,k})$; we mention two such differences. For fixed n, $d_n(H^{\alpha,k}_{\text{odd}}) \to 0$ as $k \to \infty$. This is explained in Remark 5.3. For fixed k, the order of convergence of $d_n(H^{\alpha,k}_{\text{odd}})$ is $O(n^{-k})$, as indicated in Theorem 5.4. But, in contrast to $d_n(H^{\alpha,k})$ and $d_n(V^{\alpha,k})$, the decrease in $d_n(H^{\alpha,k}_{\text{odd}})$ is not faster than exponential in the pre-asymptotic range, and the plot of $d_n(H^{\alpha,k}_{\text{odd}})$ is not S-shaped. This is illustrated in Fig. 3(a) and (b), where we have plotted $d_n(H^{\alpha,2}_{\text{odd}})$ and $d_n(H^{\alpha,7}_{\text{odd}})$. These plots should be compared with Fig. 2(a) and (b).

Other properties of $d_n(H^{\alpha,k}_{\text{odd}})$, which can be derived from (5.8), are similar to those of $d_n(H^{\alpha,k})$ and $d_n(V^{\alpha,k})$.

Table 3
The values of the n-width, $d_n(V^{\alpha,k})$, for $\alpha_j = 1$

k	n							
	1	2	3	4	5	6	7	8
1	0.577	0.378	0.277	0.218	0.180	0.152	0.132	0.117
2	0.577	0.180	0.867e−1	0.512e−1	0.339e−1	0.241e−1	0.180e−1	0.140e−1
3	0.577	0.180	0.342e−1	0.136e−1	0.690e−2	0.401e−2	0.255e−2	0.172e−2
4	0.577	0.180	0.342e−1	0.468e−2	0.161e−2	0.730e−3	0.383e−3	0.222e−3
5	0.577	0.180	0.342e−1	0.468e−2	0.499e−3	0.155e−3	0.637e−4	0.307e−4
6	0.577	0.180	0.342e−1	0.468e−2	0.499e−3	0.438e−4	0.124e−4	0.473e−5
7	0.577	0.180	0.342e−1	0.468e−2	0.499e−3	0.438e−4	0.327e−5	0.860e−6
8	0.577	0.180	0.342e−1	0.468e−2	0.499e−3	0.438e−4	0.327e−5	0.212e−6

(a) The qualitative behavior of $d_n(V^{\alpha,k})$ is very similar to that of $d_n(H^{\alpha,k})$. (b) We have $d_n(H^{\alpha,k}) < d_n(V^{\alpha,k})$, but the difference between $d_n(V^{\alpha,k})$ and $d_n(H^{\alpha,k})$ is not large.

Table 4
The values of the n-width, $d_n(H^{\alpha,k}_{\text{odd}})$, for $\alpha_j = 1$

k	n							
	1	2	3	4	5	6	7	8
1	0.537	0.303	0.208	0.157	0.126	0.106	0.906e−1	0.793e−1
2	0.323	0.961e−1	0.440e−1	0.250e−1	0.161e−1	0.112e−1	0.824e−2	0.631e−2
3	0.202	0.306e−1	0.934e−2	0.398e−2	0.205e−2	0.119e−2	0.749e−3	0.502e−3
4	0.127	0.973e−2	0.198e−2	0.633e−3	0.261e−3	0.126e−3	0.681e−4	0.400e−4
5	0.808e−1	0.310e−2	0.421e−3	0.101e−3	0.332e−4	0.134e−4	0.620e−5	0.318e−5
6	0.514e−1	0.986e−3	0.892e−4	0.160e−4	0.423e−5	0.142e−5	0.563e−6	0.253e−6
7	0.327e−1	0.314e−3	0.189e−4	0.255e−5	0.538e−6	0.151e−6	0.512e−7	0.201e−7
8	0.208e−1	0.999e−4	0.402e−5	0.406e−6	0.685e−7	0.160e−7	0.466e−8	0.160e−8

The quantitative and qualitative character of this table is different from those of Tables 2 and 3. Here, for fixed n, we see that $d_n(H^{\alpha,k}_{\text{odd}}) \to 0$ as k increases.

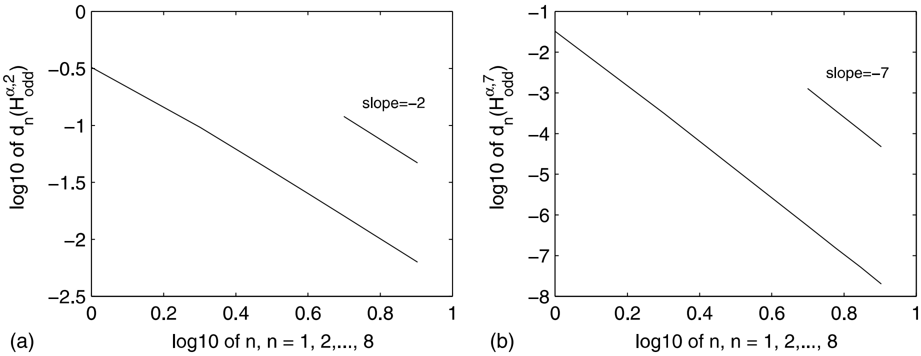

Fig. 3. The graphs of the n-width, $d_n(H^{\alpha,k}_{\mathrm{odd}})$, for $\alpha_j = 1$, and $k = 2$ and 7. The rate of convergence of $d_n(H^{\alpha,k}_{\mathrm{odd}})$, $\alpha_j = 1$ is $O(n^{-k})$. The decrease is not faster than exponential in the pre-asymptotic range.

Table 5
The values of the n-width, $d_n(H^{\alpha,k})$, for various values of α

k	n		
	2	5	8
(a) $d_n(H^{\alpha,k})$, $\alpha_j = j!$			
2	0.115	0.141e−1	0.508e−2
5	0.972e−1	0.396e−4	0.122e−5
8	0.972e−1	0.282e−4	0.909e−9
(b) $d_n(H^{\alpha,k})$, $\alpha_j = 1/j!$			
2	0.186	0.275e−1	0.101e−1
5	0.164	0.272e−2	0.133e−3
8	0.164	0.208e−2	0.167e−4
(c) $d_n(H^{\alpha,k})$, $\alpha_0 = 1$, $\alpha_k = 1$, $\alpha_j = 0$, *for* $j \neq 1, k$			
2	0.176	0.200e−1	0.720e−2
5	1.00	0.449e−3	0.134e−4
8	1.00	1.00	0.186e−6
(d) $d_n(H^{\alpha,k})$, $\alpha_j = \binom{k}{j}$, $j \leqslant k$			
2	0.133	0.195e−1	0.714e−2
5	0.439e−1	0.268e−3	0.122e−4
8	0.269e−1	0.365e−4	0.978e−7
(e) $d_n(H^{\alpha,k})$, $\alpha_j = 10^{2j}$			
2	0.178e−2	0.200e−3	0.720e−4
5	149e−2	0.449e−8	0.134e−9
8	149e−2	0.319e−8	0.186e−14
(f) $d_n(H^{\alpha,k})$, $\alpha_j = 1/10^{2j}$			
2	0.951	0.721	0.464
5	0.950	0.679	0.293
8	0.950	0.674	0.245

(a) The values of n-width depend significantly on α. (b) The qualitative features of these tables are similar, except in the case of the Table 5c where the n-width does not decrease as k increases; in fact, it can increase.

