BENCHMARK COMPUTATION OF EIGENVALUES WITH LARGE DEFECT FOR NON-SELF-ADJOINT ELLIPTIC DIFFERENTIAL OPERATORS[*]

REBEKKA GASSER[†], JOSCHA GEDICKE[‡], AND STEFAN SAUTER[†]

Abstract. In this paper we present benchmark problems for non-self-adjoint elliptic eigenvalue problems with large defect and ascent. We describe the derivation of the benchmark problem with a discontinuous coefficient and mixed boundary conditions. Numerical experiments are performed to investigate the convergence of a Galerkin finite element method with respect to the discretization parameters, the regularity of the problem, and the ascent of the eigenvalue. This allows us to verify the sharpness of the theoretical estimates from the literature with respect to these parameters. We provide numerical evidence about the size of the ascent and show that it is important to consider the mean value for the eigenvalue approximation.

1. Introduction. The spectral theory and spectral analysis for elliptic operators have numerous important practical applications in science and engineering, and there are also many mathematical applications. If the operator is non-self-adjoint and/or has complex-valued coefficients in the operator and/or boundary conditions, the arising sesquilinear form in the variational formulation is not hermitian. Such problems arise frequently, e.g., in electromagnetic scattering in lossy media, or if impedance/Sommerfeld-type boundary conditions are imposed (see, e.g., [20]). Also for the problem of modeling *mechanical* vibrations, non-self-adjoint eigenproblems arise in many applications—here, defective eigenvalues can be interpreted physically as the transition point between an oscillatory and a monotonically decaying behavior (see, e.g., [8]). As a consequence the algebraic multiplicity of an eigenvalue can differ from the geometric multiplicity, and this has strong consequences for their numerical approximation. Classical textbooks on this topic include [5], [7], [9], [21] in the mathematical and [11], [18], [25] in the engineering literature. While the numerical a priori/a posteriori analysis and the numerical simulation of eigenvalue problems for *self-adjoint* problems are fairly mature and numerous monographs and textbook chapters exist in the mathematical and engineering literature [1], [4], [5], [10], [17], [19], [22], [23], [26], [27], [28] the numerical computation of non-self-adjoint eigenvalue problems is less developed. The standard reference for the numerical analysis of the Galerkin finite element discretization is the seminal book chapter by Babuška and Osborn [1]; see also [3], [4], [5], [14], [15], [17], [19], [29]. They derive estimates for

[†]Institut für Mathematik, Universität Zürich, CH-8057 Zürich, Switzerland (rebekka.klara@hispeed.ch, stas@math.uzh.ch).

[‡]University of Vienna, Faculty of Mathematics, Oskar-Morgenstern-Platz 1, 1090 Vienna, Austria. Current address: Institut für Numerische Simulation, Universität Bonn, Endenicher Allee 19b, 53115 Bonn, Germany (gedicke@ins.uni-bonn.de).

the convergence rates depending on the mesh size, the polynomial order of the finite element space, the regularity of the elliptic operator, and also on the *ascent* of *defective* eigenvalues.

To the best of our knowledge, systematic numerical experiments on the sharpness of these estimates with respect to all parameters do not exist in the literature, and it is the goal of our paper to derive benchmark problems for elliptic eigenvalue problems with possible large defects and ascents and to verify by numerical experiments the sharpness of the estimates in [1].

The construction of elliptic eigenvalue problems with large defect is far from being trivial and very sensitive with respect to the choice of parameters in the elliptic operator and boundary condition. We have generalized the one-dimensional Green's function approach in [12] and [24] in order to construct eigenvalue problems with large ascent and defect also in higher dimension. The numerical experiments show very nicely that the estimates in [1] are sharp with respect to all parameters.

The paper is structured as follows.

In section 2 we present the elliptic eigenvalue problem with appropriate coefficients and transform it to an equivalent eigenvalue problem for a compact operator. The Galerkin finite element discretization is introduced in section 3, and we recall briefly the estimates for the convergence rates of the eigenvalues and eigenfunctions from [1]. Section 4 is devoted to the construction of elliptic eigenvalue problems with large defect and ascent. We generalize the one-dimensional Green's function approach from [12] and [24] to higher dimensions and to eigenvalues with larger defect and ascent. In section 5 we present the results of numerical experiments and compare them to the theoretical predictions. These examples show the sharpness of the estimates in [1].

2. Elliptic eigenvalue problems. The computation of eigensystems of partial differential operators is of utmost practical and mathematical importance, and their efficient numerical computation is one major field in numerical analysis and scientific computing. Compared to self-adjoint eigenvalue problems for positive definite operators, numerical methods for the solution of non-self-adjoint eigenvalue problems are less developed, in particular, for problems with *defective* eigenvalues, i.e., eigenvalues where the algebraic and geometric multiplicity of an eigenvalue are different.

As our model eigenvalue problem we consider the elliptic problem

$$- \operatorname{div}(a\nabla u) = \lambda u \quad \text{in } \Omega,$$
$$u = 0 \quad \text{on } \Gamma_{\mathrm{D}},$$
$$a\nabla u \cdot \mathbf{n} + cu = 0 \quad \text{on } \Gamma_{\mathrm{R}}$$

such that the arising variational formulation, in general, is *non-self-adjoint*. Here $\Omega \subset \mathbb{R}^d$ is a bounded Lipschitz domain whose boundary Γ is split into two disjoint measurable subsets: the Dirichlet part Γ_{D} and the Robin part Γ_{R}. We always assume that Γ_{R} has positive surface measure. The unit normal vector field $\mathbf{n} : \Gamma \to \mathbb{S}_{d-1}$ is defined almost everywhere and oriented towards the exterior of Ω. Let $L^2(\Omega)$ denote the usual Lebesgue space with (complex) scalar product $(u,v) = \int_\Omega u\overline{v}$ and norm $\|\cdot\| = (\cdot, \cdot)^{1/2}$. Let $H^1(\Omega)$ denote the standard Sobolev space. We set $V := H^1(\Omega)$ if $\Gamma_{\mathrm{D}} = \emptyset$ and $V := \left\{ u \in H^1(\Omega) \mid \gamma_{\mathrm{D}} u = 0 \right\}$ in the case that Γ_{D} has positive boundary measure. The standard trace operators are denoted by $\gamma_{\mathrm{D}} : H^1(\Omega) \to \Gamma_{\mathrm{D}}$ and $\gamma_{\mathrm{R}} : H^1(\Omega) \to \Gamma_{\mathrm{R}}$. If the $(d-1)$-dimensional surface measure $|\Gamma_{\mathrm{R}}|$ is positive, the multiplicative trace inequality holds:

$$(2.1) \qquad \|\gamma_{\mathrm{R}} u\|_{L^2(\Gamma_{\mathrm{R}})}^2 \leq C_{\mathrm{trace}} \|u\| \, \|u\|_{H^1(\Omega)}.$$

(For $d = 2, 3$, this is the last formula in [16, page 41]. For $d = 1$ it can be obtained by applying the fundamental theorem of calculus to the functions $Z |u|^2$ for a suitable chosen affine function Z).

The variational formulation of the eigenvalue problem is given as follows: Find $(u, \lambda) \in V \setminus \{0\} \times \mathbb{C}$ such that

$$A(u, v) = \lambda (u, v) \quad \forall v \in V \tag{2.2}$$

with

$$A(u, v) := (a \nabla u, \nabla v) + (cu, v)_{L^2(\Gamma_\mathrm{R})}.$$

We assume

$$a \in L^\infty(\Omega) \quad \text{and} \quad \operatorname*{ess\,inf}_{x \in \Omega} (\operatorname{Re} a(x)) \geq \alpha_0 \quad \text{for some } \alpha_0 > 0 \tag{2.3}$$

and that a is sufficiently smooth in an Ω-neighborhood of Γ_R such that the trace $\gamma_\mathrm{R}(a)$ is well defined. Finally we assume that

$$c \in L^\infty(\Gamma_\mathrm{R}) \quad \text{and set} \quad c_0 := \operatorname*{ess\,inf}_{x \in \Gamma_\mathrm{R}} (\operatorname{Re} c(x)). \tag{2.4}$$

Clearly the sesquilinear form A is non-self-adjoint if $\operatorname{Im} c \neq 0$ on Γ_R or $\operatorname{Im} a \neq 0$ in the L^∞ sense.

