SIAM J. NUMER. ANAL.
Vol. 34, No. 1, pp. 315–358, February 1997

013

FINITE ELEMENT SOLUTION OF THE HELMHOLTZ EQUATION WITH HIGH WAVE NUMBER PART II: THE *h-p* VERSION OF THE FEM*

FRANK IHLENBURG† AND IVO BABUŠKA‡

Abstract. In this paper, which is part II in a series of two, the investigation of the Galerkin finite element solution to the Helmholtz equation is continued. While part I contained results on the h version with piecewise linear approximation, the present part deals with approximation spaces of order $p \geq 1$. As in part I, the results are presented on a one-dimensional model problem with Dirichlet–Robin boundary conditions. In particular, there are proven stability estimates, both with respect to data of higher regularity and data that is bounded in lower norms. The estimates are shown both for the continuous and the discrete spaces under consideration. Further, there is proven a result on the phase difference between the exact and the Galerkin finite element solutions for arbitrary p that had been previously conjectured from numerical experiments. These results and further preparatory statements are then employed to show error estimates for the Galerkin finite element method (FEM). It becomes evident that the error estimate for higher approximation can—with certain assumptions on the data—be written in the same form as the piecewise linear case, namely, as the sum of the error of best approximation plus a pollution term that is of the order of the phase difference. The paper is concluded with a numerical evaluation.

Key words. Helmholtz equation, *h-p* finite element method, elliptic, partial differential equation

AMS subject classifications. 65N30, 65N15, 35A40, 35J05

PII. S0036142994272337

1. Introduction.

This paper is the second part (of two) of an investigation devoted to the numerical analysis of the Galerkin finite element method (FEM) for the reduced wave (Helmholtz) equation. The interest in this topic has grown over the last few years; related results have been published by a number of authors, both in the mathematical and in the engineering literature [2], [13], [14], [12], [21], [22], [15]. In part I of the present investigation [16], we analyzed the h version of the Galerkin (FEM) with piecewise linear approximation (i.e., elements of polynomial degree $p = 1$) on the following one-dimensional model problem: let $\Omega = (0, 1)$ and consider the boundary value problem (BVP)

$$u''(x) + k^2 u(x) = -f(x), \tag{1.1}$$

$$u(0) = 0, \tag{1.2}$$

$$u'(1) - iku(1) = 0, \tag{1.3}$$

or, equivalently, the variational problem: find $u \in H^1(\Omega),\ u(0) = 0$ such that

$$\mathcal{B}(u, v) = \int_0^1 \big(u'(x)\bar{v}'(x) - k^2 u(x)\bar{v}(x)\big)\, dx - iku(1)\bar{v}(1) = \int_0^1 f(x)\bar{v}(x) dx \tag{1.4}$$

holds for all $v \in H^1(\Omega),\ v(0) = 0$.

*Received by the editors August 3, 1994; accepted for publication (in revised form) April 17, 1995.
http://www.siam.org/journals/sinum/34-1/27233.html

†Institute for Physical Science and Technology, University of Maryland at College Park, College Park, MD 20742 (ihl@olgao.umd.edu). Current address: Texas Institute for Computational and Applied Mathematics, University of Texas at Austin, TAY2.400, Austin, TX 78712. The research of this author was supported in part by grant 517 402 524 3 of the German Academic Exchange Service (DAAD).

‡Institute for Physical Science and Technology, University of Maryland at College Park, College Park MD, 20742. The research of this author was supported in part by ONR grant N00014-93-I-0131. Current address: Texas Institute for Computational and Applied Mathematics, University of Texas at Austin, TAY 2.400, Austin, TX 78712.

Since we will analyze specifically the case of large wave number k, we assume throughout the paper that $k \geq 1$. We assume that the model problem is given on a domain that has been scaled to unit length; hence wave number k (as well as stepsize h of the FEM discretization) has no physical dimension.

We showed the following [16]:

A1. For data $f \in L^2(0,1)$, the boundary value problem (BVP) (1.1)–(1.3) has unique solution $u \in H^2(0,1)$.

A2. For data $f \in H^1(0,1)'$, the variational problem (VP) (1.4) has unique solution $u \in H^1(0,1)$.

A3. The form B defined by (1.4) is continuous for each fixed wave number k:

$$|B(u,v)| \leq C_o(k)|u|_1|v|_1$$

with $C_o = 1 + k + k^2$.

A4. The Babuška–Brezzi constant γ of the VP (1.4) is inversely proportional to the wave number k.

Remark 1. Obviously, the statements A1 and A2 hold true for higher regularity of the data; in general, the operators given by BVP (1.1)–(1.3) and VP (1.4) define bijective mappings from H^s to H^{s+2}.

For the Galerkin FEM with piecewise linear approximation ($p = 1$) we then proved the following:

F1. The finite element solution is (asymptotically) quasi optimal with respect to the H^1-seminorm provided k^2h is sufficiently small, where h is the size of the elements.

F2. The discrete Babuška–Brezzi (inf-sup) condition holds with $\gamma_h \sim \frac{1}{k}$.

F3. In the preasymptotic range, the error bound in the H^1-seminorm consists of a term of order kh (reflecting the error of best approximation) and a pollution term of order k^3h^2 (reflecting the phase difference between exact and finite element solutions).

Finally, we concluded from a numerical evaluation that the terms mentioned in F3 do indeed occur in the error and, consequently, the estimate (obtained from F3 by setting $\theta = kh$)

$$|e|_1 \leq C_1\theta + C_2k\theta^2, \tag{1.5}$$

where $C_1\theta$ is the minimal error in the approximation space, cannot be principally improved. The second term in the error bound indicates pollution of the optimal error for high wave number [8], [9], [10].

In the present paper we extend our study to elements of arbitrary (but fixed) polynomial order p. Due to the oscillatory character of the propagating solutions, the application of higher-order (e.g., quadratic or cubic) elements is considered a natural choice in many applied computations (cf. [21], [22], and references therein). However, except for an asymptotic error estimate, no results stemming from the numerical analysis of Galerkin-type FEM for Helmholtz problems seem to be available for $p > 1$. In the above mentioned asymptotic estimate, which is given without rigorous proof in [3], it is assumed that $k^2h \ll 1$ (cf. Theorem 3.2 of this paper). As we showed in part I, this assumption on the meshsize does not cover the size of practically applied meshes where the wave number k (but not k^2, which is a significant difference for large k) is commonly normalized[1] by the stepwidth h [21], [15].

[1] I.e., h is chosen such that the product hk remains close to some a priori fixed value.

Consequently, as in part I, it is the principal goal of the present investigation to show error estimates that hold under assumptions on hk only. We call these estimates "preasymptotic" in order to distinguish the results from the above mentioned asymptotic estimates. In section 3 of this paper, we prove two error estimates that hold in the preasymptotic range. Both estimates are written as a sum of the error of the best approximation (naturally, the estimates coincide in this first member) and a pollution term. While in the first statement the pollution term is obtained by using an L^2 estimate of the data, the second estimate employs a dual argument to raise the degree of hp^{-1} in the pollution term. Under certain additional assumptions, namely, if the exact solution is sufficiently regular and oscillating with frequency k, we see that the pollution term of the second estimate has (in k, h, and p) the order of the phase difference between the exact and the finite element solutions.

It is this phase difference that has been used as the principal criterion for the quality of various finite element approaches [15], [21], [22]. We show here (Theorem 3.2) that, for arbitrary p, the phase difference between the exact and the Galerkin finite element solutions is of order $k\mathcal{O}(\frac{kh}{2p})^{2p}$. To our knowledge, this statement has not been previously proven. It is in accordance with published computational results that have been obtained by Taylor series expansion in the finite element matrices for several fixed p [21].

Throughout the paper, special care is given to state, as precisely as possible, the dependence of the constants involved in the estimates on the principal parameters k, h, and p.

It turns out (both theoretically and in the numerical evaluation) that the constant in the best approximation estimate grows only moderately with p. On the other hand, the constant in the estimate of the pollution in dual norms (Theorem 3.7) theoretically grows significantly with increasing p. However, this growth has not been observed in several numerical experiments; it is therefore an open question whether this estimate is sharp with respect to p.

The paper is organized as follows. In section 2 we fix notations and discuss a definition of negative norms that is suited for our purposes. Further, we discuss auxiliary problems and conclude the section with the proof of stability results. In section 3 we investigate the Galerkin finite element solution to the model problem. First (subsection 3.1) we identify the approximation spaces $S_h^p(\Omega)$ and give an outline of the finite element solution procedure (subsection 3.2). In subsection 3.3 we prove the proposition on the phase lag (Theorem 3.2). Subsection 3.4 is devoted to stability propositions on the approximation subspaces. We first have a result that is true for general L^2 data f (Theorem 3.3); in the proof we employ an inf-sup condition for arbitrary order of approximation (Lemma 3.4). We then show a dual theorem for specific "bubble" data. This is motivated as follows. In the approximation theorem (Theorem 3.1) we had constructed, for a given function u, an approximating function s that is of optimal order of approximation in H^1-, L^2-, and dual norms and, beyond that, possesses specific interpolation properties with respect to u. We use these interpolation properties in the error analysis where we show that the finite element error is majorized by certain norms of the difference $u - s$. Since $u - s$ occurs as the data in variational problems related to the error estimation, we prepare the error analysis with propositions on stability. We then turn to the error analysis of the Galerkin FEM (subsection 3.5). We show an asymptotic estimate and study the finite element solution in the preasymptotic range. We show that, for oscillating solutions, the pollution term is of the order of the phase lag.

Finally, in section 4, we discuss results of selected numerical experiments that have been carried out to evaluate the error estimates of section 3.

2. Analytic solution properties.

2.1. Notations and preliminaries. We will use the following *notations*:

- **Constants.** If not stated otherwise, all constants $C, D, E, \ldots$, or C_i, where i is a natural number, are understood to be generic, independent of all parameters of the given estimate, and having, in general, different meanings in different contexts. However, some specific constants that are used repeatedly with the same meaning in different contexts are marked by literal subscript; so, we write C_a for a constant arising from the principal approximation estimate in the discrete space, and so forth.
- $L^2(0,1)$ is the space of all square-integrable complex-valued functions equipped with the inner product
$$(v,w) := \int_0^1 v(x)\bar{w}(x)dx$$
and the norm
$$\|w\| := \sqrt{(w,w)}.$$
- $H^s(0,1)$ denotes the Sobolev space
$$H^s = \left\{ u \,\middle|\, u \in L^2 \wedge \partial^i u \in L^2, i = 1,\ldots,s \right\},$$
where $\partial^i u$ are the derivatives of order i in the distribution sense. As usual, we define the subspace
$$H_o^s(0,1) = \left\{ u \in H^s(0,1) | \partial^i u(0) = \partial^i u(1) = 0, i = 0,\ldots,s-1 \right\}.$$
We will also work with subspaces $H^s_{(o}$ and $H^s_{o)}$ consisting of functions with Dirichlet data 0 given only in $x=0$ or $x=1$, resp. By $|u|_s := \|\partial^s u\|$ a seminorm is given on H^s. A norm of the space $H^s(0,1)$ is defined as $\|u\|_s := \left(\sum_{i=0}^s |u|_s^2\right)^{1/2}$. On H^s_o, $H^s_{(o}$, and $H^s_{o)}$ the seminorm $|\cdot|_s$ is a norm equivalent to $\|\cdot\|_s$.
- $H^{-s}(0,1) = (H^s(0,1))'$ denotes the dual to $H^s_{(o}$ space equipped with the norm
$$\|f\|_{-s} := \sup_{v \in H^s_{(o}} \frac{|(f,v)|}{|v|_s}. \tag{2.1}$$
- As usual, $f^{(i)}$, $i = 0,1,2,\ldots$, denotes the ith derivative of f. We generalize this notation for integration: $f^{(-i)}$ is a function subject to (s.t.) $\partial^i f^{(-i)}(x) = f(x)$. More specifically, we define for $f \in L^2(\Omega)$ and $i = 0,1,\ldots$,
$$f^{(-i-1)}(x) = -\int_x^1 f^{(-i)}(t)dt.$$

With these definitions, the dual H^l-norm of a L^2 function f is equal to the L^2-norm of $f^{(-l)}$.

LEMMA 2.1. *For $f \in L^2(\Omega)$ and $l = 0,1,2,\ldots$,*
$$\|f\|_{-l} = \|f^{(-l)}\| \tag{2.2}$$
holds.

Proof. Let $l = 1$ and write $F = f^{(-1)}$. Then, by partial integration,

$$\|F\| = \sup_{v \in L^2(\Omega)} \frac{\int_\Omega |Fv|}{\|v\|} = \sup_{v \in L^2(\Omega)} \frac{\int_\Omega |fV|}{\|v\|}$$

holds with $V := \int_0^x v(t)dt$. Obviously, $V \in H^1_{(o}(\Omega)$ and $V' = v$. On the other hand, every $V \in H^1_{(o}(\Omega)$ can be represented by an integral of an L^2 function, hence

$$\|F\| = \sup_{v \in H^1_{(o}(\Omega)} \frac{\int_\Omega fv}{\|v'\|} = \|f\|_{-1}.$$

This proves the statement for $l = 1$. The induction to higher l is obvious and the proof is completed. $\square$

Variational forms arising from the Helmholtz equation are, in general, indefinite. One can obtain, however, coercive forms if the wave number is properly restricted (we write K in order to distinguish this problem from the general case where k may be large). For later use, we consider here the case of Dirichlet boundary conditions: find $u \in V = H^1_o(\Omega)$ such that for all $v \in V$

$$B_K(u, v) = \int_\Omega u'\bar{v}' - K^2 \int_\Omega u\bar{v} = (f, v) \tag{2.3}$$

holds.

LEMMA 2.2. *Let $u \in V = H^1_o(\Omega)$ be the solution to the VP* (2.3) *with data f. Assume that $0 \le K \le \alpha < \pi$; then*

$$\|u\| \le \frac{1}{\pi^2 - \alpha^2} \|f\|, \tag{2.4}$$

$$\|u'\| \le \frac{\pi}{\pi^2 - \alpha^2} \|f\|, \tag{2.5}$$

$$\|u'\| \le \frac{\pi^2}{\pi^2 - \alpha^2} \|f^{(-1)}\|. \tag{2.6}$$

Proof. All inequalities are trivial for $u \equiv 0$, hence we may, without loss of generality, assume $\|u\| > 0$. It is easy to see that, for all $u \in H^1_o(\Omega)$,

$$\|u'\| \ge \pi\|u\|.$$

Hence

$$B_K(u, u) \ge (\pi^2 - K^2)\|u\|^2$$

and

$$B_K(u, u) \ge \frac{\pi^2 - K^2}{\pi^2} \|u'\|^2.$$

From the first inequality we easily conclude (2.4):

$$\|f\| = \sup_{v \in L^2(\Omega)} \frac{(f, v)}{\|v\|} \ge \frac{(f, u)}{\|u\|} \ge (\pi^2 - \alpha^2)\|u\|.$$

Equation (2.6) follows similarly.

Equation (2.5) follows from (2.4) by

$$\|u'\|^2 = (f, u) + K^2\|u\|^2 \le \|f\|\|u\| + \alpha^2\|u\|^2.$$

The proof is completed. $\square$

Remark 2. The statement of the lemma holds also if u and v are chosen from a Hilbert subspace $V_h \subset V$. Indeed, since obviously $\min_{u \in V}(|u|_1/\|u\|) \le \min_{u \in V_h}(|u|_1/\|u\|)$, the form B_K is strongly elliptic on the subspace and the same arguments apply.

2.2. Stability estimates for higher regularity. In part I we proved stability estimates for L^2 and H^{-1} data. We now generalize these properties in two directions. Namely, we first consider data of higher (than L^2) regularity and prove that for $l \geq 2$ the solution norm $|u|_{l+1}$ is bounded by $k^{l-1}\|f\|_{l-1}$.

We then show a dual result, i.e., we bound $|u|_1$ by $\|f^{-m}\|$. In this case we consider data of a specific type—"bubble" data—for which the integrals vanish at the boundaries of Ω. The sense of this assumption will become clear in the error analysis.

THEOREM 2.1 (*Continuous stability*). *Let f be the data and u the solution of the BVP* (1.1)–(1.3). *Assume, for $l > 1$, $f(x) \in H^{l-1}(0,1)$. Then $u \in H^{l+1}(0,1)$ and the estimate*

$$|u|_{l+1} \leq C_s(l) k^{l-1} \|f\|_{l-1} \tag{2.7}$$

holds for a positive constant $C_s(l) \leq Dl$, where D does not depend on k and l.

Remark 3. Except for the dependence on k (and l), the statement is similar to the well-known regularity result for the Laplace equation (cf. [19, pp. 52–53]).