Table 5 reports values of $d_n(H^{\alpha,k})$ for the other choices for α introduced in Section 3.

We make the following observations based on Table 5 and also give the rationale for our choices of α's.

(a) From Theorem 5.11 and Remark 5.10, we have $d_n(H^{\alpha,k+1}) \leqslant d_n(H^{\alpha,k})$ for all choices of α except (3.7c), which corresponds to Table 5(c). That $d_n(H^{\alpha,k})$ is decreasing in k is clearly seen in all tables except Table 5(c); in this table we see that $d_n(H^{\alpha,k})$ can, in fact, increase in k. That $d_n(H^{\alpha,k}) = 1$ for $n < k$ in Table 5(c), was confirmed in Theorem 5.12. The reason for choosing the α given in (3.7c) was to show that $d_n(H^{\alpha,k})$ may not be decreasing (in fact may increase) in k for certain α.

(b) For all choices for α, Theorem 5.9 shows that $d_n(H^{\alpha,k}) = \mathrm{O}(n^{-k})$ for large n (in the asymptotic range). Whereas, for all values of α, the tables show that $d_n(H^{\alpha,k})$ decreases with n, we do not see the asymptotic rate because we have only used small values of n, i.e., we are in the pre-asymptotic range. Specifically, $d_n(H^{\alpha,k})$ hardly decreases with n in Table 5(f). We chose the α given in (3.7g) to show that the decrease in $d_n(H^{\alpha,k})$ with respect to n may be extremely slow for small n.

(c) If $\alpha_j \geqslant \beta_j$, then $H^{\alpha,k} \subset H^{\beta,k}$ and $\|u\|_{H^{\alpha,k}} \leqslant 1$, implies $\|u\|_{H^{\beta,k}} \leqslant 1$, so higher weights correspond to smoother functions and smaller weights correspond to less smooth functions. Also, by the definition of n-widths, we have $d_n(H^{\alpha,k}) \leqslant d_n(H^{\beta,k})$. So n-widths are smaller for spaces with smoother functions, as noted following Table 2. Now, the weights in Table 5(a), (d), and (e) are larger than those in Table 5(b), (c), and (f), respectively. So we should expect the n-widths in Table 5(a), (d), and (e) to be smaller than those in Table 5(b), (c), and (f), respectively. We readily observe this. We chose these values of α also to show the relationship between the smoothness of the functions in the underlying spaces and the magnitude of α, and the corresponding effect on the n-widths.

(d) Since $j! \gg 10^{2j}$ for large j (see Remark 3.4), we would expect $d_n(H^{\alpha,k})$ to be much smaller for $\alpha_j = j!$ than for $\alpha_j = 10^{2j}$, based on the arguments given in item (c). We see, however, just the opposite in Table 5(a) and (e). This is not a contradiction, because $j! \not\gg 10^{2j}$ for the values of j relevant to the values of k used in these tables (i.e., $j \leqslant k \leqslant 8$). We chose α as in (3.7d) and (3.7f) also to illuminate this apparent contradiction.

(e) Using the argument in item (c), we would expect $d_n(H^{\alpha,k})$ to be much smaller for $\alpha_j = 10^{-2j}$ than for $\alpha_j = 1/j!$. But comparing Table 5(b) and (f) we see just the opposite behavior. This is again not a contradiction, and can be explained by following the arguments in item (d). α as in (3.7e) and (3.7g) was chosen to illuminate this point.

(f) This principle does not strictly apply to Table 5(a) and (d) because the weights in Table 5(a) are not larger than those in Table 5(d) for all j. But they are larger for large j, and we see that the n-widths in Table 5(a) are smaller than those in Table 5(d) for $n > 2$, but the differences are not large. So the principle does hold in this extended sense.

(g) From these tables it is clear that $d_n(H^{\alpha,k})$ depends strongly on the choice of α.

We have computed $d_n(V^{\alpha,k})$ for the α's considered in Table 5, but have not included the values in this paper because of space limitations. We note, however, that the qualitative features of $d_n(V^{\alpha,k})$ are similar to those of $d_n(H^{\alpha,k})$ in Table 5, and the observations made after Table 5 are also valid for $d_n(V^{\alpha,k})$. The values of $d_n(V^{\alpha,k})$ for these α's are given in [2].

Table 6 presents results for $d_n(H^{\alpha,k}_{\mathrm{odd}})$. These values were computed using Theorem 5.3.

We know from Remarks 5.3 and 5.4 that $d_n(H^{\alpha,k}_{\mathrm{odd}}) \to 0$ as $k \to \infty$ for all choices of α in Section 3 except $\alpha_j = 1/j!$ and $\alpha_j = 10^{-2j}$ (for $n \leqslant 6$). The values in Table 6 illustrate this. The values in Table 6(f) for $n = 8$ do not show this feature (i.e., $d_n(H^{\alpha,k}_{\mathrm{odd}}) \to 0$ as $k \to \infty$) clearly because, while the series in (5.14) in Theorem 5.6 diverges, the coefficients $\alpha_j = 10^{-2j}$ are very small and $\sum_{j=0}^{k} \alpha_j (n^{2j}\pi^{2j}/2^{2j})$ grows slowly as k increases. Also from Table 6(b), we see that the values of $d_n(H^{\alpha,k}_{\mathrm{odd}})$ are decreasing with k, but they do not indicate that $\lim_{k\to\infty} d_n(H^{\alpha,k}_{\mathrm{odd}}) \geqslant C(n) > 0$. We note that one has to take large values of k to see this effect. Thus, not all the observations made in item (a) following the Table 5 are valid for the values in Table 6. However, the observations made in items (b) through (g) are valid for the values in Table 6.