LEMMA 2.1. *Suppose* (2.3) *and* (2.4) *are satisfied. Let*

$$\Lambda := \begin{cases} 0 & \text{if } c_0 \geq 0 \wedge |\Gamma_\mathrm{D}| > 0, \\ \alpha_0 & \text{if } c_0 \geq 0 \wedge |\Gamma_\mathrm{D}| = 0, \\ \alpha_0 + \frac{C_\mathrm{trace}|c_0|^2}{\alpha_0} & \text{if } c_0 < 0. \end{cases}$$

Then there exist constants C_cont, $c_\mathrm{coer} > 0$ *such that the modified sesquilinear form* $A_\Lambda(u, v) := A(u, v) + \Lambda(u, v)$ *satisfies*

$$|A_\Lambda(u, v)| \leq C_\mathrm{cont} \|u\|_{H^1(\Omega)} \|v\|_{H^1(\Omega)} \quad \forall u, v \in V,$$

$$\operatorname{Re} A_\Lambda(u, u) \geq c_\mathrm{coer} \|u\|_{H^1(\Omega)}^2 \quad \forall u \in V. \tag{2.5}$$

Proof. **(a) Continuity.** Let $\alpha_1 := \|a\|_{L^\infty(\Omega)}$ and $c_1 := \|c\|_{L^\infty(\Gamma_\mathrm{R})}$. For $u, v \in H^1(\Omega)$ it holds that

$$\begin{aligned} |A_\Lambda(u, v)| &\leq \alpha_1 \|\nabla u\| \|\nabla v\| + \Lambda \|u\| \|v\| + c_1 \|u\|_{L^2(\Gamma_\mathrm{R})} \|v\|_{L^2(\Gamma_\mathrm{R})} \\ &\leq \alpha_1 \|\nabla u\| \|\nabla v\| + \Lambda \|u\| \|v\| + c_1 C_\mathrm{trace} \|u\|_{H^1(\Omega)} \|v\|_{H^1(\Omega)} \\ &\leq \left(\alpha_1 + \Lambda + c_1 C_\mathrm{trace}^2\right) \|u\|_{H^1(\Omega)} \|v\|_{H^1(\Omega)}. \end{aligned}$$

(b) Coercivity. To prove coercivity of A_Λ we begin with

$$\operatorname{Re} A_\Lambda(u, u) \geq \alpha_0 \|\nabla u\|^2 + \Lambda \|u\|^2 + c_0 \|u\|_{L^2(\Gamma_\mathrm{R})}^2.$$

If $|\Gamma_\mathrm{D}| > 0$ and $c_0 \geq 0$, we have $\Lambda = 0$, and the Friedrichs inequality (with constant $c_\mathrm{F} > 0$) implies

$$\operatorname{Re} A_\Lambda(u, u) \geq c_\mathrm{F} \alpha_0 \|u\|_{H^1(\Omega)}^2 \quad \forall u \in V.$$

If $c_0 \geq 0$ and $|\Gamma_D| = 0$, the choice of Λ leads to

$$\operatorname{Re} A_\Lambda(u, u) \geq \alpha_0 \|u\|_{H^1(\Omega)}^2 \quad \forall u \in V.$$

If $c_0 < 0$, we employ the multiplicative trace inequality (2.1) and a Young's inequality for $\varepsilon > 0$

$$
\begin{aligned}
\operatorname{Re} A_\Lambda(u, u) &\geq \alpha_0 \|\nabla u\|^2 + \Lambda \|u\|^2 - |c_0| \|u\|_{L^2(\Gamma_R)}^2 \\
&\geq \alpha_0 \|\nabla u\|^2 + \Lambda \|u\|^2 - C_{\text{trace}} |c_0| \|u\| \|u\|_{H^1(\Omega)} \\
&\geq \alpha_0 \|\nabla u\|^2 + \Lambda \|u\|^2 - C_{\text{trace}} |c_0| \left(\frac{\varepsilon}{2} \|\nabla u\|^2 + \left(\frac{\varepsilon}{2} + \frac{1}{2\varepsilon} \right) \|u\|^2 \right).
\end{aligned}
$$

The choice $\varepsilon = \frac{\alpha_0}{2 C_{\text{trace}} |c_0|}$ leads to

$$\operatorname{Re} A_\Lambda(u, u) \geq \frac{3\alpha_0}{4} \|\nabla u\|^2 + \left(\Lambda - \left(\frac{\alpha_0}{4} + \frac{C_{\text{trace}}^2 |c_0|^2}{\alpha_0} \right) \right) \|u\|^2 \geq \frac{3\alpha_0}{4} \|u\|_{H^1(\Omega)}^2. \qquad \square$$

Lemma 2.1 implies via the Lax–Milgram lemma that for any continuous antilinear functional $f \in V^\times$, where $V^\times$ denotes the dual of V, the problem

$$(2.6) \qquad \text{find } u \in V \text{ such that } A_\Lambda(u, v) = (f, v) \quad \forall v \in V$$

has a unique solution. Throughout the paper we identify the $L^2(\Omega)$ scalar product with its continuous extension to the antilinear pairing on $V^\times \times V$.

We say that the problem has regularity $r > 0$ if for any $0 \leq s \leq r$, there exists a constant C_s such that for any $f \in H^{s-1}(\Omega)$ the solution of (2.6) is in $H^{1+s}(\Omega) \cap V$ and satisfies

$$(2.7) \qquad \|u\|_{H^{1+s}(\Omega)} \leq C_s \|f\|_{H^{s-1}(\Omega)}.$$

We say it has adjoint regularity $r_* > 0$ if for any $0 \leq s \leq r_*$, there exists a constant C_s^* such that for any $g \in H^{s-1}(\Omega)$ the solution of the adjoint problem

$$(2.8) \qquad \text{find } z \in V \text{ such that } A_\Lambda(v, z) = (v, g) \quad \forall v \in V$$

is in $H^{1+s}(\Omega) \cap V$ and satisfies

$$(2.9) \qquad \|z\|_{H^{1+s}(\Omega)} \leq C_s^* \|g\|_{H^{s-1}(\Omega)}.$$

It is well known that

$$\min\{r, r_*\} \geq r_0 > 0$$

for some r_0 depending on the geometry of the domain, the geometry of the discontinuities in the coefficient a, as well as on α_0 and $\|a\|_{L^\infty(\Omega)}$.