Proof. Let us first consider the case $l = 2$. We have to prove $|u|_3 \leq Ck\|f\|_1$. We start from the Green's function representation of u (see part I for details):

$$u(x) = \int_0^1 G(x,s) f(s) ds, \tag{2.8}$$

where

$$G(x,s) = \frac{1}{k} \begin{cases} \sin kx e^{iks}; & 0 \leq x \leq s, \\ \sin ks e^{ikx}; & s \leq x \leq 1. \end{cases} \tag{2.9}$$

By partial integration,

$$u(x) = [H(x,s) f(s)]_{s=0}^{s=1} - \int_0^1 H(x,s) f'(s) ds \tag{2.10}$$

with

$$H(x,s) := \int G(x,s) ds = -\frac{1}{k^2} \begin{cases} i \sin kx e^{iks} + 1; & 0 \leq x \leq s, \\ \cos ks e^{ikx}; & s \leq x \leq 1. \end{cases} \tag{2.11}$$

For any fixed s (or x, resp.), $H(x,s)$ is a H^2 function of x (or s, resp.). In the boundary points, $H(x,0)$ and $H(x,1)$ are smooth (C^∞) functions. We now estimate

$$|u(x)| \leq |H(x,0)||f(0)| + |H(x,1)||f(1)| + \sup_{x,s} |H(x,s)| \|f'\|.$$

From (2.11) we have directly

$$\forall x, s: \quad |H(x,s)| \leq \frac{2}{k^2}$$

and with

$$\forall s: \quad |f(s)| \leq \sqrt{2} \|f\|_1$$

(for the proof see, e.g., [7]) we get

$$\|u\| \leq \frac{2}{k^2} \left(1 + 2\sqrt{2}\right) \|f\|_1. \tag{2.12}$$

For an estimate of the derivatives of u, differentiate in (2.10) to obtain

$$u'(x) = [H_x(x,s)f(s)]_{s=0}^{s=1} - \int_0^1 H_x(x,s)f'(s)ds \tag{2.13}$$

and from this and differentiation with respect to x in (2.11),

$$|u|_1 \leq \frac{1}{k}\left(1+2\sqrt{2}\right)\|f\|_1. \tag{2.14}$$

Similarly, since $H \in H^2(\Omega)$, we obtain from differentiating (2.13)

$$|u|_2 \leq \left(1+2\sqrt{2}\right)\|f\|_1. \tag{2.15}$$

Finally, since $u \in H^3(\Omega)$, the differential equation $u''' + k^2u' = f'$ holds (at least in the weak sense). Hence

$$|u|_3^2 \leq k^4|u|_1^2 + 2k^2|u|_1|f|_1 + |f|_1^2$$

and with (2.14) we obtain

$$|u|_3^2 \leq c^2k^2\|f\|_1^2 + 2ck\|f\|_1^2 + \|f\|_1^2$$

or, equivalently,

$$|u|_3 \leq C\,k\,\|f\|_1 \tag{2.16}$$

which proves the statement for $l = 2$.

For higher l we proceed analogously. First we integrate in (2.8) by parts $(l-1)$ times:

$$\begin{aligned} u(x) = {} & \left[G^{(-1)}(x,s)f(s)\right]_{s=0}^{s=1} - \left[G^{(-2)}(x,s)f'(s)\right]_{s=0}^{s=1} \pm \cdots \\ & + (-1)^{l-2}\left[G^{(-l+1)}(x,s)f^{(l-2)}(s)\right]_{s=0}^{s=1} \\ & + (-1)^{l-1}\int_0^1 G^{(-l+1)}(x,s)f^{(l-1)}(s)ds, \end{aligned} \tag{2.17}$$

where $G^{(-j)}(x,s) = \int G^{-(j-1)}(x,t)dt$ with $G^{(0)}(x,s) := G(x,s)$ and appropriate integration constants for continuity at $x = s$.

For fixed x, resp., s, the regularity is now $G^{(-j)}(x,s) \in H^{j+1}(0,1)$. At the boundaries we have again $G^{(-j)}(x,0), G^{(-j)}(x,1) \in C^\infty(0,1)$.

Therefore, differentiation of $G^{(-j)}(x,s)$ with respect to x is well defined at most l times. Hence for $j = 1,\ldots,l$,

$$\begin{aligned} \left|u^{(j)}(x)\right| \leq {} & \left|G^{(j-1)}(x,0)\right||f(0)| + \left|G^{(j-1)}(x,1)\right||f(1)| + \cdots \\ & + \left|G^{(j-l+1)}(x,0)\right|\left|f^{(l-2)}(0)\right| + \left|G^{(j-l+1)}(x,1)\right|\left|f^{(l-2)}(1)\right| \\ & + \left|\int_0^1 G^{(j-l+1)}(x,s)f^{(l-1)}(s)ds\right|. \end{aligned}$$

The data is bounded in any point by

$$\forall x \in [0,1],\ \forall j: \quad \left| f^{(j)}(x) \right| \leq \sqrt{2} \| f^{(j)} \|_1 \leq \sqrt{2} \| f \|_{j+1}.$$

Generalizing (2.11), the integrals of the Green's function can be written in the form

$$G^{(-m)}(x,s) = k^{-(1+m)} \varphi(x,s) + k^{-2} P_{m-1}(x,s),$$

where $\varphi(x,s)$ is an oscillating part (with $\sup |\varphi| \leq 1$) and P_{m-1} is a sum of polynomials of degree $m-1$ in s and x, resp. Consequently, P_{m-1} is bounded on Ω and

$$\left| G^{(j-m)}(x,s) \right| \leq \begin{cases} C_1(j,m) k^{-2} + C_2(j,m) k^{j-m-1} & \text{if} \quad j < m, \\ k^{j-m-1} & \text{if} \quad j \geq m. \end{cases}$$

In particular, for $j = l-1$ there exists a constant $C_3(l) = \max_{j,m}(C_1(j,m), C_2(j,m))$ s.t.

$$\left| u^{(l-1)}(x) \right| \leq C_3 \left(2 \left(k^{l-3} + k^{l-4} + \cdots + k^{-1} \right) + k^{-1} \right) \sqrt{2}\, \| f \|_{l-1}.$$

Similarly, for $j = l$ there exists $C_4(l)$ s.t.

$$\left| u^{(l)}(x) \right| \leq C_4 \left(2 \left(k^{l-2} + k^{l-3} + \cdots + 1 \right) + 1 \right) \sqrt{2}\, \| f \|_{l-1}.$$

Hence for $k > 1$ and $l > 2$ there are constants $C_5(l)$ and $C_6(l)$ not depending on k s.t.

$$|u|_{l-1} \leq C_5 k^{l-3} \| f \|_{l-1}, \tag{2.18}$$
$$|u|_l \leq C_6 k^{l-2} \| f \|_{l-1} \tag{2.19}$$

hold. By their definition, C_5 and C_6 are of order l.

Then, finally,

$$\begin{aligned} |u|_{l+1}^2 &= \int_0^1 \left(f^{(l-1)} - k^2 u^{(l-1)} \right)^2 dx \\ &\leq k^4 |u|_{l-1}^2 + 2k^2 |u|_{l-1} |f|_{l-1} + |f|_{l-1}^2 \\ &\leq \left(k^4 C_5^2 k^{2l-6} + 2\, k^2 C_5 k^{l-3} + 1 \right) \| f \|_{l-1}^2 \end{aligned}$$

and the statement readily follows. The proof is completed. ∎

We now proceed to a second stability result that is dual to the first one; i.e., we will bound lower solution norms by negative norms of the data. We employ two auxiliary problems:

1. Consider the second-order Dirichlet BVP on $\Omega = (0,1)$:

$$w'' - k^2 w = -g, \tag{2.20}$$
$$w(0) = w(1) = 0 \tag{2.21}$$

and the associated variational problem: for $g \in H^1(0,1)'$ find $u \in H_o^1(0,1)$ s.t.

$$\forall v \in H_o^1(0,1): \quad B_1(w,v) = \int_0^1 w'v' + k^2 \int_0^1 wv = \int_0^1 gv = (g,v). \tag{2.22}$$

2. Consider the fourth-order Dirichlet BVP

$$w'''' + k^4 w = g, \tag{2.23}$$

$$w(0) = w(1) = w'(0) = w'(1) = 0 \tag{2.24}$$

and the associated variational problem: for $g \in H^2(0,1)'$ find $u \in H_o^2(0,1)$ s.t.

$$\forall v \in H_o^2(0,1): \quad B_2(w,v) = \int_0^1 w''v'' + k^4 \int_0^1 wv = \int_0^1 gv = (g,v). \tag{2.25}$$

The forms B_1 and B_2 are coercive, hence (2.22) and (2.25) have unique solutions.

LEMMA 2.3. *Let u_1 and u_2 be the solutions of $B_1(w,v) = (g,v)$ and $B_2(w,v) = (g,v)$, resp.*

Then for u_1

$$\|u_1'\| \le \|g^{(-1)}\| \tag{2.26}$$

and for u_2

$$\|u_2''\| \le \|g^{(-2)}\|, \tag{2.27}$$

$$k\|u_2'\| \le \frac{1}{\sqrt{2}}\|g^{(-2)}\|, \tag{2.28}$$

$$k^2\|u_2'\| \le \frac{1}{2}\|g^{(-1)}\| \tag{2.29}$$

hold.

Proof. Let $|||\cdot|||_1$ and $|||\cdot|||_2$ be the energy norms induced by the forms B_1 and B_2, resp. In these norms, the B-B constant is trivially $\gamma = 1$ and hence

$$|||u_1|||_1 \le |||g|||_1',$$
$$|||u_2|||_2 \le |||g|||_2'.$$

It is then easy to see from the definitions of the various norms that

$$\|u_1'\| \le |||u_1|||_1 \le |||g|||_1' \le \|g\|_1' = \|g^{(-1)}\|,$$
$$\|u_2''\| \le |||u_2|||_2 \le |||g|||_2' \le \|g\|_2' = \|g^{(-2)}\|$$

hold. This proves (2.26) and (2.27).

To show (2.29), let $u \in H_o^2(\Omega)$. From

$$0 \le (\|u''\| - k^2\|u\|)^2$$

we conclude

$$2k^2\|u''\|\|u\| \le B_2(u,u)$$

and, by Schwarz inequality and partial integration,

$$B_2(u,u) \ge 2k^2|(u,u'')| = 2k^2\|u'\|^2. \tag{2.30}$$

On the other hand, also by partial integration and Schwarz inequality,

$$B_2(u,u) \le |(g^{(-1)}, u')| \le \|g^{(-1)}\|\|u'\|.$$

Hence

$$2k^2\|u'\|^2 \le \|g^{(-1)}\|\|u'\|.$$

Canceling $\|u'\|$ we obtain (2.29).

Finally, multiplying (2.30) by the trivial relation $B_2(u,u) \geq \|u''\|^2$ we have

$$2k^2\|u'\|^2\|u''\|^2 \leq B_2(u,u)^2. \tag{2.31}$$

But $B_2(u,u)$ is majorized by $\|g^{(-2)}\|\|u''\|$, and (2.28) readily follows cancelling $\|u''\|^2$ and taking roots. This completes the proof. □

THEOREM 2.2 (*A dual stability result*). *Assume that, for integer $m \geq 1$, there is given a function $f \in L^2(\Omega)$ such that $f^{(-i)}(0) = f^{(-i)}(1) = 0$ for $i = 1, \ldots, m$. Let $u \in H^1_{(o}(\Omega)$ be the solution to the VP (1.4) with data f.*

Then

$$|u|_1 \leq C_1 k^m \|f^{(-m)}\| + C_2\|f^{(-1)}\| \tag{2.32}$$

holds with C_1, C_2 independent of k.

Remark 4. The assumption on the data means that m integrals of f vanish in $x = 0$ (note that all integrals vanish at the endpoint by definition). Without this assumption we can just prove that

$$|u|_1 \leq C\ k\ \|f^{(-1)}\|$$

(see part I). In general, $|u|_1$ cannot be bounded by a term $C\|f^{(-1)}\|$ independently of k.

Proof.

(1) For $m = 1$ the statement simplifies to

$$|u|_1 \leq Ck\|f^{(-1)}\|$$

and is directly obtained from the B-B condition (see Introduction, A4) and general theory [4].

(2) Let $m = 2$. The basic ingredient of the argument is the introduction of a smoother kernel to the Green's function representation $u(x) = (G(x,s), f(s))$ of the solution. We define

$$K(x,s) := G(x,s) - H(x,s),$$

where $H(x,s)$ is the Green's function to the first auxiliary problem (2.22). Per definition of Green's functions, the equations

$$H_{xx}(x,s) - k^2\,H(x,s) = -\delta_s(x), \tag{2.33}$$

$$G_{xx}(x,s) + k^2\,G(x,s) = -\delta_s(x) \tag{2.34}$$

hold (in the sense of distributions). Hence, after subtraction,

$$K_{xx} = -k^2(G+H). \tag{2.35}$$

Since $G, H \in H^1(\Omega)$ (as functions of x for any fixed s) it follows that $K \in H^3(\Omega)$. We have constructed an equivalent integral representation of u as

$$u(x) = \int_0^1 K(x,s)f(s)ds + \int_0^1 H(x,s)f(s) := u_1(x) + u_2(x). \tag{2.36}$$

For estimation of $\|u_1'\|$ we integrate by parts

$$u_1'(x) = \int_0^1 K_x(x,s)f(s)ds = \int_0^1 K_{xss}(x,s)f^{(-2)}(s)ds$$

(note that the boundary terms vanish due to the specific assumption on f), hence by Cauchy–Schwarz inequality

$$|u_1'(x)| \leq \|K_{xss}\| \|f^{(-2)}\|.$$

From (2.35) and the symmetry property of Green's functions, $K_{xss} = -k^2(G+H)_x$ and

$$\|K_{xss}\| \leq k^2 \left(\|G_x\| + \|H_x\|\right)$$

follows. It is straightforward to show that $\|G_x\|$ and $\|H_x\|$ are bounded independently of k and we thus have

$$|u_1'(x)| \leq C\, k^2 \|f^{(-2)}\|. \tag{2.37}$$

To estimate u_2 we apply (2.26), Lemma 2.2, for $g = f$:

$$\|u_2'\| \leq \|f^{(-1)}\|$$

which, together with (2.37), yields the statement, case 2.

(3) Consider the case $3 \leq m \leq 6$. In analogy to the previous step, we introduce a still smoother kernel to the integral representation of u. Let us prove the extreme case $m = 6$. We start from (2.36) and define

$$L(x,s) := K(x,s) - J(x,s),$$

where $J(x,s)$ is a Green's function with the same singularity as K. Then

$$\begin{aligned} u(x) \;&=\; \int_0^1 L(x,s) f(s) ds + \int_0^1 J(x,s) f(s) + \int_0^1 H(x,s) f(s), \\ &:= u_o(x) + u_1(x) + u_2(x). \end{aligned} \tag{2.38}$$

To find the appropriate J, add (2.33) and (2.34):

$$(G+H)_{xx}(x,s) + k^2(G-H)(x,s) = -2\delta_s(x).$$

Multiplying this equation by $-k^2$ and substituting $K \;=\; G - H$, (2.35) leads to

$$K_{xxxx}(x,s) - k^4 K(x,s) = 2\, k^2\, \delta_s(x).$$

Hence if $J(x,s)$ is the Green's function to the second auxiliary problem with data $g(x) = 2\, k^2 f(x)$, then

$$L_{xxxx} \;=\; (K - J)_{xxxx} \;=\; k^4\,(K + J).$$

Since $K, J \in H^3(\Omega)$ it follows that $L \in H^7(\Omega)$. Thus, integrating by parts,

$$\begin{aligned} |u_o'(x)| &= \left| \int L_{xssssss}(x,s) f^{(-6)}(s) ds \right| \\ &= k^4 \left| \int (K+J)_{ssx}(x,s) f^{(-6)}(s) ds \right| \\ &= k^4 \left| \left(k^2 \int (G+H)_x(x,s) f^{(-6)}(s) ds + \int J_x f^{(-4)}(s) ds \right) \right| \end{aligned}$$

and, therefore,

$$\|u_o'\| \le Ck^6 \|f^{(-6)}\| + k^4 \|u_3'\|,$$

where u_3 is a solution to the second auxiliary problem (2.25) with data $g = 2\,k^2 f^{(-4)}$. From Lemma 2.3 and Poincaré's inequality,

$$\|u_3'\| \le |u_3|_2 \le 4k^2 \|f^{(-6)}\|,$$

hence

$$\|u_o'\| \le C_1\, k^6 \|f^{(-6)}\|.$$

By definition (2.38), u_1 is the solution to the VP (2.25) with data $g = 2\,k^2 f$ and u_2 solves the VP (2.22) with data f. Then it follows from Lemma 2.2. that

$$\|u_1'\| \le \frac{c}{k^2}\, \|g^{(-1)}\| = C\, \|f^{(-1)}\|$$

and

$$\|u_2'\| \le \|f^{(-1)}\|$$

so that finally u can be estimated as

$$|u|_1 \le C_1\, k^6\, \|f^{(-6)}\| + C_2 \|f^{(-1)}\|$$

which proves the statement for $m = 6$.

For $m = 3, \ldots, 5$ the argument is completely analogous, the only difference being the number of partial integrations in the representation of u_o. For $m \ge 7$, we introduce further smootheners and proceed in the same manner. The proof is completed. ∎

3. Finite element solution. In this section we analyze the h version of the Galerkin FEM with approximation order p. We first identify the finite dimensional approximation space $V_h \subset H^1_{(o}(\Omega)$ and prove two properties of the best approximation in the H^1-seminorm. After an outline of the finite element solution procedure (subsection 3.2) we turn to the analysis of an h-p Galerkin FEM for the numerical solution of the model problem. In subsection 3.3 we prove a result on the phase difference between the exact and the numerical solutions. In subsection 3.4 we give several stability estimates for the numerical solution u_{fe}. We prove a discrete inf-sup condition for general p and conclude stability of u_{fe} with respect to L^2 data, measured in L^2- or H^{-1}-norm, resp. We then proceed to the proof of a dual estimate in still lower norms, given for specific "bubble" data as it is encountered in the error estimation for $p > 1$. All stability constants are explicitly computed with respect to the wave number k. The analysis of the FEM is concluded with error estimates (subsection 3.5).