Table 6
The values of the n-width, $d_n(H^{\alpha,k}_{\mathrm{odd}})$, for various values of α

k	n		
	2	5	8
(a) $d_n(H^{\alpha,k}_{odd})$, $\alpha_j = j!$			
2	0.697e−1	0.114e−1	0.447e−2
5	0.295e−3	0.305e−5	0.291e−6
8	0.522e−6	0.344e−9	0.801e−11
(b) $d_n(H^{\alpha,k}_{odd})$, $\alpha_j = 1/j!$			
2	0.130	0.226e−1	0.890e−2
5	0.268e−1	0.352e−3	0.344e−4
8	0.122e−1	0.130e−4	0.315e−6
(c) $d_n(H^{\alpha,k}_{odd})$, $\alpha_0 = 1$, $\alpha_k = 1$, $\alpha_j = 0$, $j \neq 1,k$			
2	0.101	0.162e−1	0.633e−2
5	0.327e−2	0.335e−4	0.319e−5
8	0.105e−3	0.691e−7	0.161e−8
(d) $d_n(H^{\alpha,k}_{odd})$, $\alpha_j = \binom{k}{j}$, $j \leqslant k$			
2	0.920e−1	0.160e−1	0.629e−2
5	0.257e−2	0.321e−4	0.314e−5
8	0.716e−4	0.648e−7	0.157e−8
(e) $d_n(H^{\alpha,k}_{odd})$, $\alpha_j = 10^{2j}$			
2	0.101e−2	0.162e−3	0.633e−4
5	0.327e−7	0.335e−9	0.319e−10
8	0.105e−11	0.691e−15	0.161e−16
(f) $d_n(H^{\alpha,k}_{odd})$, $\alpha_j = 1/10^{2j}$			
2	0.950	0.706	0.444
5	0.949	0.637	0.200
8	0.949	0.623	0.982e−1

In all the tables, except Table 6(f), $d_n(H^{\alpha,k}_{\mathrm{odd}}) \to 0$, as k increases.

So far we have addressed the n-widths, i.e., the performance of the best possible shape functions. In Table 7 we report the values of the sup–inf, $\Psi(\mathscr{P}, n, H^{\alpha,k})$, for $\alpha_j = 1$, which indicate the approximability of polynomials in the spaces $H^{\alpha,k}$. We refer to [2] for the computation of Ψ.

Comparing Tables 7 and 2, we see that their qualitative and quantitative features are very similar. We observe that $d_n(H^{\alpha,k}) \leqslant \Psi(\mathscr{P}, n, H^{\alpha,k})$, which was indicated in (5.37) of Theorem 5.15 with $V = H^{\alpha,k}$, $H = L_2$ and $\Phi = \mathscr{P}$. We also observe that $\Psi(\mathscr{P}, n, H^{\alpha,k+1}) \leqslant \Psi(\mathscr{P}, n, H^{\alpha,k})$, which was indicated by (5.49) of Theorem 5.21. Fig. 4(a) and (b) are log–log plots of $\Psi(\mathscr{P}, n, H^{\alpha,k})$ versus n for $k = 2$ and 7, respectively. For $k = 2$, we see that $\Psi(\mathscr{P}, n, H^{\alpha,2})$ is in the asymptotic range and the rate of convergence is $\mathrm{O}(n^{-2})$, which was indicated by Theorem 5.19. Fig. 4(b) reveals that the rate of decrease of $\Psi(\mathscr{P}, n, H^{\alpha,7})$ is faster than exponential in the pre-asymptotic range, as suggested in Remark 5.14 (see also Remark 5.9). In this case we also have a typical S-shaped curve. We note that $\Psi(\mathscr{P}, n, H^{\alpha,7})$ has not yet reached the asymptotic range for $n \leqslant 8$. The features in Fig. 4 are similar to those in Fig. 2.

Table 8 reports the values of $\Psi(P, n, H^{\alpha,k})$ for various values of α.

Comparison of Tables 8 and 5 shows that $d_n(H^{\alpha,k}) \leqslant \Psi(\mathscr{P}, n, H^{\alpha,k})$, as expected for all our choices of α. But the values of $d_n(H^{\alpha,k})$ and $\Psi(\mathscr{P}, n, H^{\alpha,k})$ are quite similar. We note that the observations made after Table 5 are also valid for the values in Table 8.

Tables 2, 3, 5, 7 and 8 show that the values of $d_n(V)$ and $\Psi(\mathscr{P}, n, V)$ depend strongly on the spaces $V = H^{\alpha,k}$ or $V^{\alpha,k}$. Therefore, the performance of a family of shape functions Φ should be assessed relative to

Table 7
The values of the sup–inf, $\Psi(\mathscr{P}, n, H^{\alpha,k})$, for $\alpha_j = 1$

k	n							
	1	2	3	4	5	6	7	8
1	0.537	0.303	0.215	0.170	0.141	0.121	0.106	0.946e−1
2	0.503	0.150	0.617e−1	0.349e−1	0.230e−1	0.165e−1	0.124e−1	0.973e−2
3	0.501	0.130	0.272e−1	0.911e−2	0.438e−2	0.251e−2	0.159e−2	0.108e−2
4	0.501	0.129	0.222e−1	0.365e−2	0.105e−2	0.444e−3	0.228e−3	0.131e−3
5	0.501	0.128	0.218e−1	0.283e−2	0.386e−3	0.985e−4	0.377e−4	0.176e−4
6	0.501	0.128	0.217e−1	0.275e−2	0.286e−3	0.337e−4	0.781e−5	0.274e−5
7	0.501	0.128	0.217e−1	0.274e−2	0.276e−3	0.241e−4	0.251e−5	0.535e−6
8	0.501	0.128	0.217e−1	0.274e−2	0.276e−3	0.231e−4	0.173e−5	0.163e−6

(a) The qualitative and quantitative features of this table are very similar to those of the Table 2. (b) We note that $d_n(H^{\alpha,k}) \leqslant \Psi(\mathscr{P}, n, H^{\alpha,k})$. (c) Also $\Psi(\mathscr{P}, n, H^{\alpha,k+1}) \leqslant \Psi(\mathscr{P}, n, H^{\alpha,k})$.