From the compact embedding $V \overset{c}{\hookrightarrow} L^2(\Omega)$ and Lemma 2.1 it follows that there exists a compact operator $K_\Lambda : V \to V$ such that

$$A_\Lambda(K_\Lambda u, v) = (u, v) \qquad \forall u, v \in V.$$

From the theory of compact operators we deduce that (2.2) is equivalent to the following eigenvalue problem: Find $(u, \mu) \in V \setminus \{0\} \times \mathbb{C}$ such that

$$(2.10) \qquad K_\Lambda u = \mu u.$$

The eigenfunctions are the same as for the original problem (2.2), and the eigenvalues are related by

$$\frac{1}{\mu} = \lambda + \Lambda.$$

This allows us to apply the spectral theory for compact operators to our problem: From (2.5) we conclude that $\operatorname{Re}\mu > 0$. The smallest integer α such that $\mathcal{N}((K_\Lambda - \mu I)^\alpha) = \mathcal{N}((K_\Lambda - \mu I)^{\alpha+1})$ (where $\mathcal{N}$ denotes the null space) is called *ascent of* $K_\Lambda - \mu I$ and is finite for compact operators. The integer $m_{\mathrm{alg}} = \dim \mathcal{N}((K_\Lambda - \mu I)^\alpha)$ is the *algebraic multiplicity* of μ and is finite. The subspace $\mathcal{N}((K_\Lambda - \mu I)^\alpha)$ is called the space of *generalized eigenfunctions* corresponding to the eigenvalue μ. The *geometric multiplicity* is equal to $m_{\mathrm{geo}} := \dim \mathcal{N}(K_\Lambda - \mu I)$ and is always less than or equal to m_{alg}. If $m_{\mathrm{geo}} < m_{\mathrm{alg}}$ we say that the eigenvalue μ is *defective*.

3. Finite element discretization. Let $\mathcal{T} = \{K_i, 1 \le i \le N\}$ denote a conforming finite element mesh for the domain Ω (see, e.g., [2], [6]) consisting of (closed) simplices K. Let $h_K := \operatorname{diam} K$ and $h := \max\{h_K : K \in \mathcal{T}\}$, and let ρ_K denote the diameter of the largest inscribed ball in K. We assume that the mesh is shape regular, i.e., all constants in the error estimates, in general, depend continuously on the shape-regularity constant

$$c_{\mathrm{sr}} := \max\left\{\frac{h_K}{\rho_K} : K \in \mathcal{T}\right\}$$

and, possibly, increase for large c_{sr}. The finite element space is defined by

$$V_h := \left\{u \in V \mid \forall K \in \mathcal{T} : u|_K \in \mathbb{P}_p\right\},$$

where $\mathbb{P}_p$ denotes the space of d-variate polynomials of total degree p.

The Galerkin finite element method to discretize the eigenvalue problem is given by

(3.1)
$$\text{find } (u_{j,h}, \lambda_{j,h}) \in V_h \backslash \{0\} \times \mathbb{C} \quad \text{such that} \quad A(u_{j,h}, v) = \lambda_{j,h}(u_{j,h}, v) \quad \forall v \in V_h.$$

As in the continuous setting (2.10), this problem can be reformulated as an operator equation. Let $K_{\Lambda,h} : V_h \to V_h$ be given by

$$A_\Lambda(K_{\Lambda,h}u, v) = (u, v) \qquad \forall u, v \in V_h.$$

Then (3.1) is equivalent to the following: Find $(u_{j,h}, \mu_{j,h}) \in V_h \backslash \{0\} \times \mathbb{C}$ such that

$$K_{\Lambda,h}u_{j,h} = \mu_{j,h}u_{j,h}$$

and the relation $1/\mu_{j,h} = \lambda_{j,h} + \Lambda$ holds.

In the seminal work by Babuška and Osborn [1] the theory for the numerical solution of eigenvalue problems for elliptic, possibly non-self-adjoint differential operators has been developed. One important result is that, for a defective eigenvalue, the convergence rate suffers from an ascent which is larger than one. From the a priori error analysis of the finite element method [1, Theorem 8.3], we have that

$$|\lambda_j - \lambda_{j,h}| \le C h^{\min(r+r^*, 2p)/\alpha}$$

for the regularity $\min\{r, r^*\} > 0$ (cf. (2.7), (2.9)) of the original and adjoint sesquilinear form, the polynomial degree $p > 0$, and the ascent α of the eigenvalue λ. Note that

the convergence of single eigenvalues deteriorates for large α. In contrast [1, Theorem 8.2] states that there are eigenvalues $\lambda_{j_s,h}$, $1 \leq s \leq m_{\mathrm{alg}}$, that converge towards λ_j, and the convergence rate of the mean eigenvalue is independent of the defect

$$\left| \lambda_j - \left(\frac{1}{m_{\mathrm{alg}}} \sum_{s=1}^{m_{\mathrm{alg}}} \lambda_{j_s,h}^{-1} \right)^{-1} \right| \leq Ch^{\min(r+r^*,2p)}.$$

Moreover, for the convergence of the corresponding eigenfunctions, we have the following result [1, Theorem 8.4]. Suppose that the discrete generalized eigenfunction $u_{j,h}$ satisfies $(K_{\Lambda,h} - \mu_{j,h}I)^k u_{j,h} = 0$ for some $0 < k \leq \alpha$. Then there exists for any $k \leq \ell \leq \alpha$ a generalized eigenfunction u_j in the continuous eigenspace such that $(K_\Lambda - \mu_j I)^\ell u_j = 0$ and

$$\|u_j - u_{j,h}\|_{H^1(\Omega)} \leq Ch^{\min(r,p)(\ell-k+1)/\alpha}.$$

In particular we have for $\ell = k$ that

$$\|u_j - u_{j,h}\|_{H^1(\Omega)} \leq Ch^{\min(r,p)/\alpha},$$

or for $\ell = \alpha$ that

$$\|u_j - u_{j,h}\|_{H^1(\Omega)} \leq Ch^{\min(r,p)(\alpha-k+1)/\alpha}.$$

Suppose that $k = 1$, i.e., $u_{j,h}$ is an eigenfunction; then for $\ell = \alpha$ we get the expected convergence rate similar to simple eigenvalues:

$$\|u_j - u_{j,h}\|_{H^1(\Omega)} \leq Ch^{\min(r,p)}.$$

Since the rate of convergence of the eigenvalue error is usually related to the rate of convergence of the associated eigenfunction, the question arises if any rate between those two extreme cases can be observed in practice. This motivates the construction of benchmark examples with (at least) one defective eigenvalue.

4. One-dimensional benchmark problems. In this section, we employ the general Green's function approach for the construction of defective eigenvalues. In [12], [24] this approach has been used to set up a one-dimensional boundary value problem with eigenvalues of defect 2. Here we consider a more general one-dimensional boundary value problem to construct eigenvalues with defect 3. Later, in section 5, this will be generalized to higher dimensional problems with even larger eigenvalue defects.

Let $\Omega = (0,1)$ be split into subdomains $\Omega_1 := (0,b)$ and $\Omega_2 := (b,1)$ for $b \in (0,1)$. Let $a := \begin{cases} 1 & \text{in } \Omega_1 \\ a_{\mathrm{R}} & \text{in } \Omega_2 \end{cases}$ for $a_{\mathrm{R}} \in \mathbb{C}$ with $\operatorname{Re} a_{\mathrm{R}} > 0$. We consider the following transmission problem:

$$(4.1) \qquad \begin{aligned} L_a u &= \lambda u && \text{in } \Omega \backslash \{b\}, \\ u(0) &= 0, \quad a_R u'(1) + c u(1) = 0, \\ [u]_b &= [au']_b = 0 \end{aligned}$$

with $L_a u := -(au')'$ and the jump $[\cdot]_b$ across b.

Remark 4.1. We do not discuss the case that $\lambda = 0$ is an eigenvalue. This case can be treated by the following analysis by adding Λu on both sides of the first equation of (4.1) for some $\Lambda > 0$ so that $(L_a + \Lambda) u = \tilde{\lambda} u$ with $\tilde{\lambda} = \lambda + \Lambda > 0$.