3.1. The approximation space V_h. Assume that the solution domain Ω has been uniformly divided into n disjunct intervals $\Delta_i = (x_{i-1}, x_i)$ called finite elements. Let $h = x_i - x_{i-1}$ (stepwidth). The set of nodal points

$$X_h = \{0 = x_o, x_1, \ldots, x_n = 1\} \tag{3.1}$$

will be called the finite element mesh. We will consider mesh functions defined on X_h and refer to them by subscript h; for the nodal value of a mesh function u_h in a node $x_i \in X_h$ we will write shortly $u_i := u_h(x_i)$.

Let $p > 0$ be integer and Δ a finite element. We denote by $S^p(\Delta)$ the linear space of all polynomials with domain of definition $\bar{\Delta}$ and degree $\leq p$. For given mesh, we define the space of piecewise polynomial, continuous functions

$$V_h := S_h^p(0,1) := \left\{ s \in H^1_{(o}(\Omega), \ \ s(x)\,|_{\Delta_i} \in S^p(\Delta_i), \ \ i = 1, \ldots, n \right\}.$$

Thus by definition $V_h \subset H^1_{(o}(0,1)$. If not stated otherwise, we will assume that V_h is equipped with the H^1-seminorm.

Within the elements we introduce the local coordinate ξ by linear mapping $\Delta_i \to I = (-1,1)$. The polynomials in $S^p(I)$ are then written as linear combinations of the *nodal* shape functions

$$N_1(\xi) = \frac{1-\xi}{2} \ , \ N_2(\xi) = \frac{1+\xi}{2} \ ; \ -1 \leq \xi \leq 1$$

and (if $p > 1$) the *internal* shape functions

$$N_l(\xi) = \phi_{l-1}(\xi) \ ; \ l = 3, 4, \ldots, p+1,$$

where ϕ_l is written in terms of the Legendre polynomials P_j

$$\phi_l(\xi) = \sqrt{\frac{2l-1}{2}} \int_{-1}^{\xi} P_{l-1}(t)dt$$

(see [20, pp. 38–39]). The internal shape functions vanish at the element boundaries, forming the subspace $S_o^p(I) = \operatorname{span}\{N_3, N_4, \ldots, N_{p+1}\} \subset S^p(I)$.

Let $\Delta_i = (x_{i-1}, x_i)$ be a finite element of length h. Denote by $(\cdot,\cdot)_\Delta$ the L^2 inner product on Δ and by $\|\cdot\|_\Delta$ the induced local L^2-norm. Similarly we use the notation $\|\cdot\|_I$ on $I = (-1,1)$. On $S^p(I)$, an inverse inequality (in p) is given by the well-known Markov theorem[2] stating for $y \in S^p(I)$ the inequality

$$\|y'\|_I \leq C_{inv}(p)\, p^2 \|y\|_I \tag{3.2}$$

with $C_{inv}(p) = (p+1)^2/(p^2\sqrt{2})$ [11]. Hence $C_{inv}(p) \leq 4\sqrt{2}$ and $C_{inv}(p) \to 1/\sqrt{2}$ for $p \to \infty$.

We now prove that for any $u \in H^{l+1}(\Omega)$ one can find a piecewise polynomial, nodally exact (on X_h) function $s \in V_h$ s.t. the following:

- s is an optimal (in h and p) approximation of u in the H^1-seminorm; and
- the integrals of s are nodally exact, quasi-optimal approximations of the integrals of u.

THEOREM 3.1 (*Approximation in V_h*). *Let l, p be integers with $1 \leq l \leq p$ and let $u \in H^{l+1}(0,1)$. There exists an $s \in V_h = S_h^p(0,1)$ s.t. the following:*

1. *(nodally exact approximation)*

$$\forall x_i \in X_h : \quad s^{(m)}(x_i) = u^{(m)}(x_i)\,, \quad m = -p+1, \ldots, 0; \tag{3.3}$$

2. *(order of approximation)*

$$\|(u-s)^{(m)}\| \leq C_a(l) C_a(-m) \left(\frac{h}{2p}\right)^{l-m+1} |u|_{l+1}\,, \quad m = -p+1, \ldots, 1, \tag{3.4}$$

[2]This inequality is usually given in the L^∞-norm (cf., e.g., [18, p. 124]). We use an L^2 variant [11].

where C_a satisfies the following:

1. $C_a(-1) = 1$ *(formal definition);*
2. $C_a(0) = 1$*;*
3. C_a *decreases for* $0 \leq l \leq \sqrt{p}$*;*
4. C_a *increases for* $l > \sqrt{p}$*; and*
5.

$$C_a(p) = \left(\frac{e}{2}\right)^p (\pi p)^{-\frac{1}{4}} \tag{3.5}$$

is the maximum of $C_a(l)$ *over* $l \in \{0, 1, \ldots, p\}$.

Remark 5. With respect to the stepwidth h, the estimate (3.4) is the standard approximation result of finite element theory (see, e.g., [19, pp. 46–49]).

Remark 6. For $l = 0, 1$ the statements are proven in [7]. The following argument is a generalization of this proof.

Proof. We start on the local level. Let Δ_i be a finite element and let $I = (-1, 1)$. We write $u'(\xi) \in H^l(I) \subset L^2(I)$ as

$$u'(\xi) = \sum_{i=0}^{\infty} a_i P_i(\xi),$$

where $P_i(\xi)$ are the Legendre polynomials of order i and equality is understood in the L^2 sense. Set

$$s'(\xi) := \sum_{i=0}^{p-1} a_i P_i(\xi)$$

and define the integrals ($i = 0, 1, 2, \ldots$):

$$u^{(-i)}(\xi) = u^{(-i)}(1) - \int_{\xi}^{1} u^{(-i+1)}(\tau) d\tau, \tag{3.6}$$

$$s^{(-i)}(\xi) = u^{(-i)}(1) - \int_{\xi}^{1} s^{(-i+1)}(\tau) d\tau. \tag{3.7}$$

We will now prove that (3.3) holds. First let $i = 0$; then from (3.6), (3.7) we have trivially $u(1) = s(1)$. Further, by definition,

$$\begin{aligned} u(-1) = u(1) - \int_{-1}^{1} u'(\tau) d\tau &= u(1) - \sum_{j=0}^{\infty} a_j \int_{-1}^{1} P_j(\tau) d\tau = u(1) - 2a_o \\ = u(1) - \int_{-1}^{1} s'(t) dt &= u(1) - \sum_{j=0}^{p-1} a_j \int_{-1}^{1} P_j(\tau) d\tau = s(-1). \end{aligned}$$

Now we integrate $u'(\xi)$, using the well-known relation

$$P_i(\tau) = \frac{(P'_{i-1}(\tau) - P'_{i+1}(\tau))}{(2i+1)}$$

to obtain

$$u(\xi) = u(1) + a_o(P_1(\xi) - P_o(\xi)) + \sum_{i=1}^{\infty} a_i \frac{P_{i+1}(\xi) - P_{i-1}(\xi)}{2i+1}.$$

Integrating once more (we write $U := u^{(-1)}$),

$$U(-1) = U(1) - \int_{-1}^{1} u(\tau)d\tau = U(1) - 2u(1) + 2a_o + \frac{2a_1}{3}.$$

Obviously, the same result is obtained from the integration of the polynomial $s(\xi)$ since only the coefficient of P_o influences the result of integration over the whole interval I. By similar argument we conclude that integration of the polynomial s on the one hand and the function u on the other hand leads to the same result exactly $p-1$ times. Indeed, by replacing repeatedly $P_i(\tau) = (P'_{i-1}(\tau) - P'_{i+1}(\tau))/(2i+1)$, we see that with the ith successive integration of $u(\xi)$ or $s(\xi)$ the coefficient a_i enters the set of coefficients multiplying P_o. Since the norms of $u^{(i)}$ and $s^{(i)}$ depend only on the coefficient of P_o, both norms are equal until P_o is multiplied by a_{p-1}; i.e., in general $u^{(-p+1)}(\xi) = s^{(-p+1)}(\xi)$ and $u^{(-p)}(\xi) \neq s^{(-p)}(\xi)$. Thus nodal exactness (3.3) is proved on an arbitrary element and hence it holds globally.

Let us now prepare the proof of estimate (3.4). With above definitions, the error of approximation is

$$e'(\xi) := u'(\xi) - s'(\xi) = \sum_{i=p}^{\infty} a_i P_i(\xi)$$

and from the orthogonality property of the Legendre polynomials we have

$$\|e'\|^2 = \sum_{i=p}^{\infty} \frac{2}{2i+1} a_i^2. \tag{3.8}$$

It can be proven (see [7, Chapter 3]) that s' is the best L^2 approximation to u' on I and the estimate

$$\|u' - s'\| \leq \frac{C_a(l)}{p^l} \, |u|_{l+1} \tag{3.9}$$

holds for $0 \leq l \leq p$, where the constant C_a has the properties 2–5 given in the theorem.

Integrating the error e', we get

$$e(\xi) = \int_{\xi}^{1} (u'(t) - s'(t))\, dt = \sum_{i=p}^{\infty} a_i \int_{\xi}^{1} P_i(t)dt = -\sum_{i=p}^{\infty} \frac{a_i}{2i+1} \left(P_{i+1}(\xi) - P_{i-1}(\xi)\right).$$

After reordering,

$$e(\xi) = \sum_{i=p+1}^{\infty} b_i P_i(\xi) + \frac{a_p}{2p+1} P_{p-1}(\xi) + \frac{a_{p+1}}{2p+3} P_p(\xi)$$

with

$$b_i = \frac{a_{i+1}}{2i+3} - \frac{a_{i-1}}{2i-1}$$

and the norm is

$$\|e\|^2 = \sum_{i=p+1}^{\infty} \frac{2}{2i+1} b_i^2 + \frac{a_p^2}{(2p+1)^2} \frac{2}{2p-1} + \frac{a_{p+1}^2}{(2p+3)^2} \frac{2}{2p+1}. \tag{3.10}$$

We apply the relation $(a-b)^2 \le 2a^2 + 2b^2$ to obtain (for $i \ge p+1$)

$$b_i^2 \le \frac{2a_{i-1}^2}{(2i-1)^2} + \frac{2a_{i+1}^2}{(2i+3)^2}$$

and, thus,

$$\sum_{i=p+1}^{\infty} b_i^2 \frac{2}{2i+1} \le \frac{1}{2p^2} \sum_{i=p}^{\infty} a_i^2 \frac{2}{2i+1} + \frac{1}{2p^2} \sum_{i=p+2}^{\infty} a_i^2 \frac{2}{2i+1}$$

holds. Now taking into account the second and third member in the right-hand side of (3.10) we get

$$\|e\|^2 \le \frac{1}{p^2} \sum_{i=p}^{\infty} a_i^2 \frac{2}{2i+1}$$

and hence

$$\|e\| \le \frac{1}{p}\|e'\|. \tag{3.11}$$

From (3.9) it then follows that

$$\|e\| \le \frac{C_a(l)}{p^{l+1}}|u|_{l+1} \tag{3.12}$$

holds for $1 < l \le p$.

Let us conclude the local analysis showing an orthogonality property for e. Since s' is the L^2 projection of u' on $S^{p-1}(I)$,

$$\int_{-1}^{1} (u'(\xi) - s'(\xi))\xi^m d\xi = 0 \tag{3.13}$$

holds for $m = 0, 1, \ldots, p-1$. We claim that $e(\xi) = \int_\xi^1 (u'(t) - s'(t))dt$ is orthogonal to $S^{p-2}(I)$.

Indeed, for $m \ge 0$ we compute

$$\begin{aligned}\int_{-1}^{1} e(\xi)\xi^m d\xi &= \int_{-1}^{1} \left(\int_\xi^1 (u'(t) - s'(t))dt \right) \xi^m d\xi \\ &= \int_{-1}^{1} \left((u'(t) - s'(t)) \int_{-1}^{t} \xi^m d\xi \right) dt \\ &= \frac{1}{m+1} \int_{-1}^{1} (u'(t) - s'(t))(t^{m+1} + 1)dt\end{aligned}$$

which, together with (3.13), proves that $e \perp S^{p-2}(I)$.

This completes the local analysis. By back transform $I \to \Delta$ and summation over the elements we conclude (3.4) for H^1- and L^2-norm, i.e., cases $m = 1, 0$ in (3.4).

It remains to prove (3.4) for dual norms. We apply a standard argument [19]. By definition, for $m \ge 1$,

$$\|e\|_{-m} = \sup_{v \in H^m_{(o}} \frac{(e, v)}{|v|_m}.$$

Let $P^m v \in S_h^m(\Omega)$ be the L^2 projection of $v \in H_{(o}^m$ on $S_h^m(\Omega)$. Then by orthogonality, as proven in step 1,

$$\|e\|_{-m} = \sup_{v \in H_{(o}^m} \frac{(e, v - P^{m-1}v)}{|v|_m}$$

holds for $1 \le m \le p-1$. Applying Schwarz inequality and (3.12) we conclude for $1 \le l \le p$ and $1 \le m \le p-1$ the estimate

$$\|e\|_{-m} \le C_a(l) \left(\frac{h}{2p}\right)^{l+1} |u|_{l+1}\, C_a(m) \left(\frac{h}{2p}\right)^{m} \frac{|v|_m}{|v|_m} \le C(l,m) \left(\frac{h}{2p}\right)^{l+m+1} |u|_{l+1},$$

where

$$C(l,m) = C_a(l) C_a(m) \le \left(\frac{e}{2}\right)^{2p} (\pi p)^{-\frac{1}{2}}.$$

This completes the proof of Theorem 3.1. □

3.2. The FEM. Let $V_h \subset H_{(o}^1(\Omega) =: V$ be the approximation space introduced in the previous subsection. As usual, a function $u_{fe} \in V_h$ is called the finite element solution of the VP (1.4) if

$$\forall v \in V_h: \quad B(u_{fe}, v) = (f, v). \tag{3.14}$$

The approximation space $V_h = S_h^p(\Omega)$ can be written as a direct sum of two subspaces, namely,

$$S_h^p(\Omega) = S_h^1(\Omega) \oplus S_o^p(\Omega),$$

where S_h^1 is the space of continuous, piecewise linear functions and

$$S_o^p(\Omega) = \bigoplus_{j=1}^{n} S_o^p(\Delta_j)$$

with

$$S_o^p(\Delta_j) = \operatorname{span}\{N_3^j, \ldots, N_{p+1}^j\}.$$

Here, Δ_j are the finite elements, hence $S_o^p(\Delta_j)$ are local spaces of "bubble" polynomials.

Now writing $u_{fe} = u_h + u_p, v = v_h + v_p$, where $u_h, v_h \in S_h^1(\Omega)$ and $u_p, v_p \in S_o^p(\Omega)$, we have from eq (3.14)

$$\forall v_h \in S_h^1(\Omega): \quad B(u_{fe}, v_h) = (f, v_h) \tag{3.15}$$

and

$$\forall v_p \in S_o^p(\Omega): \quad B(u_{fe}, v_p) = (f, v_p). \tag{3.16}$$

From (3.15) we conclude

$$\forall v_h \in S_h^1(\Omega): \quad B(u_h, v_h) = (f + k^2 u_p, v_h). \tag{3.17}$$

On the other hand, since $S_o^p(\Omega) = \oplus_1^n S_o^p(\Delta)$ (3.16) decouples into n independent local systems

$$\forall w \in S_o^p(\Delta): \quad B_\Delta(u_p, w) = (f + k^2 u_h, w)_\Delta, \tag{3.18}$$

where B_Δ is the restriction of the form B to the finite element Δ. By formally solving these equations, we express u_p in terms of f and u_h. The result can be inserted into (3.17), giving rise to

$$\tilde{B}(u_h, v_h) = (\tilde{f}, v_h). \tag{3.19}$$

Discretizing this equation by standard approach, we obtain the linear system

$$[L_h]\{u_h\} = \{r_h\}, \tag{3.20}$$

where $[L_h]$ is an $(n \times n)$ matrix, usually called the condensed stiffness matrix, and $\{u_h\} = u_{fe}|_{X_h}$ is the vector of nodal values of the finite element solution on the mesh X_h. The piecewise linear part of u_{fe} is determined by (3.20), provided $[L_h]$ is nonsingular. The internal part u_p can then be found locally by (3.18).

For the sake of further analysis, we outline below the details of the solution procedure.