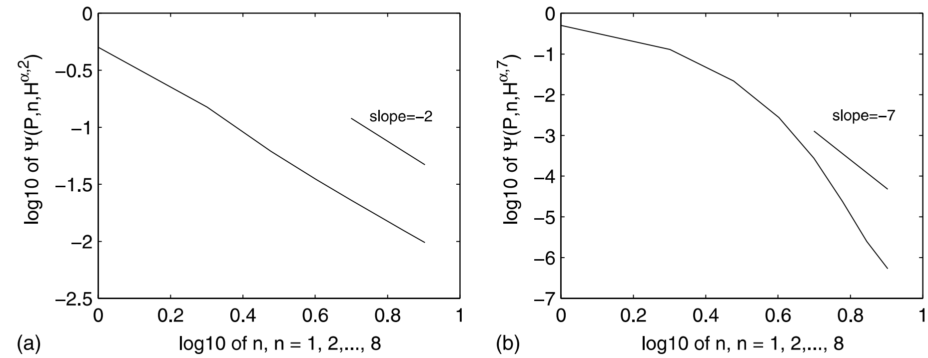

Fig. 4. The graphs of the sup–inf, $\Psi(\mathscr{P}, n, H^{\alpha,k})$, for $\alpha_j = 1$, and $k = 2$ and 7. (a) We see that the rate of convergence of $\Psi(\mathscr{P}, n, H^{\alpha,2})$ with respect to n is $O(n^{-2})$ and (b) we see that the rate of decrease of $\Psi(\mathscr{P}, n, H^{\alpha,7})$ is higher than exponential in the pre-asymptotic range, and that, $\Psi(\mathscr{P}, n, H^{\alpha,7})$ has yet not reached the asymptotic range.

the performance of the optimal shape functions for V. Hence $\Lambda(\mathscr{P}, n, V) = \Psi(\mathscr{P}, n, V)/d_n(H^{\alpha,k})$, introduced in Section 4, for $V = H^{\alpha,k}$ and $V^{\alpha,k}$, are effective measurements of the performance of polynomials in the spaces $H^{\alpha,k}$ and $V^{\alpha,k}$, respectively. Since the polynomials are optimal shape functions for the spaces $V^{\alpha,k}$, as shown in Theorem 5.8, we have

$$\Lambda(\mathscr{P}, n, V^{\alpha,k}) = 1.$$

We know from Theorem 5.23 that for $0 < \sigma \leqslant 1$, $\Lambda_\sigma(\mathscr{P}, n, H^{\alpha,k})$ is "almost bounded" in k and n provided $\alpha_j \leqslant \alpha_{j+1}$. But we note that $\Lambda_\sigma(\mathscr{P}, n, H^{\alpha,k}) = \Lambda(\mathscr{P}, n, H^{\alpha,k})$ when $\sigma = 0$. From Tables 5 and 8, we see that $\Lambda(\mathscr{P}, n, H^{\alpha,k})$ is reasonably bounded for all our choices of α. This shows that polynomials are robust shape functions for the spaces $H^{\alpha,k}$.

We now consider the performance of the trigonometric polynomials, i.e., we consider the basis $\mathscr{C} = \{\cos(i\pi(x+1)/2), i = 0, 1, \ldots\}$ in the space $H^{\alpha,k}$. Table 9 reports the values of $\Psi(\mathscr{C}, n, H^{\alpha,k})$ for $\alpha_j = 1$.

The values in Table 9 are quite different from those in Table 7. We see that for fixed n, $\Psi(\mathscr{C}, n, H^{\alpha,k})$ is nearly independent of k for $k > 3$, a feature not observed in Table 7. In Fig. 5(a) and (b), we have plotted $\Psi(\mathscr{C}, n, H^{\alpha,k})$ versus n for $k = 1$ and 5. For $k = 1$ we see from Fig. 5(a) that the order of convergence of

Table 8
The values of the sup–inf, $\Psi(\mathscr{P}, n, H^{\alpha,k})$, for various values α

k	n		
	2	5	8
(a) $\Psi(\mathscr{P}, n, H^{\alpha,k})$, $\alpha_j = j!$			
2	0.115	0.164e−1	0.690e−2
5	0.972e−1	0.397e−4	0.162e−5
8	0.972e−1	0.282e−4	0.909e−9
(b) $\Psi(\mathscr{P}, n, H^{\alpha,k})$, $\alpha_j = 1/j!$			
2	0.186	0.320e−1	0.137e−1
5	0.164	0.285e−2	0.183e−3
8	0.164	0.216e−2	0.192e−4
(c) $\Psi(\mathscr{P}, n, H^{\alpha,k})$, $\alpha_0 = 1$, $\alpha_k = 1$, $\alpha_j = 0$ for $j \neq 1, k$			
2	0.176	0.234e−1	0.979e−2
5	1.00	0.449e−3	0.178e−4
8	1.00	1.00	0.186e−6
(d) $\Psi(\mathscr{P}, n, H^{\alpha,k})$, $\alpha_j = \binom{k}{j}$, $j \leqslant k$			
2	0.133	0.227e−1	0.967e−2
5	0.439e−1	0.275e−3	0.168e−4
8	0.269e−1	0.366e−4	0.105e−6
(e) $\Psi(\mathscr{P}, n, H^{\alpha,k})$, $\alpha_j = 10^{2j}$			
2	0.178e−2	0.234e−3	0.979e−4
5	0.149e−2	0.449e−8	0.178e−9
8	0.149e−2	0.319e−8	0.186e−14
(f) $\Psi(\mathscr{P}, n, H^{\alpha,k})$, $\alpha_j = 1/10^{2j}$			
2	0.951	0.744	0.540
5	0.950	0.707	0.387
8	0.950	0.703	0.343

(a) The values of $\Psi(\mathscr{P}, n, H^{\alpha,k})$ are sensitive to the choice of spaces. (b) The values of $\Psi(\mathscr{P}, n, H^{\alpha,k})$ are very similar to $d_n(H^{\alpha,k})$. (c) We see that $d_n(H^{\alpha,k}) \leqslant \Psi(\mathscr{P}, n, H^{\alpha,k})$ for all α. (d) In all the cases, except the case (c), we have $\Psi(\mathscr{P}, n, H^{\alpha,k+1}) \leqslant \Psi(\mathscr{P}, n, H^{\alpha,k})$.

Table 9
The values of the sup–inf, $\Psi(\mathscr{C}, n, H^{\alpha,k})$, for $\alpha_j = 1$

k	n							
	1	2	3	4	5	6	7	8
1	0.537	0.303	0.208	0.157	0.126	0.106	0.906e−1	0.793e−1
2	0.503	0.150	0.811e−1	0.418e−1	0.319e−1	0.199e−1	0.175e−1	0.119e−1
3	0.501	0.130	0.646e−1	0.355e−1	0.256e−1	0.173e−1	0.143e−1	0.106e−1
4	0.501	0.129	0.631e−1	0.353e−1	0.253e−1	0.173e−1	0.142e−1	0.106e−1
5	0.501	0.128	0.629e−1	0.353e−1	0.253e−1	0.173e−1	0.142e−1	0.106e−1
6	0.501	0.128	0.629e−1	0.353e−1	0.253e−1	0.173e−1	0.142e−1	0.106e−1
7	0.501	0.128	0.629e−1	0.353e−1	0.253e−1	0.173e−1	0.142e−1	0.106e−1
8	0.501	0.128	0.629e−1	0.353e−1	0.253e−1	0.173e−1	0.142e−1	0.106e−1

(a) For large n and k, the values in Tables 7 and 9 are quite different. (b) For fixed n, the values $\Psi(\mathscr{C}, n, H^{\alpha,k})$ are nearly independent of k for $k > 3$. (c) The basis $\mathscr{C}$ performs worse than the basis $\mathscr{P}$ for n, $k \geqslant 3$, and much worse for n, k large, but for $k = 1$, $\mathscr{C}$ performs better than $\mathscr{P}$.