Let $V = \left\{ v \in H^1\left(\Omega\right) \mid v\left(0\right) = 0 \right\}$, and let $V^\times$ denote the space of antilinear functionals on V. Define the sesquilinear form $A : V \times V \to \mathbb{C}$ by

$$A\left(u, v\right) := \left(au', v'\right) + cu\left(1\right)\bar{v}\left(1\right) \qquad \forall v \in V.$$

The continuity of A (cf. Lemma 2.1) implies that there exists an operator $\mathcal{A} : V \to V^\times$ such that

$$\left(\mathcal{A}u, v\right) = A\left(u, v\right) \qquad \forall u, v \in V.$$

The weak form of (4.1) is given by the following: Find $\left(u, \lambda\right) \in V \setminus \{0\} \times \mathbb{C}$ such that

$$A\left(u, v\right) = \lambda\left(u, v\right) \qquad \forall v \in V$$

or in operator form

$$\left(\mathcal{A} - \lambda I\right)u = 0.$$

In the following we will derive a representation of the exact solutions for this problem which will allow us to determine choices of parameters a_R, b, c such that an eigenvalue becomes defective. Let $\mu := \begin{cases} \mu_{\mathrm{L}} := \sqrt{\lambda} & \text{in } \Omega_1 \\ \mu_{\mathrm{R}} := \sqrt{\lambda/a_{\mathrm{R}}} & \text{in } \Omega_2 \end{cases}$ and $v_{a,b,c}^{\mathrm{L}}\left(\lambda, x\right) := \sin\mu_{\mathrm{L}}x$. We employ the ansatz

$$(4.2) \qquad u\left(x\right) = \begin{cases} A_1 v_{a,b,c}^{\mathrm{L}}\left(\lambda, x\right) & x \in \Omega_1, \\ A_2 v_{a,b,c}^{\mathrm{R}}\left(\lambda, x\right) & x \in \Omega_2, \end{cases}$$

where the coefficients c_1, c_2 in $v_{a,b,c}^{\mathrm{R}}\left(\lambda, x\right) = c_1 \sin\mu_{\mathrm{R}}x + c_2 \cos\mu_{\mathrm{R}}x$, $x \in \Omega_2$, are chosen such that $a_R \partial_x v_{a,b,c}^{\mathrm{R}}\left(\lambda, 1\right) + c v_{a,b,c}^{\mathrm{R}}\left(\lambda, 1\right) = 0$ is satisfied, i.e.,

$$(4.3) \quad v_{a,b,c}^{\mathrm{R}}\left(\lambda, x\right) = \left(c\cos\mu_{\mathrm{R}} + \mu_{\mathrm{R}}\sin\mu_{\mathrm{R}}\right)\sin\mu_{\mathrm{R}}x + \left(\mu_{\mathrm{R}}\cos\mu_{\mathrm{R}} - c\sin\mu_{\mathrm{R}}\right)\cos\mu_{\mathrm{R}}x.$$

We employ the transmission conditions $[u]_b = [au']_b = 0$ to see that the coefficients A_1, A_2 in (4.2) satisfy the linear relation

$$(4.4)$$
$$\mathbf{M}_{a,b,c}\left(\lambda\right)\begin{pmatrix} A_1 \\ A_2 \end{pmatrix} = \mathbf{0} \quad \text{with} \quad \mathbf{M}_{a,b,c}\left(\lambda\right) := \begin{bmatrix} v_{a,b,c}^{\mathrm{L}}\left(\lambda, b\right) & -v_{a,b,c}^{\mathrm{R}}\left(\lambda, b\right) \\ \partial_x v_{a,b,c}^{\mathrm{L}}\left(\lambda, b\right) & -a\partial_x v_{a,b,c}^{\mathrm{R}}\left(\lambda, b\right) \end{bmatrix}.$$

Hence, $\lambda \in \mathbb{C}$ is an eigenvalue of problem (4.1) if and only if $\det\mathbf{M}_{a,b,c}\left(\lambda\right) = 0$ since then (4.4) has nontrivial solutions.

Remark 4.2. For $\lambda \neq 0$, the matrix $\mathbf{M}_{a,b,c}\left(\lambda\right)$ is not the zero matrix. Hence, the eigenspace of any eigenvalue $\lambda_j \neq 0$ has dimension 1.

LEMMA 4.3. *Let λ_j be an eigenvalue of* (4.1). *Then, there exists a neighborhood $\mathcal{U}\left(\lambda_j\right)$ such that $v_{a,b,c}^{\mathrm{R}}\left(\lambda_j, \cdot\right)$ is not the zero function for all $\lambda \in \mathcal{U}\left(\lambda_j\right)$, i.e., the coefficients $c_1\left(\lambda\right) = c\cos\mu_{\mathrm{R}} + \mu_{\mathrm{R}}\sin\mu_{\mathrm{R}}$ and $c_2\left(\lambda\right) = \mu_{\mathrm{R}}\cos\mu_{\mathrm{R}} - c\sin\mu_{\mathrm{R}}$ in* (4.3) *are not vanishing simultaneously in this neighborhood.*

Proof. If the coefficients a, b, c are such that $v_{a,b,c}^{\mathrm{R}}\left(\lambda_j, \cdot\right)$ is the zero function, then $v_{a,b,c}^{\mathrm{L}}\left(\lambda_j, \cdot\right)$ satisfies

$$-\partial_x^2 v_{a,b,c}^{\mathrm{L}}\left(\lambda_j, \cdot\right) = \lambda_j v_{a,b,c}^{\mathrm{L}}\left(\lambda_j, \cdot\right) \quad \text{in } \Omega_1,$$
$$v_{a,b,c}^{\mathrm{L}}\left(\lambda_j, 0\right) = v_{a,b,c}^{\mathrm{L}}\left(\lambda_j, b\right) = \partial_x v_{a,b,c}^{\mathrm{L}}\left(\lambda_j, b\right) = 0.$$

However, this implies $v_{a,b,c}^{\mathrm{L}}(\lambda_j, \cdot) = 0$, and hence λ_j cannot be an eigenvalue of (4.1). By contradiction we may conclude that $v_{a,b,c}^{\mathrm{R}}(\lambda_j, \cdot)$ is not the zero function. Its definition implies that the coefficients $c_1(\lambda)$, $c_2(\lambda)$ cannot vanish simultaneously at $\lambda = \lambda_j$. Since c_1, c_2 depend continuously on λ this property carries over to a neighborhood $\mathcal{U}(\lambda_j)$ of λ_j. $\qquad\square$

In order to determine the defect and ascent of the eigenvalue and a basis for the generalized eigenspace, we will employ the Green's function for problem (4.4):

$$L_{a,x}G_{a,b,c}(\lambda, x, y) - \lambda G_{a,b,c}(\lambda, x, y) = \delta(x - y), \quad \text{for } (x, y) \in \Omega \backslash \{b\} \times \Omega,$$
$$G_{a,b,c}(\lambda, 0, y) = 0 \quad \text{and} \quad a_R \partial_x G_{a,b,c}(\lambda, 1, y) + c G_{a,b,c}(\lambda, 1, y) = 0, \quad \text{in } \Omega,$$
$$[G_{a,b,c}(\lambda, \cdot, y)]_b = [a \partial_x G_{a,b,c}(\lambda, \cdot, y)]_b = 0, \qquad\qquad \text{in } \Omega.$$