Step 1 *(Local approximation and static condensation).* On any element Δ_j, the trial function u and the test function v are written as scalar products of shape functions $\{N_1^j, N_2^j, \ldots, N_{p+1}^j\}$ and the vectors of unknown coefficients $\{a^j\} = \{a_1^j, a_2^j, \ldots, a_{p+1}^j\}^T$ and $\{b^j\} = \{b_1^j, b_2^j, \ldots, b_{p+1}^j\}^T$, respectively. Identifying $a_1^j = u(x_{j-1}), a_2^j = u(x_j)$, and $b_1^j = v_{j-1}, b_2^j = v_j$, we have $u_h|_{\Delta_j} = a_1^j N_1^j + a_2^j N_2^j$. The condition that u_{fe} be the solution of the VP (1.4) for all $v \in V_h$ leads locally (i.e., on Δ_j) to

$$\{\bar{b}^j\}^T [B^j]\{a^j\} = \{\bar{b}^j\}^T\{r^j\} \tag{3.21}$$

with the $(p+1) \times (p+1)$ square matrix $(l, m = 1, \ldots, p+1)$

$$[B^j] = \left[\left\{\int_{\Delta_j} N_l^j(x)'\, N_m^j(x)' dx - k^2 \int_{\Delta_j} N_l^j(x)\, N_m^j(x) dx - ikN_l^j(1)N_m^j(1)\right\}\right] \tag{3.22}$$

and the right-hand side

$$\{r^j\} = \left\{(f(x), N_l^j(x))_{\Delta_j}, l = 1, \ldots, p+1\right\}^T.$$

Now, decomposing

$$[B^j] = \begin{bmatrix} \left[B_{11}^j\right] & \left[B_{12}^j\right] \\ \\ \left[B_{21}^j\right] & \left[B_{22}^j\right] \end{bmatrix}, \tag{3.23}$$

where $[B_{11}^j]$ is the left upper 2×2 submatrix of $[B^j]$, and assuming for the moment that $[B_{22}^j]$ is nonsingular, we define

$$[CB^j] = \left[B_{11}^j\right] - \left[B_{12}^j\right]\left[B_{22}^j\right]^{-1}\left[B_{21}^j\right]. \tag{3.24}$$

Then, by local variation of $\{b_3^j, \ldots, b_{p+1}^j\}^T$, we find—cf. (3.18)—

$$\{v_{j-1}\ v_j\}\left[CB^j\right]\left\{\begin{array}{c} u_{j-1} \\ u_j \end{array}\right\} = \{v_{j-1}\ v_j\}\left\{\begin{array}{c} \tilde{r}_{j-1} \\ \tilde{r}_j \end{array}\right\}, \tag{3.25}$$

where

$$\left\{ \begin{array}{c} \tilde{r}_{j-1} \\ \tilde{r}_j \end{array} \right\} = \left\{ \begin{array}{c} r_1^j \\ r_2^j \end{array} \right\} - \left[B_{12}^j\right]\left[B_{22}^j\right]^{-1} \left\{ \begin{array}{c} r_3^j \\ \vdots \\ r_{p+1}^j \end{array} \right\}. \tag{3.26}$$

On uniform mesh, the local matrices $[CB^j]$ are identical on all elements and can be written in the form

$$[CB] = \left[\begin{array}{cc} S_p(K) & T_p(K) \\ T_p(K) & S_p(K) \end{array} \right], \tag{3.27}$$

where $S_p(K)$ and $T_p(K)$ are rational polynomial functions of $K = kh/2$.

Remark 7. The analogy to the h version with piecewise linear approximation is given by the following consideration. On Δ_j, the homogeneous h-p finite element solution is written as

$$u_{fe}(x) = u_{fe}(x_{j-1})N_1^p(x) + u_{fe}(x_j)N_2^p(x), \tag{3.28}$$

where the condensed shape functions $N_j^p \in S^p(\Delta_j)$ are local variational solutions to the homogeneous Helmholtz equation with appropriate boundary conditions.

The local stiffness matrix is written as[3]

$$[CB^j] = \left[\begin{array}{cc} B(N_1^p, N_1^p) & B(N_1^p, N_2^p) \\ & \\ B(N_2^p, N_1^p) & B(N_2^p, N_2^p) \end{array} \right]$$

which is again the formal analogon to the h version (cf. part I).

Step 2 *(Global assembling and solution for u_h).* Enforcing continuity of the test functions in the nodal points of X_h we obtain the set of linear equations (3.20). The discrete operator L_h is an $n \times n$ tridiagonal matrix

$$[L_h] = \left[\begin{array}{ccccc} 2S_p(K) & T_p(K) & & & \\ T_p(K) & 2S_p(K) & T_p(K) & & \\ & & \ddots & & \\ & & T_p(K) & 2S_p(K) & T_p(K) \\ & & & T_p(K) & S_p(K) - iK \end{array} \right], \tag{3.29}$$

the global right-hand side vector is

$$R_h = \left\{ \begin{array}{c} \tilde{r}_1 \\ \vdots \\ \tilde{r}_{j-1} + \tilde{r}_j \\ \vdots \\ \tilde{r}_n \end{array} \right\} \tag{3.30}$$

and $K = kh/2$ is a measure for the number of elements per wavelength (see part I, Remark 8).

Step 3 *(Local "decondensation").* The discrete equivalent of (3.18)

$$\left[B_{22}^j\right] \left\{ \begin{array}{c} a_3^j \\ \vdots \\ a_{p+1}^j \end{array} \right\} = \left\{ \begin{array}{c} r_3^j \\ \vdots \\ r_{p+1}^j \end{array} \right\} - \left[B_{21}^j\right] \left\{ \begin{array}{c} u_{j-1} \\ u_j \end{array} \right\} \tag{3.31}$$

[3]See the next subsection for the proof.

can be inverted—provided $[B_{22}^j]$ is regular—to determine locally the bubble modes of the finite element solution.

3.3. Discrete Green's function and discrete wave number. The nodal values of the finite element solution on X_h are found from the tridiagonal linear system (3.20). On uniform mesh, this system consists of formally identical difference stencils

$$T_p(K)u_h(x_{j-1}) + 2S_p(K)u_h(x_j) + T_p(K)u_h(x_{j+1}).$$

Equating these stencils to zero we find the homogeneous solutions

$$y_{h1} = \exp(ik'x_h), \quad y_{h2} = \exp(-ik'x_h), \tag{3.32}$$

where the parameter k' is determined as a function of k, h, and p by

$$\cos(k'h) = -\frac{S_p(K)}{T_p(K)}. \tag{3.33}$$

The discrete wave number k' is, in general, different from k, causing a phase difference between the finite element and exact solutions.

The discrete finite element solution $\{u_h\}$ can be written as

$$u_h(x_i) = h\sum_{j=1}^{n} G_h(x_i, s_j)R_h(s_j) \tag{3.34}$$

with the discrete Green's function (cf. part I)

$$G_h(k', x_i, s_j) = \frac{1}{\sin k'h}\begin{cases} \sin k'x_i\,(A\sin k's_j + \cos k's_j)\,, & x \le s, \\ \sin k's_j(A\sin k'x_i + \cos k'x_i), & s \le x \le 1. \end{cases} \tag{3.35}$$

Note that the Green's function does not depend directly on p. The dependence occurs only implicitly through the parameter k'. This means that the estimates of G that are given in part I for piecewise linear approximation carry over to higher p without modification, except for the estimation of k' in terms of k, h, and p.

THEOREM 3.2 (*Phase difference*). *Let $p \ge 1$ and k' be the parameter in the fundamental system* (3.32) *of the set of linear equations* (3.20).

Then, if $hk < 1$,

$$|k' - k| \le k\,C\left(\frac{C_a(p)}{2}\right)^2\left(\frac{hk}{2p}\right)^{2p}, \tag{3.36}$$

where k is the exact wave number, C_a is the approximation constant from Theorem 3.1*, and C does not depend on k, h, and p.*

Remark 8. The phase difference between the exact and the finite element solution has been extensively studied in [15] (for $p = 1$) and [21] (for $p = 1, 2, 3$). As a conclusion from numerical experiments, the statement of the theorem has been induced with respect to k and h in [21]. We now prove that with increasing p the phase difference is also going down with a factor $\left(\frac{e}{2}\right)^{2p}(\pi p)^{-1/2}(2p)^{-2p}$; i.e., the improvement with higher approximation is still more significant than it was assumed in above-named references.

We will give the proof of this theorem after the following preliminary discussion: first we observe that any homogeneous solution to the Helmholtz equation can be written on each inner element $\Delta_j \subset \Omega$ as

$$u(x) = u^1 t_1(x) + u^2 t_2(x), \tag{3.37}$$

where $u^1 := u(x_{j-1}), u^2 := u(x_j)$ are the nodal values of u on X_h, whereas t_1, t_2 satisfy

$$t'' + k^2 t = 0 \quad \text{on} \ \ \Delta_j \tag{3.38}$$

with inhomogeneous local Dirichlet data

$$t_1(x_{j-1}) = 1, \quad t_1(x_j) = 0 \tag{3.39}$$

or

$$t_2(x_{j-1}) = 0, \quad t_2(x_j) = 1, \tag{3.40}$$

resp.

By discrete evaluation of the VP (1.4) in the nodal points of X_h we find that, for $j = 1, \ldots, n-1$,

$$T_o(K)u(x_{j+1}) + 2S_o(K)u(x_j) + T_o(K)u(x_{j-1}) = 0$$

holds with

$$\begin{aligned} T_o(K) &= B(t_1, t_2) = B(t_2, t_1), \\ 2S_o(K) &= B(t_1, t_1) + B(t_2, t_2). \end{aligned}$$

The fundamental solutions are

$$z_{h1} = \exp(ikx_h), \quad z_{h2} = \exp(-ikx_h)$$

and

$$\cos(kh) = -\frac{S_o(K)}{T_o(K)}$$

holds.

Second, writing the finite element solution on Δ_j as

$$u_{fe}(x) = u^1_{fe} N^p_1(x) + u^2_{fe} N^p_2(x),$$

where $N^p_1, N^p_2 \in S^p(\Delta_j)$ are approximate solutions to (3.38) with boundary conditions ((3.39) or (3.40), resp.) we have

$$T_p = B(N^p_1, N^p_2) = B(N^p_2, N^p_1), \tag{3.41}$$

$$2S_p = B(N^p_1, N^p_1) + B(N^p_2, N^p_2). \tag{3.42}$$

To see this, let us analyze the BVPs (3.38), (3.39) and (3.38), (3.40) on $I = (-1, 1)$. After linear transformation $\Delta_j \to I$ we arrive at

$$t''(\xi) + K^2 t(\xi) = 0. \tag{3.43}$$

We assume that $K \leq \alpha < 2\pi$. The boundary conditions are

$$t(-1) = 1, \quad t(1) = 0 \tag{3.44}$$

or

$$t(-1) = 0, \quad t(1) = 1, \tag{3.45}$$

resp.

To formulate an equivalent variational problem, we write the admissible functions as

$$t = N_1 + \phi_1 \quad \text{or} \quad t = N_2 + \phi_2,$$

resp., where $\phi_i \in H_o^1(I)$ and N_1, N_2 are the linear shape functions. The objective is then to find $\phi_i^o \in H_o^1(I)$ such that

$$\forall \tau \in H_o^1(I): \quad B_K(\phi_i^o, \tau) = -B_K(N_i, \tau) \tag{3.46}$$

holds for $i = 1, 2$, resp., (cf. [20, pp. 16–17]).

The real bilinear form

$$B_K(u, v) = \int_I u'v' - K^2 \int_I uv$$

is symmetric and coercive (cf. Lemma 2.2). We can find uniquely defined "one element solutions" N_1^p, N_2^p by solving the following: find $\phi_i^p \in V_h = S^p(I) \cap H_o^1(I)$ s.t.

$$\forall \sigma \in V_h: \quad B_K(\phi_i^p, \sigma) = -B_K(N_i, \sigma) \tag{3.47}$$

holds.

We now show (3.41), (3.42). Writing (applying the usual summation convention for $j = 3, \ldots, p+1$)

$$N_1^p(\xi) = N_1(\xi) + a_1^j N_j(\xi),$$
$$N_2^p(\xi) = N_2(\xi) + a_2^j N_j(\xi)$$

we find from (3.31) the vectors $\{a_1\}, \{a_2\}$ to be

$$\{a_1\} = -\left[B_{22}\right]^{-1}\left[B_{21}\right]\begin{Bmatrix} 1 \\ 0 \end{Bmatrix}; \quad \{a_2\} = -\left[B_{22}\right]^{-1}\left[B_{21}\right]\begin{Bmatrix} 0 \\ 1 \end{Bmatrix}. \tag{3.48}$$

It is now easy to see that

$$[CB] = \begin{bmatrix} B_K(N_1^p, N_1^p) & B_K(N_1^p, N_2^p) \\ B_K(N_2^p, N_1^p) & B_K(N_2^p, N_2^p) \end{bmatrix}. \tag{3.49}$$

Indeed,

$$\begin{aligned}
B_K(N_1^p, N_1^p) &= B_K(N_1, N_1) + B_K(N_1, a_1^m N_m) + B_K(a_1^l N_l, N_1) + B_K(a_1^l N_l, a_1^m N_m) \\
&= B_K(N_1, N_1) + 2\{B_K(N_1, N_m)\}\{a_1\} + \{a_1\}^T[\{B_K(N_l, N_m)\}]\{a_1\} \\
&= B_K(N_1, N_1) - 2\left(\begin{Bmatrix} 1 \\ 0 \end{Bmatrix}^T\left[B_{12}\right]\right)\left(\left[B_{22}\right]^{-1}\left[B_{21}\right]\begin{Bmatrix} 1 \\ 0 \end{Bmatrix}\right) \\
&\quad + \left(\begin{Bmatrix} 1 \\ 0 \end{Bmatrix}^T\left[B_{21}\right]^T\left[B_{22}\right]^{-T}\right)\left[B_{22}\right]\left(\left[B_{22}\right]^{-1}\left[B_{21}\right]\begin{Bmatrix} 1 \\ 0 \end{Bmatrix}\right) \\
&= B_{11}[1,1] - \begin{Bmatrix} 1 \\ 0 \end{Bmatrix}^T\left[B_{12}\right]\left[B_{22}\right]^{-1}\left[B_{21}\right]\begin{Bmatrix} 1 \\ 0 \end{Bmatrix} \\
&= CB[1,1]
\end{aligned}$$

and so forth for $CB[1,2], CB[2,1]$, and $CB[2,2]$. In the equation chain above, we have repeatedly used both the symmetry of the form B_K and the local stiffness matrix.

This validates (3.49). Transforming back to global coordinates and assembling, we obtain (3.41), (3.42).

In the proof of Theorem 3.2. we will also use the following result.

LEMMA 3.1. *Let t_1, t_2 and N_1^p, N_2^p be exact and finite element solutions to* (3.43) *with Dirichlet data* (3.44) *or* (3.45)*, resp.*

Then for $i = 1,2; j = 1,2$

$$B_K(t_i - N_i^p, t_j - N_j^p) = B_K(N_i^p, N_j^p) - B_K(t_i, t_j) \tag{3.50}$$

holds.

Proof. Let us specify in (3.46)

$$i = 1, \tau = \phi_1^o: \quad B_K(N_1, \phi_1^o) = -B_K(\phi_1^o, \phi_1^o), \tag{3.51}$$

$$i = 2, \tau = \phi_1^o: \quad B_K(N_2, \phi_1^o) = -B_K(\phi_2^o, \phi_1^o), \tag{3.52}$$

$$i = 1, \tau = \phi_2^o: \quad B_K(N_1, \phi_2^o) = -B_K(\phi_1^o, \phi_2^o), \tag{3.53}$$

and, similarly, in (3.47)

$$i = 1, \sigma = \phi_1^p: \quad B_K(N_1, \phi_1^p) = -B_K(\phi_1^p, \phi_1^p), \tag{3.54}$$

$$i = 2, \sigma = \phi_1^p: \quad B_K(N_2, \phi_1^p) = -B_K(\phi_2^p, \phi_1^p), \tag{3.55}$$

$$i = 1, \sigma = \phi_2^p: \quad B_K(N_1, \phi_2^p) = -B_K(\phi_1^p, \phi_2^p). \tag{3.56}$$

Furthermore, since $S_o^p \subset H_o^1$,

$$i = 1, \tau = \phi_1^p: \quad B_K(N_1, \phi_1^p) = -B_K(\phi_1^o, \phi_1^p), \tag{3.57}$$

$$i = 2, \tau = \phi_1^p: \quad B_K(N_2, \phi_1^p) = -B_K(\phi_2^o, \phi_1^p) \tag{3.58}$$

which, together with (3.54) and (3.55), shows

$$B_K(\phi_1^o, \phi_1^p) = B_K(\phi_1^p, \phi_1^p), \tag{3.59}$$

$$B_K(\phi_2^o, \phi_1^p) = B_K(\phi_2^p, \phi_1^p). \tag{3.60}$$

We now can check the statement of the lemma by direct computation. For $i = j = 1$ the left-hand side is

$$\begin{aligned} B_K(t_1 - N_1^p, t_1 - N_1^p) &= B_K(N_1 + \phi_1^o - N_1 - \phi_1^p, N_1 + \phi_1^o - N_1 - \phi_1^p) \\ &= B_K(\phi_1^o, \phi_1^o) - 2B_K(\phi_1^o, \phi_1^p) + B_K(\phi_1^p, \phi_1^p) \\ &= B_K(\phi_1^o, \phi_1^o) - B_K(\phi_1^p, \phi_1^p) \end{aligned}$$

by (3.59). The right-hand side is

$$\begin{aligned} B_K(N_1^p, N_1^p) - B_K(t_1, t_1) &= B_K(N_1 + \phi_1^p, N_1 + \phi_1^p) - B_K(N_1 + \phi_1^o, N_1 + \phi_1^o) \\ &= B_K(N_1, N_1) + 2B_K(N_1, \phi_1^p) + B_K(\phi_1^p, \phi_1^p) \\ &\quad - B_K(N_1, N_1) - 2B_K(N_1, \phi_1^o) - B_K(\phi_1^o, \phi_1^o) \\ &= B_K(\phi_1^o, \phi_1^o) - B_K(\phi_1^p, \phi_1^p), \end{aligned}$$

where we have used (3.54) and (3.51).

This validates the statement for $i = j = 1$. The computation of the remaining cases $i = 1, j = 2$ and $i = 2, j = 1, 2$ is entirely analogous. $\square$

We are ready to give the proof of Theorem 3.2.