$\Psi(\mathscr{C}, n, H^{\alpha,1})$ is $O(n^{-1})$, whereas, for $k = 5$, the order of convergence of $\Psi(\mathscr{C}, n, H^{\alpha,5})$ is $O(n^{-1.5})$ as shown in Fig. 5(b). These orders of convergence were confirmed in Theorem 5.20. Moreover, we note that these plots

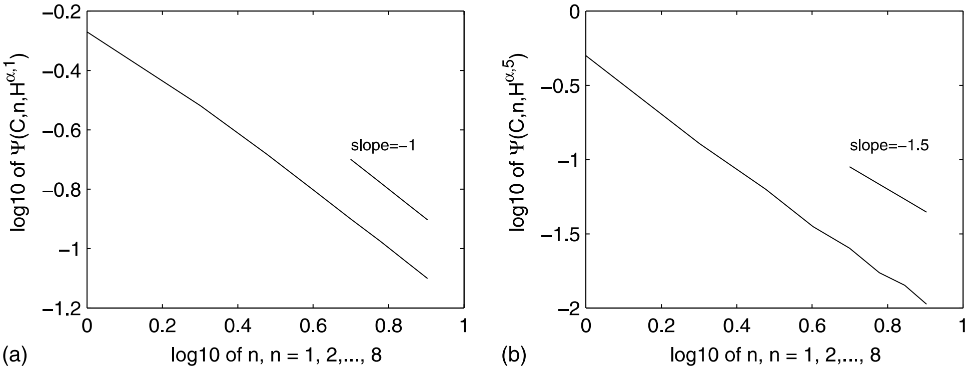

Fig. 5. The graphs of the sup–inf, $\Psi(\mathscr{C}, n, H^{\alpha,k})$, for $\alpha_j = 1$, and $k = 1$ and $k = 5$. (a) We see that the rate of convergence of $\Psi(\mathscr{P}, n, H^{\alpha,1})$ with respect to n is $O(n^{-1})$ and (b) we see that the rate of decrease of $\Psi(\mathscr{C}, n, H^{\alpha,5})$ is $O(n^{-1.5})$.

do not have the S-shape that we saw in Fig. 4(b), which means that the rate of decrease of $\Psi(\mathscr{C}, n, H^{\alpha,k})$ is not faster than exponential in the pre-asymptotic range for a fixed k.

We also see, comparing Tables 7 and 9, that the basis $\mathscr{C}$ performs worse than the basis $\mathscr{P}$ for n, $k \geqslant 3$, and much worse for n, k large. In fact, using Theorems 5.19 and 5.20, one can show that $\kappa(\mathscr{C}, \mathscr{P}, n, H^{\alpha,k}) \to \infty$ as $n \to \infty$ for our choices of α and for $k \geqslant 2$, where κ was introduced in Section 4. Therefore, based on the recommendation in Section 4, the basis $\mathscr{P}$ should be preferred over the basis $\mathscr{C}$ for $k \geqslant 2$ when $\alpha_j = 1$. For $k = 1$, we have $\kappa(\mathscr{C}, \mathscr{P}, n, H^{\alpha,k}) \leqslant 1$, because $\mathscr{C}$ is the optimal basis for $k = 1$. Also we see from Tables 7 and 9 that $\kappa \approx 0.84$ (close to 1). Therefore we have no reason to prefer the basis $\mathscr{P}$ over the basis $\mathscr{C}$ or vice versa in the pre-asymptotic range (e.g., $n \leqslant 8$) when $k = 1$. We note however that κ, for $k = 1$, will become smaller for large n, and we should prefer the basis $\mathscr{C}$ over the basis $\mathscr{P}$ in this case.

In the Table 10 we report the values of $\Psi(\mathscr{C}, n, H^{\alpha,k})$ for various values of α.

Comparison of Tables 8 and 10 shows that algebraic polynomials perform slightly better than trigonometric polynomials when k is small ($k = 2$), and algebraic polynomials perform much better when k is moderately large ($k = 5$ or 8)—except when $\alpha_j = 1/10^{2j}$. In this case we see that $\Psi(\mathscr{P}, n, H^{\alpha,k}) \approx \Psi(\mathscr{C}, n, H^{\alpha,k})$ for $k = 2, 5, 8$. Furthermore, we note that for all other α, $\Psi(\mathscr{P}, n, H^{\alpha,k})$ sharply decreases with increasing k, but $\Psi(\mathscr{P}, n, H^{\alpha,k})$ is nearly constant in k when $\alpha_j = 1/10^{2j}$. This difference in behavior of $\Psi(\mathscr{P}, n, H^{\alpha,k})$ can be explained as follows. $\Psi(\Phi, n, V)$ measures the approximability of Φ for all functions in the V-unit ball. For $V = H^{\alpha,k}$ with α_j very small, as when $\alpha_j = 1/10^{2j}$, functions in the unit ball may have very large derivatives, and hence be unsmooth. Thus the values of $\Psi(\mathscr{P}, n, H^{\alpha,k})$ for moderate sized k are about the same as $\Psi(\mathscr{P}, n, H^{\alpha,2})$, explaining why $\Psi(\mathscr{P}, n, H^{\alpha,k})$ is nearly constant in k for small to moderate sized k, for this α. One might expect the same for the case $\alpha_j = 1/j!$. But $1/j!$ is not nearly as small as $1/10^{2j}$ for moderate sized j ($j \leqslant k \leqslant 8$), and this effect does not so clearly show up in the values in Table 8(b).

We now consider the performance of polynomials, i.e., $\Phi = \mathscr{P}$, in the spaces $H^{\alpha,k}_{\text{odd}}$, where $\alpha_j = 1$. We note that the functions in these spaces are restricted at the boundary, and $\Phi = \mathscr{C}$ is the optimal basis for these spaces, as indicated in Theorem 5.3. In Table 11 we report the values of $\Psi(\mathscr{P}, n, H^{\alpha,k}_{\text{odd}})$ for $\alpha_j = 1$.