Here, the subscript x in $L_{a,x}$ indicates that the differential operator is applied with respect to the x variable. It is an easy exercise to prove that the Green's function is given by

$$G_{a,b,c}(\lambda, x, y) := G_{a,b,c}^{\mathrm{free}}(\lambda, x, y) + G_{a,b,c}^{\mathrm{hom}}(\lambda, x, y)$$

with

$$G_{a,b,c}^{\mathrm{free}}(\lambda, x, y) := -\frac{\mathrm{e}^{\mathrm{i}\,\mu|x-y|}}{2\,\mathrm{i}\,\mu} + \begin{cases} \frac{\mathrm{e}^{\mathrm{i}\,\mu_{\mathrm{L}}y}}{2\,\mathrm{i}\,\mu_{\mathrm{L}}}\cos\mu_{\mathrm{L}}x & \text{in } \Omega_1, \\ \frac{\mathrm{e}^{-\mathrm{i}\,\mu_{\mathrm{R}}y}}{2\,\mathrm{i}\,\mu_{\mathrm{R}}}\mathrm{e}^{\mathrm{i}\,\mu_{\mathrm{R}}x} & \text{in } \Omega_2 \end{cases}$$

and

$$G_{a,b,c}^{\mathrm{hom}}(\lambda, x, y) := \begin{cases} G_1(y)\,v_{a,b,c}^{\mathrm{L}}(\lambda, x) & \text{in } \Omega_1 \times \Omega, \\ G_2(y)\,v_{a,b,c}^{\mathrm{R}}(\lambda, x) & \text{in } \Omega_2 \times \Omega. \end{cases}$$

Note that the *boundary conditions* are already incorporated into $G_{a,b,c}^{\mathrm{free}}$. The coefficient functions G_1, G_2 are the solution of the system of linear equations

(4.5)

$$\mathbf{M}_{a,b,c}(\lambda)\begin{pmatrix} G_1(y) \\ G_2(y) \end{pmatrix} = \begin{pmatrix} g_{a,b,c}(\lambda, y) \\ f_{a,b,c}(\lambda, y) \end{pmatrix}; \quad \begin{cases} g_{a,b,c}(\lambda, y) := \left[G_{a,b,c}^{\mathrm{free}}(\lambda, \cdot, y)\right]_b, \\ f_{a,b,c}(\lambda, y) := \left[a \partial_x G_{a,b,c}^{\mathrm{free}}(\lambda, \cdot, y)\right]_b. \end{cases}$$

LEMMA 4.4. *Let $\lambda \neq 0$. Then, the functions $g_{a,b,c}(\lambda, \cdot)$ and $f_{a,b,c}(\lambda, \cdot)$ are linearly independent.*

Proof. We have

$$g_{a,b,c}(\lambda, y) = \frac{\mathrm{e}^{\mathrm{i}\,\mu_{\mathrm{R}}(b-y)} - \mathrm{e}^{\mathrm{i}\,\mu_{\mathrm{R}}|b-y|}}{2\,\mathrm{i}\,\mu_{\mathrm{R}}} + \frac{\mathrm{e}^{\mathrm{i}\,\mu_{\mathrm{L}}|b-y|} - \mathrm{e}^{\mathrm{i}\,\mu_{\mathrm{L}}y}\cos\mu_{\mathrm{L}}b}{2\,\mathrm{i}\,\mu_{\mathrm{L}}},$$
$$f_{a,b,c}(\lambda, y) = \frac{\mathrm{sign}(b-y)}{2}\delta(y) + a_{\mathrm{R}}\frac{\mathrm{e}^{\mathrm{i}\,\mu_{\mathrm{R}}(b-y)}}{2} + \frac{\mathrm{e}^{\mathrm{i}\,\mu_{\mathrm{L}}y}}{2\,\mathrm{i}}\sin\mu_{\mathrm{L}}b$$

for $\delta(y) := \mathrm{e}^{\mathrm{i}\,\mu_{\mathrm{L}}|b-y|} - a_{\mathrm{R}}\mathrm{e}^{\mathrm{i}\,\mu_{\mathrm{R}}|b-y|}$.

First case. $a_{\mathrm{R}} = 1$ so that $\mu = \mu_{\mathrm{R}} = \mu_{\mathrm{L}} = \sqrt{\lambda}$. Then

$$g_{a,b,c}(\lambda, y) = \frac{\mathrm{e}^{\mathrm{i}\,\mu(b-y)} - \mathrm{e}^{\mathrm{i}\,\mu y}\cos\mu b}{2\,\mathrm{i}\,\mu},$$
$$f_{a,b,c}(\lambda, y) = \frac{\mathrm{e}^{\mathrm{i}\,\mu(b-y)}}{2} + \frac{\mathrm{e}^{\mathrm{i}\,\mu y}}{2\,\mathrm{i}}\sin\mu b.$$

These functions are linearly independent provided $(\alpha, \beta) = (0, 0)$ is the only solution of

$$(4.6) \qquad (\alpha + \mathrm{i}\,\beta)\,\mathrm{e}^{\mathrm{i}\,\mu(b-y)} + (\beta \sin \mu b - \alpha \cos \mu b)\,\mathrm{e}^{\mathrm{i}\,\mu y} = 0.$$

Since $\mu \neq 0$ the functions $\mathrm{e}^{\mathrm{i}\,\mu(b-y)}$ and $\mathrm{e}^{\mathrm{i}\,\mu y}$ are linearly independent so that (4.6) implies $\alpha = -\mathrm{i}\,\beta$ and $\beta \sin \mu b = \alpha \cos \mu b$. It is a simple exercise to verify that $(\alpha, \beta) = (0,0)$ is the only solution, so we proved the lemma for the first case.

Second case. $a_{\mathrm{R}} \neq 1$ so that $\mu_{\mathrm{R}} \neq \mu_{\mathrm{L}}$. Observe that the function $g_{a,b,c}(\lambda, \cdot)$ is continuous while $f_{a,b,c}(\lambda, \cdot)$ is discontinuous since $\delta(b) = 1 - a_{\mathrm{R}} \neq 0$. Hence, they are linearly independent provided $g_{a,b,c}$ is not the zero function. Let $0 \leq y \leq b$ so that

$$g_{a,b,c}(\lambda, y) = \frac{\mathrm{e}^{\mathrm{i}\,\mu_{\mathrm{L}} b}\,\mathrm{e}^{-\mathrm{i}\,\mu_{\mathrm{L}} y} - \cos \mu_{\mathrm{L}} b\,\mathrm{e}^{\mathrm{i}\,\mu_{\mathrm{L}} y}}{2\,\mathrm{i}\,\mu_{\mathrm{L}}}.$$

Since $\mathrm{e}^{-\mathrm{i}\,\mu_{\mathrm{L}} y}$ and $\mathrm{e}^{\mathrm{i}\,\mu_{\mathrm{L}} y}$ are linearly independent the function $g_{a,b,c}(\lambda, \cdot)$ is the zero function if and only if $\mathrm{e}^{\mathrm{i}\,\mu_{\mathrm{L}} b}$ and $\cos \mu_{\mathrm{L}} b$ are zero. However, this is not possible, and we proved the lemma also for the second case. $\qquad \square$

If λ is not an eigenvalue of (4.1) the system (4.5) has a unique solution, and the Green's function is well defined. If λ approaches an eigenvalue λ_j the Green's function has a singularity which is related only to the part $G_{a,b,c}^{\mathrm{hom}}(\lambda, \cdot, \cdot)$ since $G_{a,b,c}^{\mathrm{free}}$ is a bounded function with respect to λ. The order of singularity (as $\lambda \to \lambda_j$) depends on the order of the zero of $\det \mathbf{M}_{a,b,c}(\lambda)$ at $\lambda = \lambda_j$. Expansion of $\det \mathbf{M}_{a,b,c}(\lambda)$ about some λ_j leads to

$$(4.7) \qquad \det \mathbf{M}_{a,b,c}(\lambda) = \sum_{\ell=0}^{\infty} \gamma_\ell (\lambda - \lambda_j)^\ell \quad \text{for some } \gamma_\ell = \gamma_\ell(a, b, c, \lambda_j).$$

THEOREM 4.5. *Let* $\operatorname{Re} a_{\mathrm{R}} > 0$, *and let* $(u_j, \lambda_j) \in V \setminus \{0\} \times \mathbb{C} \setminus \{0\}$ *denote an eigenpair of* (4.1), *and let the determinant of* $\mathbf{M}_{a,b,c}(\lambda)$ *be expanded according to* (4.7). *Let* ν *denote the largest integer such that* $\gamma_\ell = 0$ *for* $0 \leq \ell \leq \nu$ *and we assume* $\nu < \infty$, *i.e.,* $\det \mathbf{M}_{a,b,c}$ *is not the zero function to avoid pathological cases. Then, the ascent* α_j *of* λ_j *equals* ν. *The dimension of the generalized eigenspace is* $m_{\mathrm{alg}} = \nu$ *and spanned by*

$$w_\ell := \left.\frac{d^{\ell-1} u_j}{d\lambda^{\ell-1}}\right|_{\lambda = \lambda_j} \quad \text{for } 1 \leq \ell \leq \nu.$$

The functions w_ℓ *belong to* $\mathcal{N}((\mathcal{A} - \lambda_j I)^\kappa)$ *for* $\kappa = \ell$ *but not for* $\kappa < \ell$.