Proof (*Theorem* 3.2). We will show that, neglecting higher-order terms of (kh),

$$|\cos(k'h) - \cos(kh)| \le Ck^2h^2 \frac{C_a^2(p)}{4}\left(\left(\frac{hk}{2p}\right)^{2p}\right) \tag{3.61}$$

holds with an independent constant C. Assuming this relation for the moment we continue with ($C_1(p) := CC_a^2(p)/4$)

$$\left|2\sin\frac{k'+k}{2}h\sin\frac{k'-k}{2}h\right| = |\cos(kh) - \cos(k'h)| \le k^2h^2C_1(p)\left(\frac{hk}{2p}\right)^{2p}.$$

Let $k'h - kh = \varepsilon$; then

$$\left|2\sin\frac{\varepsilon}{2}\right| \le \left|\frac{k^2h^2C_1(p)\left(\frac{hk}{2p}\right)^{2p}}{\sin(kh+\frac{\varepsilon}{2})}\right| \le \frac{k^2h^2C_1(p)\left(\frac{hk}{2p}\right)^{2p}}{kh}.$$

Since by assumption $(hk/2p)^{2p} \ll 1$ we see that ε is small and we may neglect higher-order terms in the Taylor expansion of $2\sin\frac{\varepsilon}{2}$. Thus we obtain

$$\varepsilon \le hkC_1(p)\left(\frac{hk}{2p}\right)^{2p}$$

and replacing $\varepsilon = k'h - kh$ we conclude the statement of the theorem.

We now have to prove (3.61) which we write in the form

$$\left|\frac{S_p}{T_p} - \frac{S_o}{T_o}\right| \le k^2h^2C_1(p)\left(\frac{hk}{2p}\right)^{2p}. \tag{3.62}$$

Consider any internal element Δ being mapped on the master element $I = (-1, 1)$. On I, the exact homogeneous solution of the VP (1.4) is represented by

$$u_{ex}(\xi) = u_{ex}^1 t_1(\xi) + u_{ex}^2 t_2(\xi),$$

where

$$t_1(\xi) = -\frac{\sin K\xi}{2\sin K} + \frac{\cos K\xi}{2\cos K},$$

$$t_2(\xi) = \frac{\sin K\xi}{2\sin K} + \frac{\cos K\xi}{2\cos K}$$

are solutions of the BVPs (3.43), (3.44) or (3.43), (3.45), resp., with $K = kh/2$.

The finite element solution is written on I as

$$u_{fe}(\xi) = u_{fe}^1 N_1^p(\xi) + u_{fe}^2 N_2^p(\xi),$$

where N_1^p, N_2^p are approximate solutions to the BVPs (3.43), (3.44) or (3.43), (3.45) as discussed in the preliminaries.

By continuity of B_K,

$$|B_K(t_l - N_l, t_m - N_m)| \le (1 + K^2)|t_l - N_l|_1|t_m - N_m|_1,$$

and now applying (3.9) we have

$$B_K(t_l - N_l, t_m - N_m) \le \frac{C_a(p)^2}{p^{2p}}\,|t_l|_{p+1}|t_m|_{p+1} \tag{3.63}$$

for $l, m = 1, 2$.

By direct computation,

$$|t_l|_{p+1}^2 = \begin{cases} K^{2p+2}\|t_l\|^2 & \text{if} \quad p \quad \text{is odd,} \\ K^{2p}|t_l|_1^2 & \text{if} \quad p \quad \text{is even} \end{cases} \tag{3.64}$$

with

$$\|t_l\|^2 = \frac{2}{3} + \mathcal{O}(K^2); \quad |t_l|_1^2 = \frac{1}{2} + \mathcal{O}(K^4).$$

Here and in the following, $\mathcal{O}(K^2)$ means an expression of the form $C_1K^2 + C_2K^4 + \cdots$ with constants C_i not depending on h, k, and p.

Also by direct computation,

$$B_K(t_1, t_1) = B_K(t_2, t_2) = \frac{1}{2} + \mathcal{O}(K^2) \tag{3.65}$$

and

$$B_K(t_1, t_2) = B_K(t_2, t_1) = -\frac{1}{2} + \mathcal{O}(K^2). \tag{3.66}$$

Finally, we recall from Lemma 3.1 that

$$B_K(t_l - N_l, t_m - N_m) = B_K(N_l, N_m) - B_K(t_l, t_m) \tag{3.67}$$

holds for $l, m = 1, 2$.

Returning to the proof of (3.62), we have

$$\begin{aligned} |S_pT_o - S_oT_p| &= |B_K(t_1, t_1)B_K(N_1, N_2) - B_K(t_1, t_2)B_K(N_1, N_1)| \\ &= |B_K(t_1, t_1)\left(B_K(t_1 - N_1, t_2 - N_2) - B_K(t_1, t_2)\right) \\ &\quad -B_K(t_1, t_2)\left(B_K(t_1 - N_1, t_1 - N_1) - B_K(t_1, t_1)\right)| \\ &= |B_K(t_1, t_1)B_K(t_1 - N_1, t_2 - N_2) \\ &\quad -B_K(t_1, t_2)B_K(t_1 - N_1, t_1 - N_1)|. \end{aligned} \tag{3.68}$$

Here, we applied (3.67) to expand the expression on the right-hand side. Thus

$$\begin{aligned} |S_pT_o - S_oT_p| \leq\ & |B_K(t_1, t_1)||B_K(t_1 - N_1, t_2 - N_2)| \\ & +|B_K(t_1, t_2)||B_K(t_1 - N_1, t_1 - N_1)| \end{aligned}$$

and applying (3.65), (3.66) leads to

$$|S_pT_o - S_oT_p| \leq \left(\frac{1}{2} + \mathcal{O}(K^2)\right)\left((1 + K^2)\frac{C_a(p)^2}{p^{2p}}\left(|t_1|_{p+1}|t_2|_{p+1} + |t_1|_{p+1}^2\right)\right).$$

For odd p we then have directly by (3.64)

$$\begin{aligned} |S_pT_o - S_oT_p| &\leq \left(\frac{1}{2} + \mathcal{O}(K^2)\right)\left((1 + K^2)\frac{C_a(p)^2}{p^{2p}}K^{2p+2}\left(\frac{4}{3} + \mathcal{O}(K^2)\right)\right) \\ &\leq \frac{C_a(p)^2}{p^{2p}}K^{2p+2}, \end{aligned} \tag{3.69}$$

where we neglected terms of order $\mathcal{O}(K^2)$. Then also (3.62), and hence the statement, holds since

$$\left|\frac{S_p}{T_p} - \frac{S_o}{T_o}\right| = \frac{|S_pT_o - S_oT_p|}{|T_oT_p|}$$

and it can easily be seen that

$$|T_oT_p| = |B_K(t_1,t_2)||B_K(t1 - N1, t_2 - N_2) + B_K(t_1,t_2)|$$

is bounded from below by a constant independently from K, p.

If p is even, then we insert (3.65), (3.66) into (3.68) to get

$$\begin{aligned}|S_pT_o - S_oT_p| &\leq \left(\frac{1}{2} + \mathcal{O}(K^2)\right)|B_K(t_1 - N_1, t_2 - N_2)| \\ &\quad + \left(\frac{1}{2} + \mathcal{O}(K^2)\right)|B_K(t_1 - N_1, t_1 - N_1)| \\ &= \frac{1}{2}\,|B_K(t_1 - N_1, t_1 + t_2 - (N_1 + N_2))| \\ &\quad + \mathcal{O}(K^2)\,|B_K(t_1 - N_1, t_2 - N_2) + B_K(t_1 - N_1, t_1 - N_1)|\,.\end{aligned}$$

For the first term in this equation we have

$$\begin{aligned}|B_K(t_1 - N_1, t_1 + t_2 - (N_1 + N_2))| &\leq \frac{C_a(p)^2}{p^{2p}}|t_1|_{p+1}\,|(t_1 + t_2)|_{p+1} \\ &= \frac{C_a(p)^2}{p^{2p}}K^{2p}|t_1|\left|\frac{\cos K\xi}{\cos K}\right|_1 \\ &\leq CC_a(p)^2\left(\frac{hk}{2}\right)^2\left(\frac{hk}{2p}\right)^{2p}\end{aligned}$$

with C not depending on K, p. Again, terms of order $\mathcal{O}(K^2)$ have been neglected. For the second term, a similar estimate follows directly from (3.63), (3.64). Thus the estimate (3.69) holds also for even p, and the statement follows by similar argument. $\square$

3.4. Discrete stability. In part I, when investigating the h version with $p = 1$, we proved the stability estimate

$$\|u'_{fe}\| \leq C\|f\|$$

and showed that the inf-sup constant on the discrete subspace is $\gamma_h = Ck^{-1}$. A standard corollary then yields the stability estimate

$$\|u'_{fe}\| \leq Ck\|f^{-1}\|.$$

We will now show that both results carry over to higher p.

Further, we will define in this subsection a specific data subspace that we will encounter in the error analysis. In this subspace we will prove stability with respect to higher integrals of the data—the discrete analogon to Theorem 2.2.

Hence let $u_{fe} \in S_h^p$ be the finite element solution to the VP (1.4) for data $f \in L^2(\Omega)$. We write

$$u_{fe} = u_h + u_p,$$

where u_h is based on the nodal shape functions and u_p on the internal ones. From the definition of the shape functions N_j and orthogonality property of the Legendre polynomials it follows that

$$\|u'_{fe}\|^2 = \|u'_h\|^2 + \|u'_p\|^2. \tag{3.70}$$

Furthermore, u_p satisfies on each element Δ (3.18). Transforming $\Delta \to I^+ = (0,1)$, we get

$$\forall w \in S_o^p(\Delta): \quad B_K(u_p, w) = h^2(f, w) + K^2(u_h, w). \tag{3.71}$$

We now prove a first stability lemma on $\|u'_p\|$ for data $f \in L^2(\Omega)$.

LEMMA 3.2. *Let u_{fe} be the finite element solution to the VP* (1.4) *with data $f \in L^2(\Omega)$. Assume that $hk \leq \alpha < \pi$.*

Then

$$\|u'_{fe}\| \leq C\|f\| \tag{3.72}$$

holds with a constant C independent of h, k, and p.

Remark 9. The assumption $hk < \alpha \leq \pi$ guarantees invertibility of the submatrix B_{22}^j in the condensation procedure since π is obviously a lower bound for the eigenvalues of B_K, (3.71). For $kh > \pi$, one has to consider the discrete eigenvalues as well as the "stopping bands" [22], i.e., intervals where $|S_p(K)/T_p(K)| > 1$. For details, see [17] where focus is on computational aspects of *h-p* finite element solutions for the Helmholtz equation.

Proof. Let $u_{fe} = u_h + u_p$ as defined above. We know (cf. part I, Lemma 3) by straight estimation of the discrete Green's function representation that $\|u'_h\| \leq C_1\|f\|$, where C_1 does not depend on h, k, and p. Applying (2.5) from Lemma 2.2 to (3.71), we have

$$\|u'_p\| \leq D\left(h^2\|f\| + K^2\|u_h\|\right),$$

where $D = \pi/(\pi^2 - \alpha^2)$. Back transform to Δ then yields

$$\|u'_p\|_\Delta \leq Dh\left(\|f\|_\Delta + k^2\|u_h\|_\Delta\right).$$

Summing up and applying Schwarz inequality, we get

$$\|u'_p\| \leq Dh\left(\|f\| + k^2\|u_h\|\right) \leq \|f\|h\left(D + C_2 k\right),$$

where we applied $k\|u_h\| \leq C_2\|f\|$ with C_2 not depending on h, k, p. Thus

$$\|u'_{fe}\| \leq \left(C_1 + Dh + C_2 kh\right)\|f\|,$$

and the statement is readily obtained by neglecting Dh and setting $C = C_1 + C_2\alpha$. The proof is completed. ∎

We now prove the inf-sup condition for $V_h = S_h^p(\Omega)$.

LEMMA 3.3 (*Discrete* inf-sup *condition*). *Let $V_h = S_h^p(\Omega)$ and let $B : V_h \times V_h \to \mathbf{C}$ be the sesquilinear form defined by* (1.4).

Then, if h is such that $hk \leq \alpha < \pi$,

$$\inf_{u \in V_h} \sup_{v \in V_h} \frac{|B(u,v)|}{|u|_1 |v|_1} \geq \frac{C}{k}, \tag{3.73}$$

where C does not depend on h, k, and p.

Proof. The argument is similar to the case $p = 1$. For arbitrarily fixed $u \in V_h$ set $v := u + z$, where $z \in V_h$ is the solution to the VP

$$\forall w \in V_h : \qquad B(w, z) = k^2(w, u).$$

Let $z := z_h + z_p$ as above. In part I it was shown that

$$\|z_h'\| \le C_2 k \left(\frac{k}{k'}\right) \|u'\|.$$

The ratio k/k' is bounded for $kh < \pi$ and $p \ge 2$ by Theorem 3.2. Further, z_p solves locally, i.e., for all $w \in S_o^p(\Delta)$,

$$B_\Delta(z_p, w) = k^2(u + z_h, w)_\Delta.$$

Applying Lemma 2.2 and Remark 2,

$$\|z_p'\| \le C_1 k(\|u\| + \|z_h\|).$$

Again by discrete Green's function representation, we can show that $\|z_h\| \le C_2\|u'\|$, with C_2 not depending on $h, k,$ and p. Applying a Poincaré inequality for $\|u\|$, we conclude

$$\|z'\| \le C_3 k \|u'\|$$

and the statement is concluded, using the particular choice of z (cf. part I). ∎

It follows by standard theory that $|u|_1 \le Ck\|f^{(-1)}\|$.

We collect both stability estimates in the following proposition.

THEOREM 3.3 (*Stability of the finite element solution* I). *Let $f \in L^2(\Omega)$ and let $u_{fe} \in V_h = S_h^p(\Omega)$ be the finite element solution to the VP* (1.4).

Then, if h is such that $hk \le \alpha < \pi$, the stability estimates

$$|u_{fe}|_1 \le C_1 \|f\| \tag{3.74}$$

and

$$|u_{fe}|_1 \le C_2 k \|f^{(-1)}\| \tag{3.75}$$

hold for C_1, C_2 not depending on $h, k,$ and p.

Next we formulate a dual stability property that we will use to prove a preasymptotic error estimate for the p degree finite element solution. To this end, we define a specific data subspace.

DEFINITION 3.1. *For integer $l \ge 0$ we define a subspace $F_o^l(\Omega) \subseteq L^2(\Omega)$ by*

$$F_o^l(\Omega) = \left\{ f \in L^2(\Omega) \;\middle|\; f^{(-i)}|_{X_h} = 0 \quad \textit{for} \quad i = 1, \ldots, l \right\}$$

with $F_o^o(\Omega) := L^2(\Omega)$.

Observe that for this space we can show, by adding up local Poincaré inequalities,

$$\|f^{(-i)}\| \le h \|f^{(-i+1)}\|. \tag{3.76}$$

In particular, for $f \in F_o^1$ and $hk \le \alpha$, inequality (3.74) directly follows from (3.75).

For the discrete data obtained from $f \in F_o^l(\Omega)$ by the finite element procedure, the following proposition is true.

LEMMA 3.4. *Consider the VP* (1.4) *on $V_h = S_h^p(\Omega)$ with data $f \in F_o^{p-1}(\Omega)$. Let Δ_j be an arbitrary finite element and let $\{\tilde{r}_{j-1}, \tilde{r}_j\}^T$ be the condensed right-hand side* *vector given by* (3.26)*. Assume further that the stepwidth h is sufficiently small so that $hk \leq \alpha < \pi$.*

Then

$$|\tilde{r}_j| \leq C_d(p,m) h^{\frac{1}{2}} k^m \|f^{(-m)}\|_{\Delta_i} \tag{3.77}$$

holds for even $m = 0, 2, \dots, \leq p-1$ with

$$C_d(p,0) = 1$$

and

$$C_d(p,m) = C_1 + C_2 \alpha^{p-m}\, 2^{\frac{(m-1)}{2}} \frac{(p+1)!(p+1)}{((p-m+1)!)^2}, \quad m \geq 2,$$

where C_1, C_2 do not depend on $h, k,$ and p.

Proof. First let $m = 0$. We omit in the notation the element number j and renumber formally $\{\tilde{r}_{j-1}, \tilde{r}_j\} \to \{\tilde{r}_1, \tilde{r}_2\}$. Equation (3.26) then reads

$$\left\{ \begin{array}{c} \tilde{r}_1 \\ \tilde{r}_2 \end{array} \right\} = \left\{ \begin{array}{c} r_1 \\ r_2 \end{array} \right\} - [B_{12}]\,[B_{22}]^{-1} \left\{ \begin{array}{c} r_3 \\ \vdots \\ r_{p+1} \end{array} \right\}.$$

It is straightforward to show that $r_1 = (f, N_1)_\Delta$ and $r_2 = (f, N_2)_\Delta$ satisfy $|r_j| \leq C_1 h^{\frac{1}{2}} \|f\|_\Delta,\ j = 1, 2.$

Define $\{y\} := [B_{22}]^{-1}\{r_3, \dots, r_{p+1}\}^T$. This vector is the discrete solution to the VP

$$\forall w \in S_o^p(\Delta): \quad B_\Delta(y, w) = (f, w)_\Delta$$

and by Lemma 2.2 and Remark 2,

$$\|y'\| \leq h \frac{\pi}{\pi^2 - \alpha^2} \|f\|_\Delta.$$

By simple computation, $\|y'\|_\Delta^2 \geq \frac{1}{h} |\{y\}|^2$, hence

$$|\{y\}| \leq C_1 h^{\frac{3}{2}} \|f\|_\Delta,$$

where $C_1 = \pi/(\pi^2 - \alpha^2)$ does not depend on $h, k,$ and p.