We observe from Table 11 that for fixed n, $\Psi(\mathscr{P}, n, H^{\alpha,k}_{\text{odd}})$ decreases in k. As expected, comparing the Tables 11 and 4, we see that the values in Table 11 are greater than the corresponding values in Table 4 for all n, k. Also, by a simple calculation, one sees that $\Lambda(\mathscr{P}, n, H^{\alpha,k}_{\text{odd}})$ increases in k for fixed n, as indicated by Theorem 5.24. A careful calculation of $\Lambda(\mathscr{P}, n, H^{\alpha,k}_{\text{odd}})$ (which we have not included here) shows that for fixed k, $\Lambda(\mathscr{P}, n, H^{\alpha,k}_{\text{odd}})$ first increases, then decreases, and again starts to increase slowly with increasing odd values of n. The same is true also for n even. This shows that the polynomials do not perform well in $H^{\alpha,k}_{\text{odd}}$ in the

Table 10
(a–f) The values of the sup–inf, $\Psi(\mathscr{C}, n, H^{\alpha,k})$, for various values of α

k	n		
	2	5	8
(a) $\Psi(\mathscr{C}, n, H^{\alpha,k})$, $\alpha_j = j!$			
2	0.115	0.283e−1	102e−1
5	0.972e−1	0.242e−1	0.889e−2
8	0.972e−1	0.242e−1	0.889e−2
(b) $\Psi(\mathscr{C}, n, H^{\alpha,k})$, $\alpha_j = 1/j!$			
2	0.186	0.377e−1	0.155e−1
5	0.164	0.287e−1	0.134e−1
8	0.164	0.287e−1	0.134e−1
(c) $\Psi(\mathscr{C}, n, H^{\alpha,k})$, $\alpha_0 = 1$, $\alpha_k = 1$, $\alpha_j = 0$, *for* $j \neq 1, k$			
2	0.176	0.535e−1	0.194e−1
5	0.999	0.565	0.362
8	0.999	0.999	0.848
(d) $\Psi(\mathscr{C}, n, H^{\alpha,k})$, $\alpha_j = \binom{k}{j}$, $j \leqslant k$			
2	0.133	0.272e−1	0.110e−1
5	0.438e−1	0.122e−1	0.449e−2
8	0.269e−1	0.965e−2	0.354e−2
(e) $\Psi(\mathscr{C}, n, H^{\alpha,k})$, $\alpha_j = 10^{2j}$			
2	0.695e−2	0.277e−2	0.102e−2
5	0.693e−2	0.276e−2	0.101e−2
8	0.693e−2	0.276e−2	0.101e−2
(f) $\Psi(\mathscr{C}, n, H^{\alpha,k})$, $\alpha_j = 1/10^{2j}$			
2	0.951	0.722	0.468
5	0.950	0.680	0.334
8	0.950	0.676	0.310

Comparison of this table with Table 8 shows that the trigonometric polynomials perform much worse than the algebraic polynomials except in the case $\alpha_j = 1/10^{2j}$, where trigonometric polynomials perform slightly better.

Table 11
The values of the sup–inf, $\Psi(\mathscr{P}, n, H^{\alpha,k}_{\mathrm{odd}})$, for $\alpha_j = 1$

k	n							
	1	2	3	4	5	6	7	8
1	0.537	0.303	0.215	0.170	0.141	0.121	0.106	0.946e−1
2	0.323	0.961e−1	0.574e−1	0.348e−1	0.230e−1	0.165e−1	0.124e−1	0.973e−2
3	0.202	0.306e−1	0.257e−1	0.911e−2	0.437e−2	0.251e−2	0.159e−2	0.108e−2
4	0.127	0.154e−1	0.154e−1	0.274e−2	0.100e−2	0.443e−3	0.228e−3	0.131e−3
5	0.808e−1	0.973e−2	0.973e−2	0.859e−3	0.364e−3	0.984e−4	0.376e−4	0.176e−4
6	0.514e−1	0.618e−2	0.618e−2	0.272e−3	0.205e−3	0.271e−4	0.757e−5	0.274e−5
7	0.327e−1	0.393e−2	0.393e−2	0.128e−3	0.128e−3	0.827e−5	0.238e−5	0.535e−6
8	0.208e−1	0.250e−2	0.250e−2	0.816e−4	0.816e−4	0.260e−5	0.127e−5	0.136e−6

(a) For a fixed n, $\Psi(\mathscr{P}, n, H^{\alpha,k}_{\mathrm{odd}})$ decreases as k increases. (b) For large n and k, the values in Table 11 are much bigger than the corresponding values in Table 4. (c) The system $\mathscr{P}$ performs worse than the system $\mathscr{C}$, which is optimal, for large values of n and k.

pre-asymptotic region, as indicated in Remark 5.22. The slow increase of $\Lambda(\mathscr{P}, n, H^{\alpha,k}_{\mathrm{odd}})$ with respect to n also indicates the necessity of the assumption $\sigma > 0$ in Theorem 5.25.

We did not discuss the performance of the basis $\Phi = \mathscr{S}$ in the space $H^{\alpha,k}_{\mathrm{odd}}$ as its performance is very similar to $\Phi = \mathscr{C}$ in the space $H^{\alpha,k}_{\mathrm{odd}}$; see Remark 5.5. Also, we have not discussed the performance of the basis $\Phi = \mathscr{S}$ in the space $H^{\alpha,k}$.

From the Tables 2–11 we can draw the following conclusions:

(a) Algebraic polynomials perform almost as well as the optimal shape functions in the spaces $H^{\alpha,k}$ and hence they are robust.
(b) Trigonometric polynomials perform much worse than algebraic polynomials if the approximated function belongs to $H^{\alpha,k}$, with $k \geqslant 2$. For $0 < k \leqslant 1$ (more precisely for $0 < k \leqslant 1.5$), trigonometric polynomials perform marginally better than algebraic polynomials. In fact, trigonometric polynomials are the optimal shape functions when $k = 1$, and hence they are comparable with polynomials only when $k = 1$, and are not comparable for $k > 1$.
(c) Polynomials do not perform as well as the optimal shape functions in the spaces $H^{\alpha,k}_{\mathrm{odd}}$ in the pre-asymptotic range.
(d) We note that the spaces $H^{\alpha,k}$ are uniform in x. The above conclusions are not necessarily correct when the underlying spaces are not uniform in x, e.g., in spaces with norms $\left(\sum_{j=0}^{k} \alpha_j \|\rho_j(x) u^{(j)}\|_0^2\right)^{1/2}$, where the weight functions $\rho_j(x)$ depend on x. We note, however, that the conclusions are valid for the spaces $V^{\alpha,k}$ (with Jacobi weights), which are not uniform in x.