Proof. First, we determine the order of singularity at $\lambda = \lambda_j$ of the Green's function. It suffices to study the part $G_{a,b,c}^{\mathrm{hom}}(\lambda, \cdot, \cdot)$ since $G_{a,b,c}^{\mathrm{free}}(\lambda, \cdot, \cdot)$ does not introduce poles. The singularity at λ_j is induced via the coefficients G_1, G_2 as the solution of (4.5). For λ being not an eigenvalue of (4.1) we have[1]

(4.8)

$$G_{a,b,c}^{\mathrm{hom}}(\lambda, x, y) := \begin{cases} \dfrac{\left\langle \mathbf{M}_{a,b,c}^{\times}(\lambda)\begin{pmatrix} g_{a,b,c}(\lambda, y) \\ f_{a,b,c}(\lambda, y) \end{pmatrix}, \begin{pmatrix} v_{a,b,c}^{\mathrm{L}}(\lambda, x) \\ 0 \end{pmatrix} \right\rangle}{\det \mathbf{M}_{a,b,c}(\lambda)} & \text{in } \Omega_1 \times \Omega, \\[2em] \dfrac{\left\langle \mathbf{M}_{a,b,c}^{\times}(\lambda)\begin{pmatrix} g_{a,b,c}(\lambda, y) \\ f_{a,b,c}(\lambda, y) \end{pmatrix}, \begin{pmatrix} 0 \\ v_{a,b,c}^{\mathrm{R}}(\lambda, x) \end{pmatrix} \right\rangle}{\det \mathbf{M}_{a,b,c}(\lambda)} & \text{in } \Omega_2 \times \Omega. \end{cases}$$

---

[1] For a 2×2 matrix $\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ we set $\mathbf{A}^{\times} := \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$. For $\mathbf{u} = (u_i)_{i=1}^2$, $\mathbf{v} = (v_i)_{i=1}^2 \in \mathbb{C}^2$, we denote by $\langle \cdot, \cdot \rangle$ the *bilinear* form $\langle \mathbf{u}, \mathbf{v} \rangle = u_1 v_1 + u_2 v_2$.

Poles are introduced to the Green's function via the zeroes of det $\mathbf{M}_{a,b,c}(\lambda)$ (cf. (4.7)); to determine their orders we also have to investigate whether the numerators in (4.8) can be the zero function for certain values of λ_j. In the following we will prove (by contradiction) that the two brackets $\langle \cdot, \cdot \rangle$ in (4.8) are *not* the zero function in a neighborhood of an eigenvalue. Since $v_{a,b,c}^{\mathrm{L}}(\lambda, \cdot)$ and $v_{a,b,c}^{\mathrm{R}}(\lambda, \cdot)$ are not the zero function in a neighborhood of an eigenvalue λ_j (cf. Lemma 4.3) we conclude that $\mathbf{M}_{a,b,c}^{\times}(\lambda) \left(\begin{smallmatrix} g_{a,b,c}(\lambda,y) \\ f_{a,b,c}(\lambda,y) \end{smallmatrix} \right)$ must be the zero function. Recall that $\lambda \neq 0$. From Lemma 4.4 we know that $g_{a,b,c}(\lambda, \cdot)$ and $f_{a,b,c}(\lambda, \cdot)$ are linearly independent so that there exist two values $y_1, y_2 \in \Omega$ such that the vectors $(g_{a,b,c}(\lambda, y_i), f_{a,b,c}(\lambda, y_i))^{\mathsf{T}}$, $i = 1, 2$, are linearly independent. Hence $\mathbf{M}_{a,b,c}(\lambda)$ must be the zero matrix in order that the two brackets $\langle \cdot, \cdot \rangle$ in (4.8) could be the zero function. However, for $\lambda \neq 0$ the matrix entries $v_{a,b,c}^{\mathrm{L}}(\lambda, b) = \sin \mu_{\mathrm{L}} b$ and $\partial_x v_{a,b,c}^{\mathrm{L}}(\lambda, b) = \mu_{\mathrm{L}} \cos \mu_{\mathrm{L}} b$ cannot be zero simultaneously, and, hence, the matrix $\mathbf{M}_{a,b,c}(\lambda)$ cannot be the zero matrix (cf. Remark 4.2). From (4.7) and $\gamma_\ell = 0$ for $0 \leq \ell \leq \nu$ we now can conclude that the order of the pole of the Green's function at λ_j equals ν.

From [24, Theorem 3.1] we know that the ascent of λ_j equals the order of the pole ν.

Since the geometric multiplicity of eigenvalues for problem (4.1) equals 1 (cf. Remark 4.2) we get by induction that the space of generalized eigenfunctions is spanned by the solutions of the following sequence of problems: Set $H_0 := \{0\}$. For $1 \leq \ell \leq \nu$, let $w_\ell \in V \backslash H_{\ell-1}$ be a solution of

$$(\mathcal{A} - \lambda_j I)^\ell w_\ell = 0 \quad \text{and set} \quad H_\ell := H_{\ell-1} + \operatorname{span}\{w_\ell\}.$$

Clearly, we have $w_1 = u_j$, and H_1 is the (one-dimensional) eigenspace of λ_j. For $\ell = 2, \ldots, \nu$ we obtain by induction

$$(\mathcal{A} - \lambda_j I)^\ell w_{\ell-1} = (\mathcal{A} - \lambda_j I) \left((\mathcal{A} - \lambda_j I)^{\ell-1} w_{\ell-1} \right) = 0.$$

We differentiate this equation with respect to λ_j and obtain

$$-\ell (\mathcal{A} - \lambda_j I)^{\ell-1} w_{\ell-1} + (\mathcal{A} - \lambda_j I)^\ell \partial_{\lambda_j} w_{\ell-1} = 0.$$

The first summand vanishes by induction so that $w_\ell = \partial_{\lambda_j} w_{\ell-1}$ is in H_ℓ if we prove that it is not the zero function. Since u_j is an eigenfunction of (4.1) it is the non-zero function: on Ω_1 it is a multiple of $\sin \sqrt{\lambda_j} x$ and on Ω_2 a linear combination of $\cos \sqrt{\lambda_j / a_{\mathrm{R}}} x$ and $\sin \sqrt{\lambda_j / a_{\mathrm{R}}} x$ so that no derivative with respect to λ is the zero function. Hence $w_\ell \in H_\ell$. $\qquad \square$

Remark 4.6. The functions $\gamma_\ell(a, b, c, \lambda)$ in (4.7) are transcendental complex-valued functions, and it is a nontrivial task to determine coefficients a, b, c, λ such that γ_ℓ is zero for $\ell = 0, \ldots, \nu$ for some $\nu > 0$. In [13] a procedure is described how such parameter configurations can be computed to high precision. In the setting of our paper, we were able to choose these parameters, using the nonlinear solver of Mathematica applied to the symbolic expression of $\gamma_\ell(a, b, c, \lambda)$ and carefully chosen starting values, such that $\gamma_\ell(a, b, c) = 0$ for $\ell = 0, 1, 2, 3$. We conjecture that it is not possible to find configurations for problem (4.1) such that γ_ℓ vanishes at a higher order.