Now consider the term $\{z\} := [B_{12}][B_{22}]^{-1}\{r_3, \dots, r_{p+1}\}^T = [B_{12}]\{y\}$. The coefficients of the matrix $[B_{12}]$ are $b_{ij} = B_\Delta(N_i, N_j)$ with $i = 1, 2$ and $j = 3, \dots, p+1$. For these i, j we can show, using integration by parts, that $(N_i', N_j')_\Delta = 0$ and hence $b_{ij} = -k^2 (N_i, N_j)_\Delta$. Obviously, the euclidian norm of the rows $\{b_i\}, i = 1, 2$ can be bounded as

$$|\{b_i\}| \leq C_2 k^2 h.$$

The constant C_2 does not depend either on $h, k,$ or p since the bandwidth of the local mass matrix does not increase with p for $p \geq 3$ (cf. [20, p. 46]).

Thus, with the previous estimate of $\|\{y\}\|$, for $i = 1, 2$,

$$\begin{aligned} |z_i| = |b_{ij} y_j| &\leq |\{b_i\}|_\Delta |\{y\}|_\Delta \\ &\leq C_1 C_2 k^2 h^{\frac{5}{2}} \|f\|_\Delta \\ &\leq C_3 h^{\frac{1}{2}} \|f\|_\Delta \end{aligned} \tag{3.78}$$

with $C_3 = C_1 C_2 \alpha^2$. Together with the observation for r_1, r_2, the last estimate proves the statement for $m = 0$.

Next let $m \geq 2$ and assume for convenience that Δ has been mapped to $I^+ = (0,1)$. Then, for $j = 1, 2$,

$$\tilde{r}_j = h \int_0^1 f(\theta) \varphi_j(\theta) d\theta, \tag{3.79}$$

where $\varphi_1, \varphi_2 \in S^p_{(o}(I^+)$ or $S^p_{o)}(I^+)$, resp., are the "one-element" solutions to the homogeneous Helmholtz equation

$$u'' + k^2 h^2 u = 0 \tag{3.80}$$

with the boundary conditions

$$u(0) = 1, \quad u(1) = 0 \tag{3.81}$$

or

$$u(0) = 0, \quad u(1) = 1, \tag{3.82}$$

resp. The exact solutions t_1 and t_2 of these BVPs are

$$t_1(\theta) = \cos kh\theta - \cot kh \sin kh\theta, \tag{3.83}$$

$$t_2(\theta) = \frac{\sin kh\theta}{\sin kh}. \tag{3.84}$$

Integration by parts in (3.79) leads to

$$|\tilde{r}_j| = h \left| \int_0^1 f^{(-m)}(\theta) \varphi^{(m)}(\theta) d\theta \right|;$$

no boundary terms occur for $m \leq p - 1$ since $f \in F_o^{p-1}(\Omega)$. Consequently we have for $j = 1, 2$

$$|\tilde{r}_j| \leq h \, \|f^{(-m)}\|_{I^+} \|\varphi_j^{(m)}\|_{I^+}.$$

For the estimation of $\|\varphi_j^{(m)}\|_{I^+}$, we define $\chi_j \in S^p_{(o}(I^+)$ or $S^p_{o)}(I^+)$, resp., by

$$\chi_1(\theta) := \tau_1(\theta) - \theta \tau_1(1)$$

and

$$\chi_2(\theta) := \tau_2(\theta) + \theta(1 - \tau_2(1)),$$

where $\tau_1, \tau_2 \in S^p(I^+)$ are the Taylor polynomials of order p in $\theta_o = 0$ for $t_1(\theta)$ and $t_2(\theta)$, resp. Now trivially

$$\|\varphi_j^{(m)}\|_{I^+} \leq \|\varphi_j^{(m)} - \chi_j^{(m)}\|_{I^+} + \|\chi_j^{(m)}\|_{I^+}.$$

As to the second member on the right-hand side, it can be shown by direct computation that for even $m \geq 2$

$$\|\chi_j^{(m)}\|_{I^+} \leq C_1 (hk)^m \tag{3.85}$$

holds with a constant C_1 not depending on h, k, and p. Turn to the estimation of $\|\varphi_j^{(m)} - \chi_j^{(m)}\|_{I^+}$. For $m = 1$, we have

$$|\varphi_j - \chi_j|_1 \le |t_j - \varphi_j|_1 + |t_j - \chi_j|_1 \le C_2 |t_j - \chi_j|_1 \tag{3.86}$$

by Céa's lemma.[4] By construction of χ_j,

$$|t_j - \chi_j|_1 = |t_j - \tau_j|_1 + \mathcal{O}\left(\frac{(hk)^p}{p!}\right) \le C_3 \left(\frac{(hk)^p}{p!}\right), \tag{3.87}$$

where C_3 does not depend on h, k, and p.

For $m \ge 2$, we use repeatedly the inverse inequality (3.2) to relate

$$\|\varphi_j^{(m)} - \chi_j^{(m)}\|_{I^+} \le \left(\frac{(p+1)!}{(p-m+1)!}\right)^2 2^{\frac{(m-1)}{2}} |\varphi_j - \chi_j|_1.$$

Inserting this estimate into (3.86) vs. (3.87), we conclude that

$$\|\varphi_j^{(m)} - \chi_j^{(m)}\|_{I^+} \le C_4 \frac{((p+1)!)^2 \, 2^{\frac{(m-1)}{2}}}{((p-m+1)!)^2 p!} (hk)^p$$

where $C_4 = C_2 C_3$ does not depend on h, k, and p.

Thus

$$\|\varphi_j^{(m)}\|_{I+} \le C_1 (hk)^m + C_4 \frac{(p+1)!(p+1)}{((p-m+1)!)^2} \, 2^{\frac{(m-1)}{2}} (hk)^p$$

and, finally,

$$|\tilde{r}_j| \le h(hk)^m \left(C_1 + C_4 \frac{(p+1)!(p+1)}{((p-m+1)!)^2} \, 2^{\frac{(m-1)}{2}} (hk)^{p-m}\right).$$

The statement now follows by back transform $I^+ \to \Delta$. The proof is completed. ∎

Remark 10. For odd m this statement holds true with the additional assumption that hk is bounded from below, i.e., $0 < \beta \le hk$.

Let us show this for $m = 1$ and $j = 2$. We have

$$\|\chi_2'\| \le \|\tau_2'\| + \|1 - \tau_2(1)\|.$$

By construction, the second member is $\mathcal{O}(\frac{(hk)^p}{p!})$. Also $\tau_2'(\theta) = t_2'(\theta) + \mathcal{O}(\frac{(hk)^p}{p!})$. Hence, neglecting terms of higher order (note that by assumption $p > m$),

$$\begin{aligned} \|\chi_2'\| \le \|t_2'\| \; &= \; \frac{kh}{\sin kh} \|\cos kh\theta\|_{I^+} \\ &\le C_1 \left(1 + \mathcal{O}\left(\frac{(kh)^2}{2}\right)\right). \end{aligned}$$

Furthermore, $|\varphi_2 - \chi_2|_1 = \mathcal{O}(\frac{(hk)^p}{p!})$. Hence $\|\varphi_2'\|_{I^+} \le C_2$ and

$$\begin{aligned} |\tilde{r_2}| &\le C_2 \, h \, \|f^{(-1)}\|_{I^+} \\ &\le C_2 h^{-\frac{1}{2}} \|f^{(-1)}\|_\Delta \end{aligned}$$

[4]To be precise, we write χ, φ as sums of linear and bubble functions and apply the statement of Céa's lemma on an appropriate subspace $V_h \subset H_o(I^+)$.

after back transform $I^+ \to \Delta$. Multiplying and dividing now by $kh^{1/2}$, we get

$$\begin{aligned} |\tilde{r_2}| &\leq C_2 \frac{h^{\frac{1}{2}} k}{hk} \|f^{(-1)}\|_{I^+} \\ &\leq C h^{\frac{1}{2}} k \|f^{(-1)}\|_{\Delta} \end{aligned}$$

with $C = C_2 \beta^{-1}$. This shows our case; the argument for $j = 1$ or higher m is similar.

COROLLARY 3.1. *For the norm of the discrete right-hand side R_h, the estimate*

$$\|R_h\|_h \leq C_d(p,m)\, h\, k^m \|f^{(-m)}\| \tag{3.88}$$

holds for $m = 0, \ldots, \leq p-1$.

Proof. By definition (cf. part I),

$$\|R_h\|_h^2 = h \sum_{i=1}^{n} |R_i|^2,$$

where the coefficients R_i are given by (3.30). By Lemma 3.4,

$$\|R_h\|_h \leq C_d(p,m) \left(h \sum_{i=1}^{n} h k^{2m} \|f^{(-m)}\|_{\Delta_i}^2 \right)^{\frac{1}{2}}$$

and the statement readily follows. ∎

We now prove a proposition on dual stability if the data $f \in F_o^m(\Omega)$.

THEOREM 3.4 (*Stability of the finite element solution* II). *Let $u_{fe} \in V_h = S_h^p(\Omega)$ be the finite element solution to the VP* (1.4) *with data $f \in F_o^m(\Omega)$, where m is even, $m \leq p-1$. Assume that $0 < kh \leq \alpha < \pi$.*

Then

$$|u_{fe}|_1 \leq C_d(p,m) k^m \|f^{(-m)}\| + C_1 \|f^{(-1)}\|, \tag{3.89}$$

where C_d is the discrete dual stability constant (cf. Lemma 3.4*) and C_1 does not depend on h, k, and p.*

Proof. In light of Theorem 3.3. we only need to prove the statement for $m \geq 2$. As before, we write $u_{fe} = u_h + u_p$. Then

$$\|u_{fe}'\| \leq \|u_h'\| + C_1(\|f^{(-1)}\| + k^2 \|u_h^{(-1)}\|)$$

by Lemma 3.2. It can be shown, using the Green's function representation of u_h, that

$$\|u_h'\| + C_1 k^2 \|u_h^{(-1)}\| \leq \frac{C_2}{h} \|R\|_h,$$

where C_2 does not depend on h, k, and p. On the other hand,

$$\|R\| \leq C_d(p,m) h k^m \|f^{(-m)}\|,$$

from Lemma 3.4, and the statement follows. The proof is completed ∎

Remark 11. The stability theorem holds for odd m with the additional assumption that $0 < \beta \leq hk$—cf. the previous remark.

3.5. Error estimates for the finite element solution. Throughout this subsection, we denote by l the regularity of the exact solution u; i.e., we assume that $u \in V^l := H^{l+1}(\Omega) \cap V$ with $V = H^1_{(o}(\Omega)$.

In part I we proved that the finite element solution is asymptotically quasi optimal. The same result holds for higher approximation. However, the range of asymptotic behavior, as taken with respect to the meshsize h^{-1}, grows with p.

THEOREM 3.5 (*Asymptotic estimate*). *Let $l \geq 1, p \geq 1$. Further let $u \in V^l$ and $u_{fe} \in V_h = S_h^p(\Omega)$ be the exact and finite element solutions to the VP* (1.4), *resp.*

Then, if $k^2 h/p$ is sufficiently small, the quasi-optimal estimate

$$|u - u_{fe}|_1 \leq C \inf_{v \in V_h} |u - v|_1 \tag{3.90}$$

holds with

$$C = \left(\frac{4 + \left(\frac{hk}{2p}\right)^2}{\frac{1}{2} - 6k^2 \left(\frac{hk}{p}\right)^2 \left(1 + \sqrt{\frac{3}{2}} \left(\frac{hk}{p}\right)^2\right)} \right)^{\frac{1}{2}}, \tag{3.91}$$

provided the denominator of C is positive.

Proof. Denote $e := u - u_{fe}$ and let z be the solution to the VP

$$\forall w \in V : \quad B(w, z) = (w, e). \tag{3.92}$$

This problem has a unique solution $z \in H^3(\Omega)$. We can show (part I, Theorem 3) that

$$\|e\|^2 \leq 2 \left(|z - w|_1 |e|_1 + k^2 \|z - w\| \|e\|\right)$$

holds for all $w \in V$. In particular, for $w = s$, where s is the approximation of z as constructed in Theorem 3.1., we have

$$\|e\|^2 \leq 2 \left(\frac{h}{2p} |z|_2 |e|_1 + k^2 \left(\frac{h}{2p}\right)^3 |z|_3 \|e\| \right).$$

We now apply the stability estimates (part I, Lemma 1, and (2.24), Theorem 2.1) to get

$$|z|_2 \leq (1 + k)\|e\|,$$
$$|z|_3 \leq (1 + 4k)\|e\|_1 \leq \sqrt{\frac{3}{2}}(1 + 4k)|e|_1$$

by Poincaré inequality since $e \in H_{(o}(\Omega)$.

Hence

$$\|e\|^2 \leq 2 \left(\frac{h}{2p}(1 + k) + k^2 (1 + 4k) \left(\frac{h}{2p}\right)^3 \sqrt{\frac{3}{2}} \right) \|e\| \, |e|_1;$$

dividing now by $\|e\|$ and neglecting terms where h is of order higher than k, we arrive at the intermediary result

$$\|e\| \leq \frac{kh}{p} \left(1 + \sqrt{\frac{3}{2}} \left(\frac{kh}{p}\right)^2\right) |e|_1. \tag{3.93}$$

In the next step we use B orthogonality of e to show (cf. again part I, Theorem 3) that

$$\frac{1}{2}|e|_1^2 \le 6k^2\|e\|^2 + 4|u-v|_1^2 + k^2\|u-v\|^2$$

holds for all $v \in V_h$. We choose $v = s$ from Theorem 3.2; then

$$\|u-s\| \le \frac{h}{2p}|u-s|_1$$

and, now applying (3.93), we obtain

$$\frac{1}{2}|e|_1^2 - 6k^2\left(\frac{kh}{p}\right)^2\left(1+\sqrt{\frac{3}{2}}\left(\frac{kh}{p}\right)^2\right)^2|e|_1^2 \le \left(4+\frac{k^2h^2}{4p^2}\right)|u-s|_1^2,$$

hence

$$|e|_1 \le C|u-s|_1,$$

where

$$C = \left(\frac{4+\left(\frac{hk}{2p}\right)^2}{\frac{1}{2}-6k^2\left(\frac{hk}{p}\right)^2\left(1+\sqrt{\frac{3}{2}}\left(\frac{hk}{p}\right)^2\right)}\right)^{\frac{1}{2}}.$$

This completes the proof. $\square$

Remark 12. For piecewise linear approximation we proved (part I, Corollary 2)

$$|u-u_{fe}| \le C_1 \inf_{v\in V_h}|u-v|_1$$

with

$$C_1 = \frac{2\left(1+\left(\frac{hk}{2\pi}\right)^2\right)^{\frac{1}{2}}}{\left(\frac{1}{2}-6C_2^2k^2h^2(1+k)^2\right)^{\frac{1}{2}}},$$

where

$$C_2 = \frac{2}{(1-2(1+k)\frac{k^2h^2}{\pi^2})\,\pi}.$$

Obviously $C_2 \ge 2$, hence $k^4h^2 < \frac{1}{12\cdot 4}$ is necessary for welldefiniteness of C_1. A similar computation yields $k^4h^2 < \frac{p^2}{12}$ as a necessary condition for welldefiniteness of C in the theorem above.

Remark 13. In the form given in Theorem 3.5, the quasi-optimal estimate holds independently of the regularity of the solution u. The order of convergence, in terms of hp^{-1}, is obtained by introducing the approximation property of the subspace V_h from Theorem 3.1. For given p, the maximal order of convergence

$$|u-u_{fe}|_1 \le C\left(\frac{e}{2}\right)^p(\pi p)^{-\frac{1}{4}}\left(\frac{h}{2p}\right)^p|u|_{p+1}$$

is achieved when the regularity of u is $l \ge p$.

We now proceed to error estimates in the preasymptotic range (i.e., without restrictions on k^2h). Let us first relate the finite element solution to best approximations in $V_h = S_h^p(\Omega)$, as constructed in Theorem 3.1.

LEMMA 3.5. *Let, for $p \geq 1$, u and u_{fe} be the exact and finite element solutions to the VP* (1.4)*, resp., and let $s \in V_h$ be a nodally exact quasi-optimal approximation to u in the sense of Theorem* 3.1*.*

Then $z := u_{fe} - s$ is the finite element solution to the VP (1.4) *with data $k^2(u-s)$.*

Proof. Trivially $z = u_{fe} - u + u - s$, and by B orthogonality of $u - u_{fe}$ to V_h, $B(z,v) = B(u-s,v)$ holds for all $v \in V_h$. The boundary term in $B(u-s,v)$ vanishes due to $u|_{X_h} = s|_{X_h}$. Using local exactness of the integrals of s we show, repeatedly integrating by parts, that the term $((u-s)', v')$ also vanishes. Thus, for all $v \in V_h$,

$$B(u-s,v) = -k^2(u-s,v)$$

which completes the proof. □

It is now straightforward to show a first error estimate.

THEOREM 3.6 (*Preasymptotic estimate* I). *Let, for $1 \leq l \leq p$, $u \in V^l$ and $u_{fe} \in V_h$ be the solution and the finite element solution to the VP* (1.4)*, resp. Assume that $hk \leq \alpha < \pi$.*

Then for $e := u - u_{fe}$

$$|e|_1 \leq C_a(l) \left(1 + C_1 k \left(\frac{kh}{2p}\right)\right) \left(\frac{h}{2p}\right)^l |u|_{l+1} \tag{3.94}$$

holds, where C_1 does not depend on h, k, and p, whereas $C_a(l)$ is the approximation constant (Theorem 3.1*), being at most of order $\left(\frac{e}{2}\right)^p$.*

Proof. Let $z = u_{fe} - u + u - s$ as above. By Theorem 3.3 and the previous lemma, $|z|_1 \leq Ck^2\|u-s\|$, hence

$$|e|_1 = |z + u - s|_1 \leq Ck^2\|u-s\| + |u-s|_1. \tag{3.95}$$

To complete the proof, we insert the appropriate results from the approximation theorem. □

Remark 14. Note that if $k^2h/2p$ is bounded, the error estimate is equivalent to the asymptotic quasi-optimal estimate given in Theorem 3.5. This, again, is an analogy to the h version with $p = 1$ (part I, section 3.6).