Finally, we return to the functions $w_{\beta,\gamma}$ introduced in (1.11). As indicated in (4.3) we can estimate E_n^P (defined in (1.10)) using $\Psi(\mathscr{P}, n, V, H)$, with $H = L_2$ and $V = H^{\alpha,k}$ or $V^{\alpha,k}$ with $\alpha_j = 1$:

$$E_n^P(w_{\beta,\gamma}) \leqslant \Psi(\mathscr{P}, n+1, H^{\alpha,k}) \frac{\|w'_{\beta,\gamma}\|_{H^{\alpha,k}}}{\|w'_{\beta,\gamma}\|_{L_2}} \tag{6.1}$$

and

$$E_n^P(w_{\beta,\gamma}) \leqslant \Psi(\mathscr{P}, n+1, V^{\alpha,k}) \frac{\|w'_{\beta,\gamma}\|_{V^{\alpha,k}}}{\|w'_{\beta,\gamma}\|_{L_2}}, \tag{6.2}$$

for $n = 1, 2, \ldots$ Here n is the degree of the polynomial approximating $w'_{\beta,\gamma}$, and $n+1$ is the dimension of the space of polynomials of degree $\leqslant n$. Now the expression on the right-hand side of (6.1) can be calculated from of Tables 1 and 7; and the right-hand side of (6.2) can be evaluated using Tables 1 and 3 (note that $\Psi(\mathscr{P}, n+1, V^{\alpha,k}) = d_{n+1}(V^{\alpha,k})$). These values are reported in Table 12.

The values in Table 12 reflect various function space inclusions for $w'_{\beta,\gamma}$. Specifically, $w'_{\beta,\gamma}$ can be viewed as belonging to $H^{\alpha,k}$ or $V^{\alpha,k}$ for various k with $\alpha_j = 1$, each inclusion providing an estimate for $E_n^P(w_{\beta,\gamma})$. The values in Table 12 should be compared with the values for $E_n^P(w_{\beta,\gamma})$ displayed in Fig. 1. The numerical values of $E_n^P(w_{\beta,\gamma})$ are given in [2]. These comparisons show that the estimates are very pessimistic, although in a certain sense they are optimal (see Remark 5.13). They are pessimistic because there are other functions in the spaces $H^{\alpha,k}$ or $V^{\alpha,k}$ with the same norm as $w_{\beta,\gamma}$, but which are more poorly approximated by $\mathscr{P}$. This shows that an error estimator that uses the value of a particular norm as the only available information on the approximated function can be very pessimistic. Nevertheless, such estimates provide the correct rate for approximation errors.

Furthermore, from these tables we see that for a given n, there is an optimal space—$H^{\alpha,k}$ with a certain k or $V^{\alpha,k}$ with a certain k, with $\alpha_j = 1$—which leads to the best available estimate. Usually the magnitude of higher order derivatives are larger than that of the lower order derivatives. Hence a priori error estimates for low n should be based on spaces where the size of the higher derivatives is not taken into account, i.e., we should choose functions space inclusions with low k.

Table 12
The estimates of $E_n^P(w_{\beta,\gamma})$ for $\beta = 1$, $\gamma = 5$ and $\beta = 7$, $\gamma = 5$ using $\Psi(\mathscr{P}, n+1, H^{\alpha,k})$ and $\Psi(\mathscr{P}, n+1, V^{\alpha,k})$ with $\alpha_j = 1$

k	n						
	1	2	3	4	5	6	7
(a) Estimate of $E_n^P(w_{\beta,\gamma})$ for $\beta = 1$, $\gamma = 5$ using $\Psi(\mathscr{P}, n+1, H^{\alpha,k})$, $\alpha_j = 1$							
1	2.42	1.72	1.36	1.13	0.97	0.85	0.76
2	5.89	2.42	1.37	0.90	0.65	0.49	0.38
3	17.2	3.59	1.20	0.58	0.33	0.21	0.14
4	40.4	6.96	1.14	0.33	0.14	0.072	0.041
(b) Estimate of $E_n^P(w_{\beta,\gamma})$ for $\beta = 7$, $\gamma = 5$ for $\Psi(\mathscr{P}, n+1, H^{\alpha,k})$, $\alpha_j = 1$							
1	1.84	1.31	1.03	0.86	0.73	0.64	0.57
2	6.90	2.84	1.61	1.06	0.76	0.57	0.45
3	54.8	11.5	3.84	1.85	1.06	0.67	0.46
4	693.1	119.3	19.61	5.64	2.39	1.22	0.70
(c) Estimate of $E_n^P(w_{\beta,\gamma})$ for $\beta = 1$, $\gamma = 5$ for $\Psi(\mathscr{P}, n+1, V^{\alpha,k})$, $\alpha_j = 1$							
1	1.35	0.97	0.76	0.63	0.53	0.46	0.41
2	2.19	1.06	0.62	0.41	0.29	0.22	0.17
3	7.43	1.41	0.56	0.28	0.17	0.11	0.071
4	22.7	4.32	0.59	0.20	0.092	0.048	0.028
(d) Estimate of $E_n^P(w_{\beta,\gamma})$ for $\beta = 7$, $\gamma = 5$ using $\Psi(\mathscr{P}, n+1, V^{\alpha,k})$, $\alpha_j = 1$							
1	2.22	1.63	1.28	1.06	0.89	0.77	0.69
2	7.75	3.73	2.20	1.46	1.04	0.77	0.60
3	67.6	12.8	5.11	2.59	1.51	0.96	0.65
4	740.4	140.6	19.3	6.62	3.00	1.58	0.91

Note the large overestimates for many of the function space inclusions for $w_{\beta,\gamma}$.

7. Conclusions

This is the first in a series of papers addressing the problem of selection of shape functions for the GFEM. We focused on the principles that should govern the selection of effective shape functions, and elaborated in detail the one dimensional case.

Shape functions should have good approximation properties relative to the available information on the approximated function, typically the unknown solution of a boundary value problem. The available information on the approximated function is necessarily *fuzzy*. It is usually characterized by inclusion in a family $\mathscr{V}$ of function spaces V. The function spaces V could be, e.g., $H^{\alpha,k}$, $H^{\alpha,k}_{\mathrm{even}}$, $H^{\alpha,k}_{\mathrm{odd}}$, or $V^{\alpha,k}$—the spaces introduced in Section 3. The smoothness of the functions in the spaces $H^{\alpha,k}$, $H^{\alpha,k}_{\mathrm{even}}$, or $H^{\alpha,k}_{\mathrm{odd}}$ is characterized uniformly in the space variable x, i.e., the definitions of the norms in these spaces use only constant weight functions. We note, however, that the characterization of smoothness may be non-uniform in x, as with $V^{\alpha,k}$. We also note that the available information on the approximated function may include its boundary conditions, as with $H^{\alpha,k}_{\mathrm{even}}$ or $H^{\alpha,k}_{\mathrm{odd}}$.