5. Numerical experiments. In this section we present several numerical experiments that indicate that the Babuška–Osborn theory is sharp for defective eigenvalues. The numerical experiments below verify that the eigenvalue errors of a defective

eigenvalue have reduced convergence rates while the mean eigenvalue error converges with the full rate. We construct two main examples based on the construction in section 4; one has full regularity, and one has reduced regularity.

In the following figures, we display the errors in terms of the number of degrees of freedom N, where $h \approx N^{-1/d}$ for uniform meshes.

5.1. Regular example. In this example we align the jump of the diffusion coefficient with the mesh so that the diffusion coefficient is piecewise constant on refined meshes. Although the continuous eigenfunctions are not globally smooth due to the jumping coefficient, they are piecewise smooth, and if the mesh contains the jump point as a mesh point we can expect that the convergence orders are *not* reduced due to lower global regularity. In this light, we call this set of examples *regular examples*.

5.2. One-dimensional example. Let $\Omega = (0,1)$ with $a = 1$ on $[0, 1/2]$ and $a = a_{\mathrm{R}}$ on $(1/2, 1]$; using the construction of the previous section, we compute

$$a_{\mathrm{R}} = 0.1069220800406739 + 0.08937533852238478i,$$
$$c = -0.9634059612381408 + 0.5989684988897067i$$

for the first (smallest in magnitude) complex eigenvalue

$$\lambda = 5.250721274740938 + 6.750931815875402i,$$

which has algebraic multiplicity $m_{alg} = 3$ and ascent $\alpha = 3$ by construction.

In Figure 5.1 we observe convergence rates according to the theory; in case of the $\mathbb{P}_1$ finite element method the convergence is of order $\mathcal{O}(N^{-2/3})$ (due to $\alpha = 3$) for the eigenvalue errors $|\lambda - \lambda_{j,h}|$, $j = 1, 2, 3$, and optimal convergence $\mathcal{O}(N^{-2})$ for the mean eigenvalue error. For the second order $\mathbb{P}_2$ finite element method we observe twice the convergence rate, i.e., $\mathcal{O}(N^{-4/3})$ for the eigenvalue errors and $\mathcal{O}(N^{-4})$ for the mean eigenvalue error, which show that the theoretical predicted rates are sharp for these examples.

Next, we investigate the sensitivity of the defective eigenvalue λ. Since the defect is very sensitive towards the choice of the parameters a_R and c, we perturb only the real part of c by adding a small (real) value δ. In Figure 5.2, we observe that even very small perturbations δ immediately lead to a splitting of the defective eigenvalue
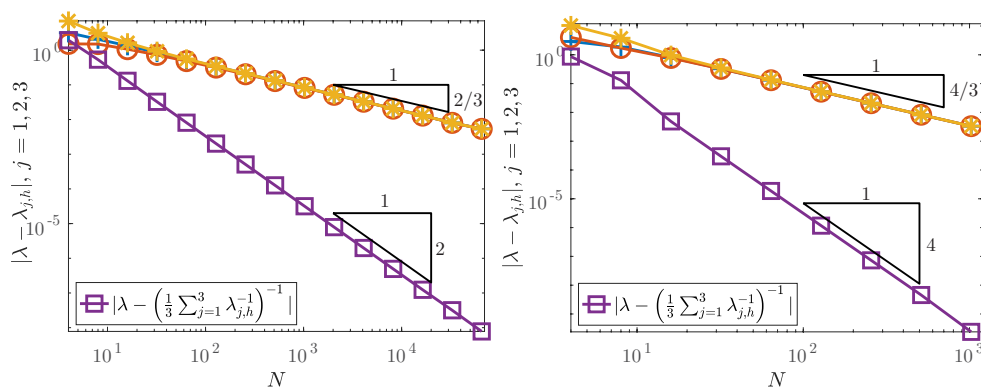

FIG. 5.1. *Regular example with* $\mathbb{P}_1$ *and* $\mathbb{P}_2$ *finite element method* (*FEM*) *in one dimension* (1d).
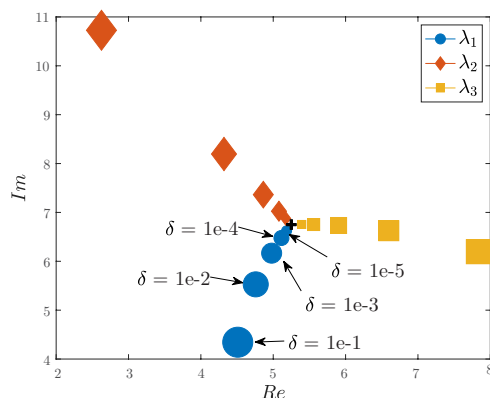
FIG. 5.2. *Sensitivity of eigenvalues due to δ perturbations of the real part of c.*
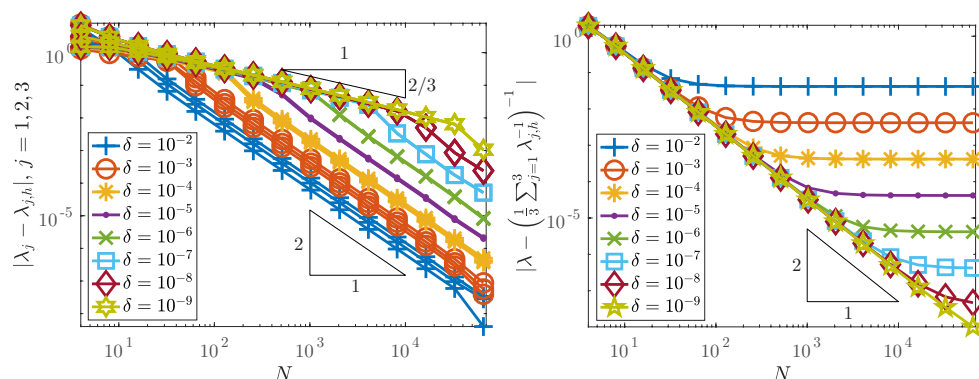

FIG. 5.3. *Convergence history for $\mathbb{P}_1$ FEM in 1d for δ perturbed problems with clustered eigenvalues.*

into three clustered eigenvalues. Even a relatively small perturbation $\delta = 10^{-2}$, of about 1%, already leads to a significant separation of the eigenvalues of size greater than 2.

We investigate the transition of the defective eigenvalue into a separated cluster of eigenvalues in more detail and make the following observations in Figure 5.3. In the left figure we display the eigenvalue errors $|\lambda_j - \lambda_{j,h}|$, $j = 1, 2, 3$, for $\mathbb{P}_1$ finite elements and different perturbations δ towards precomputed reference values λ_j for the clustered eigenvalues. We computed the reference values with higher order $\mathbb{P}_3$ finite elements on fine meshes with high accuracy. In the right figure, we show the convergence of the mean eigenvalue error towards the defective eigenvalue λ. We observe that for $\delta = 10^{-2}$ the eigenvalues are well separated; hence the three distinct eigenvalues converge with optimal rates, and the mean eigenvalue error does not converge towards the defective eigenvalue λ. Interestingly, for smaller values of δ, we observe that there seems to be a resolution barrier. Before a certain resolution is reached, we observe that the eigenvalue errors show the reduced convergence rate of approximating a defective eigenvalue, and even the mean value converges towards the defective eigenvalue λ. Once the mesh is fine enough, so that the clustered eigenvalues can be separated also on the discrete level, the eigenvalue errors converge with optimal rates, and their mean value stops converging towards the defective eigenvalue.