The estimate of the previous lemma can be generalized for $p \geq 2$, employing dual stability estimates for the data $k^2(u-s) \in F_o^{p-1}(\Omega)$.

THEOREM 3.7 (*Preasymptotic estimate* II). *Let $1 \leq l \leq p$ and $0 \leq m \leq p$, m even, with $p \geq 2$. Let $u \in V^l$ be the solution to the VP* (1.4) *with data $f \in H^{(l-1)}(\Omega)$ and let $u_{fe} \in V_h$ be the finite element solution to this problem. Assume further that the stepwidth h is such that $hk \leq \alpha < \pi$.*

Then

$$|e|_1 \leq C_a(l) \left[1 + C_1 \left(\frac{kh}{2p}\right)^2 + kC_d(p,m)C_a(m)\left(\frac{kh}{2p}\right)^{m+1}\right] \left(\frac{h}{2p}\right)^l |u|_{l+1} \tag{3.96}$$

holds with C_1 not depending on k, h, p.

Proof. Let $s \in V_h$ be a nodally exact, optimal approximation of u in the sense of Theorem 3.1 and define, as before, $z := u_{fe} - s$. We know that z solves $B(z,v) = -k^2(u-s,v)$ for all $v \in V_h$. The data of this problem is in the space F_o^{p-1}, hence by Theorem 3.4.

$$|z|_1 \leq k^2 \left(C_d(p,m)k^m\|(u-s)^{(-m)}\| + C_1\|(u-s)^{(-1)}\|\right)$$

holds for $m \leq p-1$. Inserting estimate (3.4) from Theorem 3.1,

$$\|(u-s)^{(-m)}\| \leq C_a(l)C_a(m)\left(\frac{h}{2p}\right)^{l+m+1}|u|_{l+1},$$

and we conclude the statement. □

Remark 15. With the additional assumption $0 < \beta \leq kh$, the statement holds also for odd m—cf. Remarks 9, 10. This assumption is consistent with the error estimation in the *preasymptotic* range and with computational application where the magnitude of hk is, for medium and high k, bounded from below by practical considerations.

Remark 16. We obtain estimate I from estimate II by setting $m = 0$, hence II generalizes I.

Let us specify the estimate II for a certain type of solutions, namely, those oscillating with frequency k. These solutions are of practical importance in physical applications in wave propagation and wave scattering; they are, among others, produced by Dirac data (point sources).

Thus, having in mind solutions that essentially behave like $\exp(ikx)$, we define the following.

DEFINITION 3.2. *Let, for $l \geq 1$, $u \in V^l$ be a solution to the VP* (1.4). *We call u an* oscillating *solution if*

$$|u|_{l+1} \leq Dk^l|u|_1 \tag{3.97}$$

holds with a constant D not depending on k.

With this definition, we directly have the following corollary.

COROLLARY 3.2 (*Error estimate for oscillating solutions*). *Let $1 \leq l \leq p$ and $p \geq 2$. Assume that there is given a data $f \in H^{l-1}(\Omega)$ such that the solution $u \in V^l$ to the VP* (1.4) *is oscillating. Assume further that the stepwidth h is such that $hk \leq \alpha < \pi$. Let $u_{fe} \in V_h = S_h^p(\Omega)$ be the finite element solution to the VP* (1.4).

Then the relative error $|\tilde{e}|_1 := |u - u_{fe}|_1/|u|_1$ is bounded by

$$|\tilde{e}|_1 \leq \left(\frac{hk}{2p}\right)^l\left[C_1 + C_2\left(\frac{kh}{2p}\right)^2\right] + kC_3\left(\frac{kh}{2p}\right)^{l+m+1}, \tag{3.98}$$

where $C_1 = DC_a(l), C_2 = EC_a(l)$, and $C_3 = DC_d(p,m)C_a(l)C_a(m)$, with D, E not depending on h, k, and p.

Proof. We introduce the definition of oscillatory behavior into the estimate II (3.96). □

Let us consider special cases of the previous corollary.

1. If $l = p$, we have with $\theta := (hk/2p)$

$$|\tilde{e}|_1 \leq \theta^p(C_1 + C_2\theta^2) + C_3k\theta^{2p}.$$

   This is principally the estimate that was given in the analysis of the h version—cf. Introduction, equation (1.5)—with $p = 1$ and $\theta = kh$. Note also that, formally, the error is written as the sum of best approximation error plus pollution term of the order of the phase lag. However, the constant C_3 depends on p (see the next section).
2. In the case of lower regularity ($l < p$) the pollution for higher p is relatively (i.e., compared to the best approximation order) still smaller as in the case of full regularity. Consider the lowest possible case of Dirac data. Then $l = 1$ and the estimate is

$$|\tilde{e}|_1 \leq \theta(C_1 + C_2\theta^2) + C_3k\theta^{p+1}.$$

In general, the constant C_3 depends on m. Note that, for fixed approximation order p, one is free to choose m in the range of $0, \ldots, p-1$. This can be used to optimize the size of the pollution term $C_3(p,m)(hk/2p)^{l+m+1}$.

Remark 17. We will show in the numerical evaluation that the constants C_1, C_2 are sharp. On the other hand, the theoretically predicted growth in $C_3(p)$ was not observed in several numerical examples. It is an open question whether the dual estimate is sharp in the pollution term.

We conclude this subsection with an error estimate in negative norms. We first show a mapping property; the proposition then readily follows.

LEMMA 3.6. *Let $u \in V^l$ and $u_{fe} \in V_h$ be the exact and the finite element solution to the VP* (1.4). *Then, for $1 \leq m \leq p-1$,*

$$\|e\|_{-m} \leq DC_s(m+1)\, k \left(\frac{hk}{2p}\right)^{m+1} |e|_1, \tag{3.99}$$

where C_s is the stability constant from Theorem 2.1 *and D is a constant not depending on h and k.*

Proof. Since $X = H_o^m(\Omega)$ is a Hilbert space there exists $v_o \in X$ s.t.

$$\|e\|_{-m} = \frac{|(e, v_o)|}{|v_o|_m}.$$

Let $z \in H^{m+2}(\Omega) \cup H^1_{(o}(\Omega)$ be the solution of the VP (1.4) with data v_o. Then by Theorem 2.1. we have

$$|z|_{m+2} \leq C_s(m+1)k^m \|v_o\|_m,$$

where C_s grows at most linearly with m. On $H^m_{(o}$, the full norm $\|\cdot\|_m$ is equivalent to the seminorm $|\cdot|_m$, hence there exists a constant C_1 s.t.

$$|z|_{m+2} \leq C_1 C_s(m+1)k^m |v_o|_m. \tag{3.100}$$

Further, from $B(z-\chi, e) = (v_o, e)$ for all $\chi \in V_h$ we conclude

$$\begin{aligned} \|e\|_{-m}|v_o|_m &\leq C_o(k) \inf_{\chi \in V_h} |z-\chi|_1 |e|_1 \\ &\leq Ck^2 \left(\frac{h}{2p}\right)^{m+1} |z|_{m+2}|e|_1, \end{aligned}$$

where the continuity property of the form B and the approximation property of V_h have been used. The statement now follows, inserting (3.100). ∎

THEOREM 3.8 (*Dual error estimate*). *Let $u \in V^l$ and $u_{fe} \in V_h$ be the exact and the finite element solution to the VP* (1.4). *Assume $1 \leq m \leq p-1$ and $1 \leq l \leq p$.*

Then, if $hk \leq \alpha < \pi$,

$$\|e\|_{-m} \leq C(m,l) \left[C_1 + C_2 K^2 \frac{h}{2p}\right] \left(\frac{kh}{2p}\right)^{m+l+1} \|f\|_{p-1} \tag{3.101}$$

holds, where $C(m,l) = C_a(l)C_s(l)C_s(m+1)$, whereas C_1, C_2 do not depend on h, k, and p.

Proof. Combining Lemma 3.7 with the preasymptotic estimate I, as given in Theorem 3.6, we obtain an estimate with respect to $|u|_{l+1}$. The statement is then concluded from Theorem 2.1 ∎

Remark 18. All propositions on the finite element solution contain the assumption $hk < \pi$. Essentially, we ensure herewith that the inversion of the local stiffness matrices is well defined (hk has to be smaller than the minimal eigenvalue of the condensation). However, the inversion is also well defined if hk lies between the first and the second eigenvalue and so forth (to be more precise, is bounded away both from the first and second exact and the corresponding numerical eigenvalues). It is expected that, under this condition, *more than one halfwave* can be resolved by *one* element of higher approximation in practical computations.

4. Numerical evaluation. In this final section we present some results of numerical experiments that illustrate the theoretical results obtained in the previous section. We solve the model problem for the constant data $f \equiv 1$. The exact solution

$$u(x) = \frac{1}{k^2}\left((1 - \cos kx - \sin k \sin kx) + i\,(1 - \cos k) \sin kx\right)$$

is regular and oscillating.

Best approximation in S_h^p. In Theorem 3.1 we constructed a function $s \in S_h^p(\Omega)$ that has the best H^1 approximation property to a given function $u \in H_{l+1}(\Omega)$. Writing s locally, after scaling to $I = (-1, 1)$, as

$$s(\xi) = \sum_{i=1}^{p+1} c_i^{(j)} N_i(\xi)$$

we have, by interpolating property, $c_1^{(j)} = u_{j-1}$ and $c_2^{(j)} = u_j$. Further, by orthogonality, $c_i^{(j)} = (u', N_i')_{\Delta_j}$ and $|e_{ba}|_1^2 = |u - s|_1^2 = |u|_1^2 - |s|_1^2$ with

$$|s|_1^2 = \sum_{j=1}^{n}\left(h|D^j u|^2 + \frac{2}{h}\sum_{i=3}^{p_j+1} |c_i^{(j)}|^2\right).$$

The approximation property, as given in Theorem 3.1, reads for $m = 1$ and $l = p$ as

$$|\tilde{e}|_1 := \frac{|u - s|_1}{|u|_1} \le C_a(p)\left(\frac{h}{2p}\right)^p \frac{|u|_{p+1}}{|u|_1}, \tag{4.1}$$

where C_a is the approximation constant, growing as $(e/2)^p/(\pi p)^{(1/4)}$.

In Fig. 1 the error $|\tilde{e}|_1$ is plotted for $k = 50$ and $hk < \pi$. The slopes are $-p$, in accordance with (4.1).

Consider next a plot for the approximation error on coarse grid ($hk > \pi$—Fig. 2).

We observe an interesting effect whenever the mesh locally "hits" the size of a halfwave (two, three, …, halfwaves). To explain this effect, note that for $k = n\pi$ (m integer) the solution u reduces to

$$u = \frac{1}{k^2}(1 - \cos kx).$$

Hence in this case we locally solve the problem of best H^1 approximation of $u_{loc} = \cos x$ on one, two, three, …, halfwave(s). For even n, this approximation problem is symmetric and has identical solutions for even/odd p (lines I, III in Fig. 2). For odd n, the problem is antimetric and has identical solutions for odd/even p (line II in Fig. 2). We have chosen $k = 2m\pi$ to highlight the character of the observation. The same effect occurs, however, for any k if either k or kh are close to integer multitudes of π.

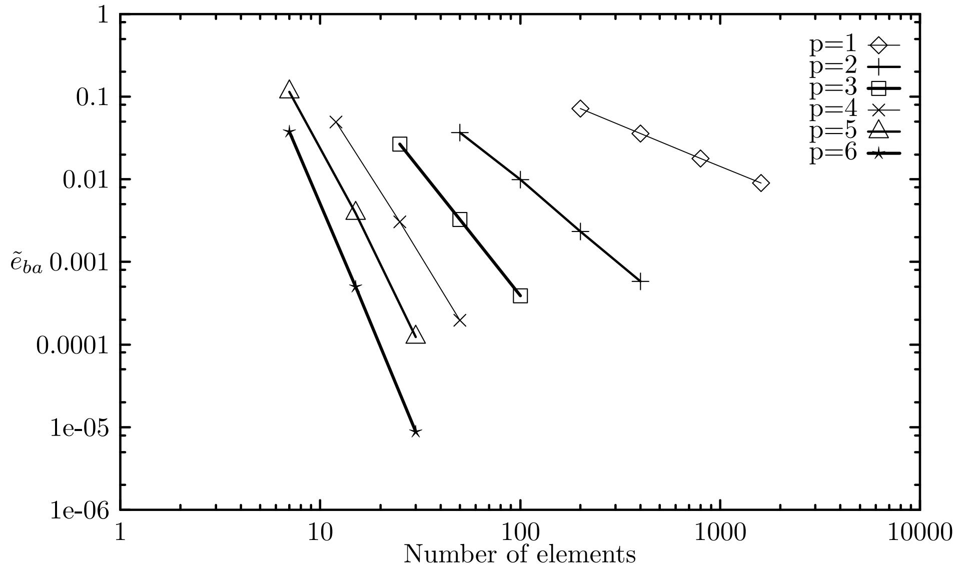

FIG. 1. *Relative error of the best H^1 approximation in $S_h^p(\Omega)$ to the exact solution. Rates of convergence in the H^1-seminorm for $p = 1, 2, \ldots, 6$, $k = 50$.*

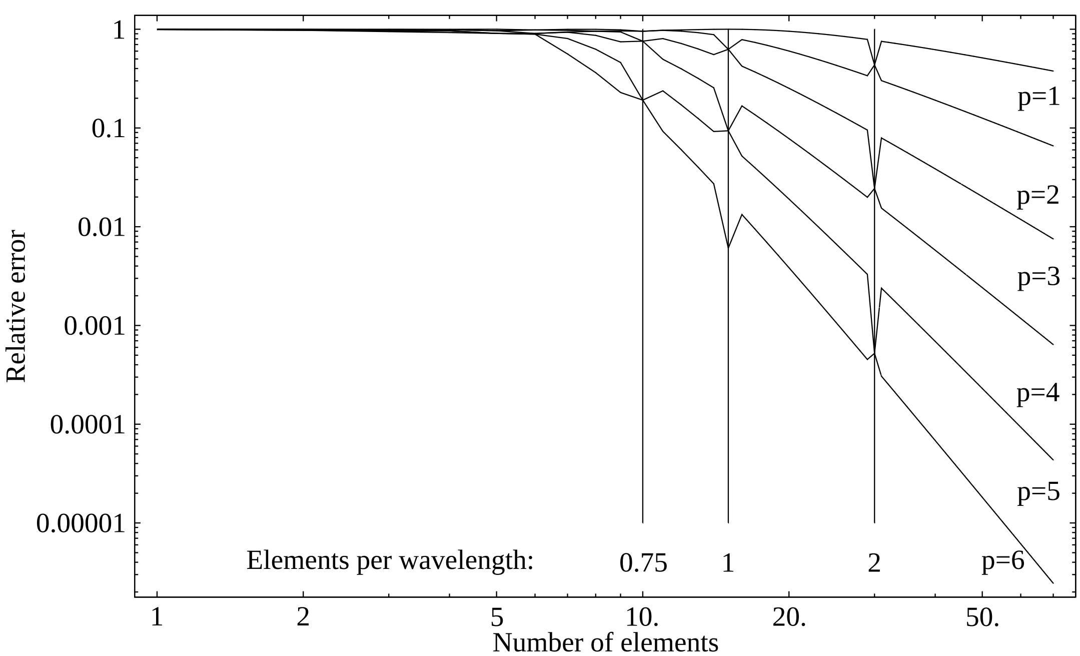

FIG. 2. *Relative error of the best approximation in the H^1-seminorm for $k = 30\pi$.*

Returning to the estimate in Theorem 3.1—here (4.1)—we next present numerical data illuminating the dependence of the constant C_a on p. Namely, for $p = 1, \ldots, 6$ and $hk = 1$ (case $n = k = 50$) we record in Table 1 the relative error, the measured value of $C_{a,meas}(p) = |\tilde{e}|_{1,meas}(2p)^{2p}$, and the predicted $C_a(p)$ from (3.5).