As pointed out in Section 2, we can view the approximation in L_2. Thus the spaces V we consider are subspaces of L_2. The shape functions (basis) $\Phi \subset L_2$, selected to approximate the unknown functions, should have good approximation properties relative to the entire family $\mathscr{V}$, i.e., they should be robust in $\mathscr{V}$. A good measure of the effectiveness of Φ in V is given by $\Lambda(\Phi, n, V)$, introduced in Section 4, which is the ratio of the worst possible approximation error in L_2 of functions in the unit ball in V using Φ, and the same error using the optimal shape functions, $\Phi(V) = \widetilde{X}_n(V, H)$, in the sense of n-widths. Then the sizes of $\Lambda(\Phi, n, V)$, for all $V \in \mathscr{V}$, is a measure of the robustness of Φ relative to $\mathscr{V}$.

Using these ideas, we have shown:

- If the only available information on the unknown function is related to smoothness characterized uniformly in x by inclusion in the class $\mathscr{V}$ of Sobolev-type spaces, $H^{\alpha,k}$, then polynomial shape functions are robust, and they perform roughly as well as the optimal shape functions, in the sense of n-widths, as determined by the spaces $V = H^{\alpha,k}$. Hence polynomial shape functions are recommended for approximation of such functions.
- If, on the other hand, some additional information is available, if, e.g., the function is constrained by certain boundary conditions, then polynomial shape functions may perform—in the sense of robustness—very poorly in the pre-asymptotic range, and some other shape functions should be used. We have shown this for $V = H^{\alpha,k}_{\text{odd}}$.

We have also observed that polynomial shape functions are optimal in the sense of n-widths when the approximated functions are in spaces $V^{\alpha,k}$. The characterization of smoothness of the functions in this space is non-uniform in x. We note that we have not studied any other class of function spaces with this feature.

These recommendations are based on the rigorous theory presented in Section 5. We have also presented detailed numerical computations in Section 6. These computations illustrate our theoretical results; and our theoretical results explain the observed features of our numerical computations.

In future papers the framework of this paper will be used to explore the problem of selection of shape functions in higher dimensions.

References

[1] S.N. Atluri, T. Zhu, A new meshless local Petrov–Galerkin (MLPG) approach in computational mechanics, Comp. Mech. 22 (1998) 117–127.
[2] I. Babuška, U. Banerjee, J. Osborn, On principles for the selection of shape functions for the generalized finite element method, TICAM Report 01-16. University of Texas at Austin, Austin, TX, June 2001.
[3] I. Babuška, G. Caloz, J. Osborn, Special finite element methods for a class of second order elliptic problems with rough coefficients, SIAM J. Numer. Anal. 31 (1994) 945–981.
[4] I. Babuška, B. Guo, Direct and inverse approximation theorems in the p-version of finite element method in the framework of weighted Besov spaces, Part 1: approximability of functions in weighted Besov spaces, TICAM Report 99-31, SIAM J. Numer. Anal., in press.
[5] I. Babuška, B. Guo, Optimal estimates on lower and upper bound of approximation error of the p-version of the finite element method in two dimensions, Numer. Math. 85 (2000) 219–255.
[6] I. Babuška, J.M. Melenk, The partition of unity finite element method, Int. J. Numer. Meth. Engrg. 40 (1997) 727–758.
[7] I. Babuška, H.-S. Oh, The p-version of finite element method for domains with corners and for infinite domains, Numer. Meth. Par. Diff. Eqns. 6 (1990) 371–392.
[8] I. Babuška, J. Osborn, Can a finite element method perform arbitrarily badly? Math. Comp. 69 (1999) 443–462.
[9] I. Babuška, J. Osborn, Eigenvalue problems, in: P.G. Ciarlet, J.C. Lions (Eds.), Handbook of Numerical Analysis, vol. 2, North-Holland, Amsterdam, 1991, pp. 643–787.
[10] R.S. Barsoum, On the use of isoparametric finite elements in linear fracture mechanics, Int. J. Numer. Meth. Engrg. 10 (1976) 25–37.
[11] T. Belytschko, Y.Y. Lu, L. Gu, Element-free Galerkin methods, Int. J. Numer. Meth. Engrg. 37 (1994) 229–256.
[12] C.A. Duarte, I. Babuška, J.T. Oden, Generalized finite element methods for three dimensional structural mechanics problems, Comput. Struct. 77 (2000) 215–232.
[13] C.A. Duarte, J.T. Oden, An h–p adaptive method using clouds, Comput. Meth. Appl. Mech. Engrg. 139 (1996) 237–262.
[14] W. Han, G.J. Wagner, W.K. Liu, Convergence analysis of a hierarchical enrichment of Dirichlet boundary conditions in a meshfree method, Int. J. Numer. Meth. Engrg. 53 (2002) 1323–1336.
[15] R.D. Henshell, K.G. Shaw, Crack tip finite elements are unnecessary, Int. J. Numer. Meth. Engrg. 9 (1975) 495–507.
[16] L.A. Karlovitz, Remarks on variational characterization of eigenvalue and n-width problems, J. Math. Anal. Appl. 53 (1976) 99–110.

[17] L.A. Karlovitz, On a class of Kolmogorov n-width problems, Atti Accad. Naz. Lincei Cl. Sci. Fis. Mat. Natur. LIII (1973) 241–245.
[18] A. Kolmogoroff, Über die beste annaäherung van funktionen einer gegebenen funkticnklasse, Ann. Math. 37 (1963) 107–110.
[19] O. Laghrouche, P. Bettes, Solving short wave problems using special finite elements, in: J.R. Whiteman (Ed.), The Mathematics of Finite Elements and Applications, Elsevier, 2000, pp. 181–194.
[20] T.J. Liszka, C.A.M. Duarte, W.W. Tworzydlo, hp-Meshless cloud method, Comput. Meth. Appl. Mech. Engrg. 139 (1996) 263–288.
[21] W.K. Liu, Y. Chen, C.T. Chang, T. Belytschko, Advances in multiple scale kernel particle method, Comp. Mech. 18 (1996) 73–111.
[22] W.K. Liu, S. Jun, Y.F. Zhang, Reproducing kernel particle methods, Int. J. Numer. Meth. Fluids 20 (1995) 1081–1106.
[23] J.M. Melenk, I. Babuška, The partition of unity finite element method: theory and application, Comput. Meth. Appl. Mech. Engrg. 139 (1996) 289–314.
[24] A. Melkman, C. Micchelli, Spline subspaces are optimal for L^2 n-width, Illinois J. Math. 22 (1978) 541–564.
[25] A. Pinkus, n-Widths in Approximation Theory, Springer-Verlag, Berlin, 1980.
[26] T. Strouboulis, I. Babuška, K. Copps, The design and analysis of the generalized finite element method, Comput. Meth. Appl. Mech. Engrg. 181 (2000) 43–69.
[27] T. Strouboulis, K. Copps, I. Babuška, The generalized finite element method, Comput. Meth. Appl. Mech. Engrg. 190 (2001) 4081–4193.
[28] B. Szabo, I. Babuška, Finite Element Analysis, John Wiley & Sons Inc., New York, 1991.