TABLE 1
Constant $C_a(p)$ in the approximation theorem: magnitude as computed from measured data compared to magnitude as computed from theoretical prediction in (3.5).

p	$\|\tilde{e}\|_{1,meas}$	$C_{a,meas}$	C_a,(3.5)
1	0.2823	0.5646	1.02
2	0.367E–1	0.5872	1.17
3	0.3095E–2	0.6685	1.43
4	0.1965E–3	0.9049	1.81
5	0.9829E–5	0.9829	2.33
6	0.4135E–6	1.2357	3.02

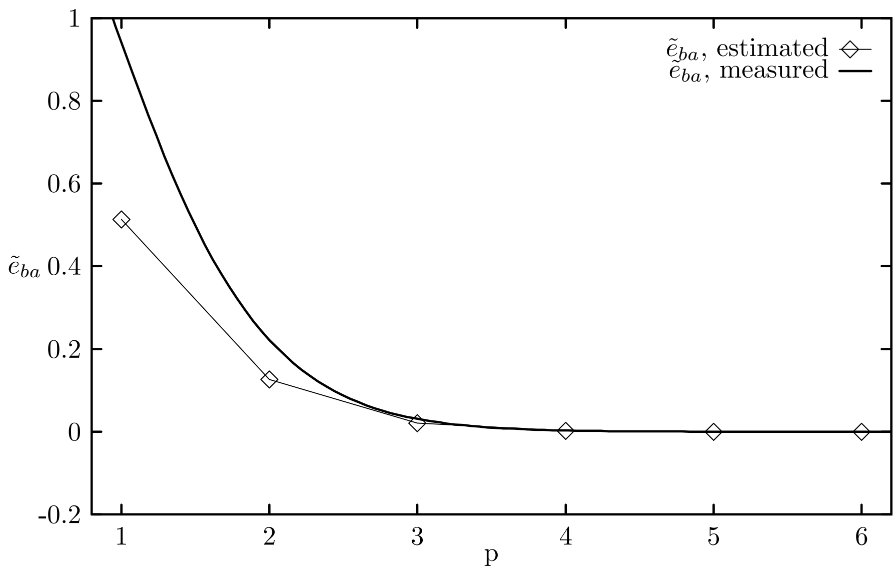

FIG. 3. *Relative error of the best approximation in the H^1-seminorm: estimated vs. measured values for $k = 30\pi$, $n = 50$, and $p = 1, 2, \ldots, 6$.*

We see that, for the particular case computed, both the magnitudes of the measured constant and its growth rates with p are lower than the upper theoretical estimates. We conclude that for the example under consideration, no further growth with p occurs in the relative error due to the ratio $|u|_{p+1}/|u|_1$. For graphic illustration we compare in Fig. 3 (for $k = 30\pi$ and $m = 50$) the estimated and measured errors for $p = 1, \ldots, 6$. Namely, we plot (setting $C_a(p) \equiv 1$)

$$\mathrm{est}_n(p) = (n \cdot 2p)^{-p}$$

for $n = 50$ and $1 \le p \le 6$. We compare with the relative error $|\tilde{e}|_1$ as obtained from computation. The measured error is in close agreement with the estimator, indicating that the estimate (4.1) is sharp.

Finally, we relate the error of best approximation to the number of degrees of freedom of the discrete model. In one dimension, this number is $d := h^{-1}p$. For fixed k, the error estimate (4.1) becomes

$$|\tilde{e}|_1 \le \frac{C(p)}{2^p} d^{-p}. \tag{4.2}$$

The predicted rate of convergence with respect to d is shown in Fig. 4.

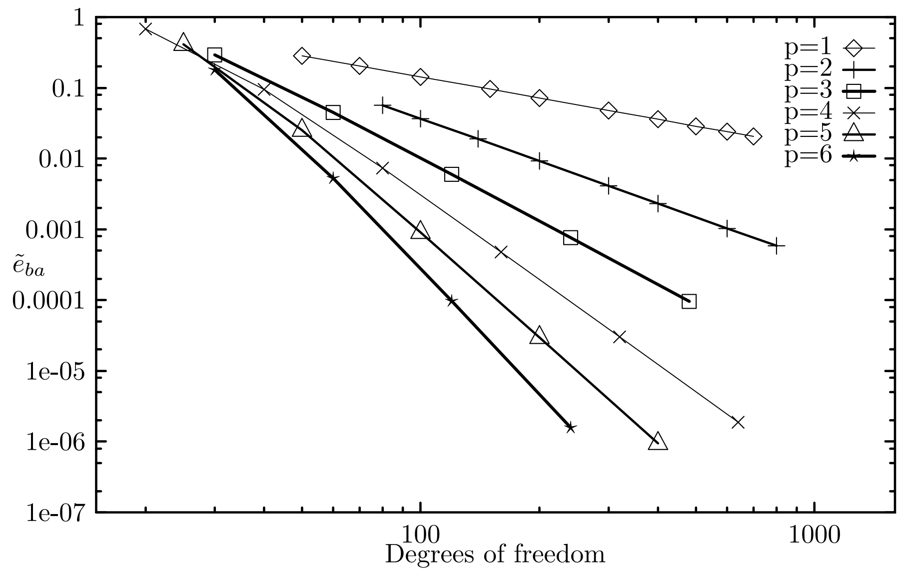

FIG. 4. *Relative error of best approximation vs. number of degrees of freedom for* $k = 50$ *and* $p = 1, 2, \ldots, 6$.

Error of the finite element solution. The error of the finite element solution is computed by

$$|e_{fe}|_1^2 = |u - u_{fe}|_1^2$$
$$= e_1^2 - \sum_{j=1}^{n} \frac{2}{\Delta_j} \sum_{i=3}^{p_j+1} \left(\bar{a}_i^{(j)} c_i^{(j)} + a_i^{(j)} \bar{c}_i^{(j)} - \bar{a}_i^{(j)} a_i^{(j)} \right), \tag{4.3}$$

where

$$e_1^2 = |u|_1^2 - h \sum_{j=1}^{n} (D_j u D_j \bar{u_h} + D_j u_h D_j \bar{u} - D_j u_h D_j \bar{u_h}) \tag{4.4}$$

is the error of piecewise linear approximation (cf. part I) and $a_i^{(j)}$ are the coefficients of the bubble modes in the local finite element ansatz; cf. subsection 3.2.

In Fig. 5, the relative error of the finite element solution is plotted against the relative error of the best approximation. The wave number is $k = 30\pi$ and the results are compared for $p = 1, \ldots, 6$. We clearly see the optimal convergence of the finite element solution for sufficiently small h. In the given example, the optimality constant C in Theorem 3.5 is asymptotically 1 as the figure shows.

To illustrate the behavior in the preasymptotic range, consider the horizontal line drawn at $\tilde{e} \equiv 0.1$. Compared with the asymptotic behavior with the optimality constant close to 1, the finite element solution is significantly polluted on the preasymptotic error level. We give the numerical results in Table 2.

The error estimate for oscillating solutions has been given in Corollary 3.2 (3.98). For our solution, the special case 1 applies, hence we have the estimate (we neglect the term $C_1(p)\theta^{1+2/p}$ of (3.98))

$$|\tilde{e}|_1 \leq C_1(p)\theta^p + C_2(p)k\theta^{2p}$$

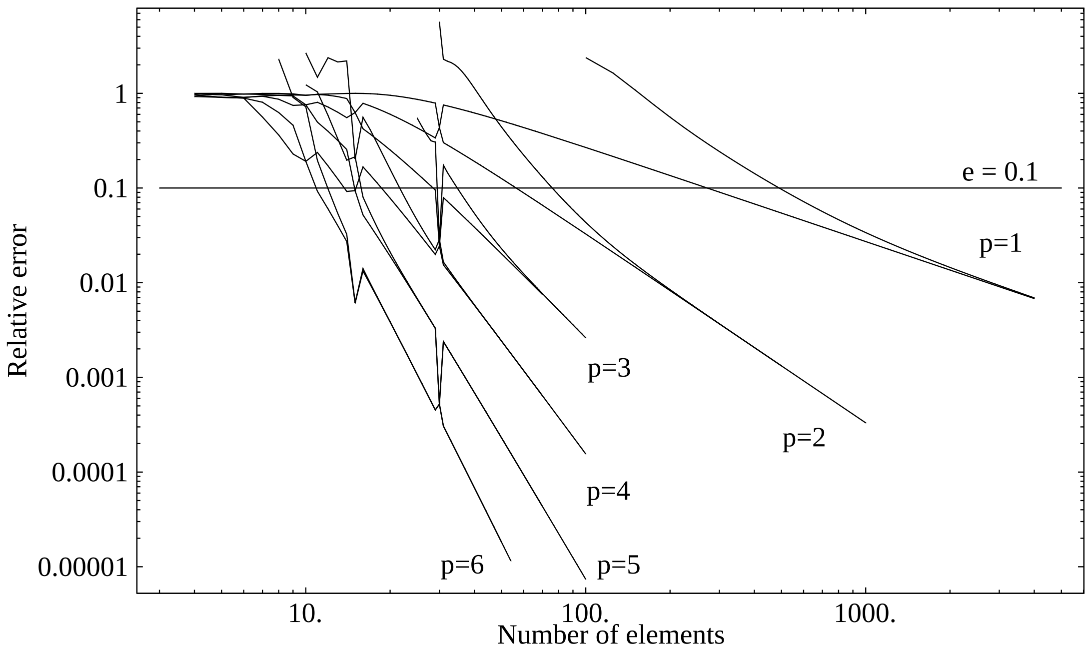

FIG. 5. *Relative error of the finite element solution versus best approximation error for* $k = 30\pi$ *and* $p = 1, 2, \ldots, 6$.

TABLE 2
Errors of finite element solution and best approximation at $\tilde{e}_{fe} \approx 0.1$ *for* $p = 1, \ldots, 6$, $k = 30\pi$.

p	1	2	3	4	5	6
$\tilde{e}_{fe}$ (meas.)	0.099947	0.09956	0.09683	0.0911	0.0829	0.09833
n (# *of elem.*)	491	76	35	22	16	12
# *of DOF*	491	152	105	88	80	72
$\theta = \left(\frac{k}{2np}\right)^p$	0.09597	0.09612	0.0904	0.0822	0.07092	0.07861
$\tilde{e}_{ba}$ (meas.)	0.05538	0.05607	0.05640	0.05543	0.0520	0.0626
$\tilde{e}_{fe} - \tilde{e}_{ba}$	0.044567	0.04349	0.04040	0.03567	0.02380	0.0380
$k\theta^2$	0.8681	0.8707	0.7702	0.6373	0.474	0.5823
$C(p) = \frac{\tilde{e}_{fe} - \tilde{e}_{ba}}{k\theta^2}$	0.0513	0.050	0.0525	0.056	0.0502	0.0656

with $\theta = kh/2p$. In this inequality, $C_1(p)\theta^p$ estimates the error of best approximation as discussed in the previous paragraph. The second member $C_2(p)k\theta^{2p}$ reflects the pollution and is of the same order as the phase lag (cf. Theorem 3.2). Theoretically, the constant C_2 may significantly grow with p; cf. Theorem 3.7 vs. Lemma 3.5. As the table shows, we do not observe this growth in the example considered.

TABLE 3

Number of elements to achieve accuracy (relative error in H^1-seminorm) of ε. Parameters: $k = 30\pi$, $p = 1, ..., 6$; n: number of elements; $DOF = n*p$: degrees of freedom; nmd: computational cost measured in numbers of multiplications and divisions.

p	1	2	3	4	5	6	ε
$\tilde{e}_{fe}$	0.49795	0.51962	0.5470	0.5582	0.5851	0.7208	
n	211	48	25	16	12	10	
DOF	211	96	75	64	60	60	0.5
nmd	1051	284	321	508	824	1296	
$\tilde{e}_{fe}$	0.099947	0.09956	0.09683	0.0911	0.0829	0.09833	
n	491	76	35	22	16	12	
DOF	491	152	105	88	80	72	0.1
nmd	2451	452	451	700	1100	1556	
$\tilde{e}_{fe}$	0.01000	0.0.01055	0.01013	0.01000	0.01026	0.00983	
n	2813	180	64	35	23	17	
DOF	2813	360	192	140	115	102	0.01
nmd	14061	1076	828	1116	1583	2206	

In the fourth row of the table, the number of degrees of freedom is displayed. We observe that the number of degrees of freedom for which the finite element error is a fixed magnitude (given tolerance) decreases significantly with increasing p.

A comparison of numerical effort is made by the following count of the number of multiplications and divisions (nmd). In the given one-dimensional case, condensation involves computing the inverse of a $(p-1) \times (p-1)$ matrix which requires $(p-1)^3$ operations [1, p. 515]. Generally, i.e., on nonuniform mesh, this has to be done on each element. Solution of the resulting tridiagonal system then requires $5n - 4$ operations [1, p. 528]. The overall number of multiplications and divisions is thus

$$nmd = 5n - 4 + n(p-1)^3.$$

In Table 3, we tabulate the numbers nmd needed to achieve a relative error of the finite element solution in H^1-seminorm of 0.1%, 0.5%, or 0.01%, respectively. We observe a significant payoff in computational effort if passing from $p = 1$ to $p = 2$ or $p = 3$. As usual in the h-p method, the optimal relation between h and p depends on the required accuracy; generally, the higher the accuracy the bigger the payoff by higher-order elements.

For the one-dimensional model problem considered here, the optimal choice for practical accuracy seems to be $p = 3$.

In the two-dimensional case, many more degrees of freedoms are needed to compute a finite element solution with 5% accuracy; cf. numerical results in [8]. The pollution effect significantly influences the computational effort. Also, the band after condensation is larger and the local reduction of degrees of freedom is smaller compared to the one-dimensional problem. In two or three dimensions, it is hence possible to expect that larger p are optimal. Iterative methods come into consideration for solving the resulting large systems of linear equations.

The structure of the global stiffness matrix for the Helmholtz equation is similar to the structure obtained for the Laplace equation. For the two-dimensional Laplace equation, implementation and performance of the h-p version of the FEM were studied in [5, 6]. The optimal relation between the number of elements and the degree of approximation is investigated in [6]. The main difference with the present problem is the numerical pollution in the error of discrete solutions of Helmholtz problems. While this aspect needs special consideration in a detailed analysis of the computational effort in the two-dimensional case, we generally expect similarity with the conclusions from the previous work.

REFERENCES

[1] K. E. Atkinson, *An Introduction to Numerical Analysis*, 2nd ed., John Wiley, New York, 1989.

[2] A. K. Aziz, R. B. Kellogg, and A. B. Stephens, *A two point boundary value problem with a rapidly oscillating solution*, Numer. Math. 53 (1988), pp. 107–121.

[3] A. Bayliss, C. I. Goldstein, and E. Turkel, *On accuracy conditions for the numerical computation of waves*, J. Comp. Phys. 59 (1985), pp. 396–404.

[4] I. Babuška and A. K. Aziz, *The mathematical foundations of the finite element method*, in The Mathematical Foundations of the Finite Element Method with Applications to Partial Differential Equations, A. K. Aziz, ed., Academic Press, New York, 1972, pp. 5–359.

[5] I. Babuška, H. C. Elman, and K. Markley, *Parallel implementation of the hp-version of the finite element method on a shared-memory architecture*, SIAM J. Sci. Statist. Comput., 13 (1992), pp. 1433–1459.

[6] I. Babuška and H. C. Elman, *Performance of the h-p-version of the finite element method with various elements*, Internat. J. Numer. Methods Engrg., 36 (1993), pp. 2503–2523.

[7] I. Babuška, I. N. Katz, and B. S. Szabó, *Finite element analysis in one dimension*, Lecture Notes, (1987).

[8] I. Babuška, F. Ihlenburg, E. Paik, and S. Sauter, *A generalized finite element method for solving the Helmholtz equation in two dimensions with minimal pollution*, Comput. Methods Appl. Mech. Engng., 128 (1995), pp. 325–359.

[9] I. Babuška and S. Sauter, *Is the pollution effect of the FEM avoidable for the Helmholtz equation considering high wave numbers*, SIAM J. Numer. Anal., to appear.

[10] I. Babuška, F. Ihlenburg, Th. Strouboulis, and S. Gangaraj, *A Posteriori Error Estimation for Finite Element Solutions of Helmholtz' Equation With High Wave Number - Part* I: *The Quality of Local Indicators; Part* II: *Aposteriori Estimation of the Pollution Effect*, Technical reports, TICAM, The University of Texas at Austin, Austin, TX, May 1996.

[11] R. Bellman, *A note on an inequality of E. Schmidt*, Bull. Amer. Math. Soc., Vol. 50, No. 10 (October 1944), pp. 734–736.

[12] D. Burnett, *A* 3*-D acoustic infinite element based on a generalized multipole expansion*, J. Acoust. Soc. Amer., 96 (1994), pp. 2798–2816.

[13] L. Demkowicz, *Asymptotic convergence in finite and boundary element methods*: *Part* I: *Theoretical results*, Comput. Math. Appl., 27 (1994), pp. 69–84.

[14] J. Douglas Jr., J. E. Santos, D. Sheen, and L. Schreiyer, *Frequency domain treatment of one-dimensional scalar waves*, Math. Models Methods in Appl. Sci., 3 (1993), pp. 171–194.

[15] I. Harari and T. J. R. Hughes, *Finite element method for the Helmholtz equation in an exterior domain: Model problems*, Comput. Methods Appl. Mech. Engrg., 87 (1991), pp. 59–96.

[16] F. Ihlenburg and I. Babuška, *Finite element solution to the Helmholtz equation with high wavenumber - part* I: *The h-version of the FEM*, Comput. Math. Appl., 30 (1995), pp. 9–37.

[17] F. Ihlenburg and I. Babuška, *Dispersion analysis and error estimation of Galerkin finite element methods for the Helmholtz equation*, Internat. J. Numer. Methods Engrg., 38 (1995), pp. 3745–3774.

[18] I. P. Natanson, *Konstruktive Funktionentheorie*, Akademieverlag, Berlin, 1955.

[19] A. H. Schatz, *An analysis of the finite element method for second order elliptic boundary value problems*, in Mathematical Theory of Finite and Boundary Element Methods, A. H. Schatz, V. T. Thomée, and W. L. Wendland, eds., Birkhäuser–Verlag, Basel, Switzerland, 1990.

[20] B. Szabó and I. Babuška, *Finite Element Analysis*, John Wiley, New York, 1991.

[21] L. L. Thompson and P. M. Pinsky, *Complex wavenumber Fourier analysis of the p-version finite element method*, Comput. Mech., 13 (1994), pp. 255–275.

[22] L. L. Thompson and P. M. Pinsky, *A Galerkin least squares finite element method for the two-dimensional Helmholtz equation*, Internat. J. Numer. Methods Engrg., 38 (1995), pp. 371–397.