In [13], other explicit choices of parameters are given such that the eigenvalue of the elliptic boundary value problem is defective.

5.3. Higher dimensions. We extend the one-dimensional example to higher dimensions $d = 2, 3$ by taking the tensor product of the (generalized) eigenfunctions in x, y, and z coordinates, which leads to the eigenvalue 2λ in two dimensions of algebraic multiplicity $m_{alg} = 9$ and the eigenvalue 3λ for $d = 3$ with algebraic multiplicity $m_{alg} = 27$. The diffusion coefficient and the boundary conditions are extended by tensorization to higher dimensions as well.

In two dimensions, we observe in Figure 5.4 for $\mathbb{P}_1$ finite elements convergence of at least $\mathcal{O}(N^{-1/5})$ for the eigenvalue errors and $\mathcal{O}(N^{-1})$ for the mean eigenvalue error. For $\mathbb{P}_2$ finite elements we observe twice the convergence, namely, at least $\mathcal{O}(N^{-2/5})$ for the eigenvalue errors and $\mathcal{O}(N^{-2})$ for the mean eigenvalue error. This shows numerically the ascent $\alpha = 5$. In addition, we observe that some discrete eigenvalues converge with rates in between those two extreme cases. Some eigenvalues converge with order close to $\mathcal{O}(N^{-2/5})$ for $\mathbb{P}_1$ finite elements and close to $\mathcal{O}(N^{-4/5})$ for $\mathbb{P}_2$ finite elements. Note that one eigenvalue seems to correspond to a discrete eigenvector; hence it converges with the optimal rate.

For $d = 3$, we observe in Figure 5.5 convergence rates of the eigenvalue errors as low as $\mathcal{O}(N^{-2/21})$ for $\mathbb{P}_1$ finite elements and $\mathcal{O}(N^{-4/21})$ for $\mathbb{P}_2$ finite elements, which indicate the ascent $\alpha = 7$. Again the mean eigenvalue errors converge optimally.
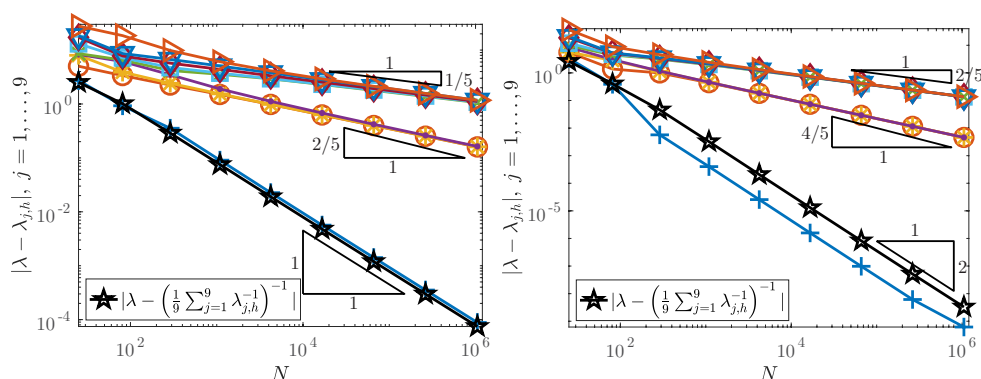

FIG. 5.4. *Regular example with $\mathbb{P}_1$ and $\mathbb{P}_2$ FEM in two dimensions (2d).*
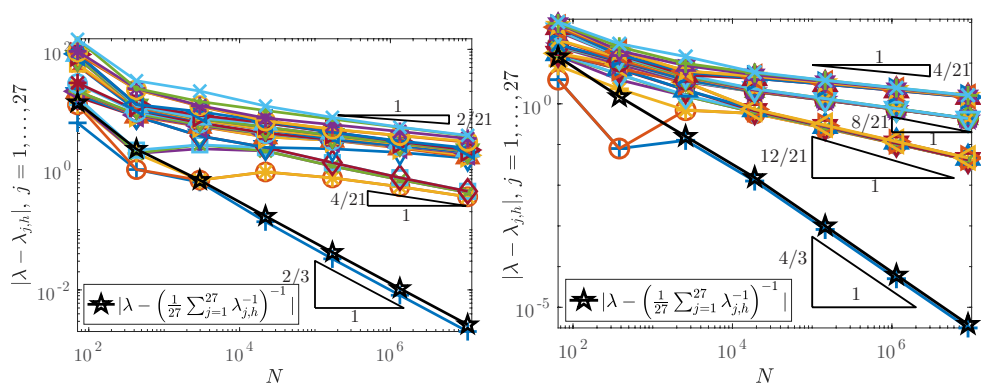

FIG. 5.5. *Regular example with $\mathbb{P}_1$ and $\mathbb{P}_2$ FEM in three dimensions.*

Note that some discrete eigenvalues converge with rates in between the optimal and reduced ones and that one eigenvalue converges with optimal rate. In particular for $\mathbb{P}_2$ finite elements we observe that some eigenvalues converge with order $\mathcal{O}(N^{-8/21})$, and some even with order $\mathcal{O}(N^{-12/21})$, which relates to convergence order of $\mathcal{O}(h^{2k/7})$ for $k = 2, 3$. This confirms the impressive sharpness of the theory and that in principle any convergence order $\mathcal{O}(h^{pk/\alpha})$ for $k = 1, \dots, \alpha$ can occur, not only in theory, but as we have demonstrated also in practical computations.

5.4. Examples with reduced regularity. Here, we choose the coefficient a such that its jump is not aligned with any (refined) mesh. Therefore, we consider $\Omega = (0,1)$ with $a = 1$ on $[0, 1/3]$ and $a = a_{\mathrm{R}}$ on $(1/3, 1]$, with

$$a_{\mathrm{R}} = 8.834634001449438 + 2.381273183203226i,$$
$$c = -23.62602259938114 + 23.10185194698031i,$$

and the first (smallest in magnitude) complex eigenvalue

$$\lambda = 72.26224904068889 + 65.85698689932984i.$$

By construction λ has algebraic multiplicity $m_{alg} = 3$ and ascent $\alpha = 3$ for $d = 1$. As in the previous example, the tensor product of the (generalized) eigenfunctions leads to the eigenvalue 2λ with algebraic multiplicity $m_{alg} = 9$ in two dimensions, and numerically we observe the ascent $\alpha = 5$.

Note that since the mesh is not aligned with the jump of the diffusion coefficient, the regularity of the (generalized) eigenfunctions is reduced to $H^{1+r}(\Omega)$ for any $0 < r < 1/2$. Therefore, we observe reduced convergence of the eigenvalues due to the reduced convergence of the (generalized) eigenfunctions on uniform meshes.

In Figure 5.6, we see that the convergence is reduced by two separate issues: the reduced regularity and the large defect of the eigenvalue. We observe the theoretically expected suboptimal convergence rates of the mean eigenvalue error of $\mathcal{O}(N^{-1})$ for both $\mathbb{P}_1$ and $\mathbb{P}_2$ finite elements due to the reduced regularity. The convergence of the eigenvalue errors is even further reduced due to the defect $\alpha = 3$; hence the convergence is only of order $\mathcal{O}(N^{-1/3})$.

The situation in two dimensions is less clear from the numerical point of view. In Figure 5.7, we observe in the left figure reduced rates of the mean eigenvalue error for $\mathbb{P}_1$ finite elements on uniform meshes, but still at worst $\mathcal{O}(N^{-1/5})$ convergence of
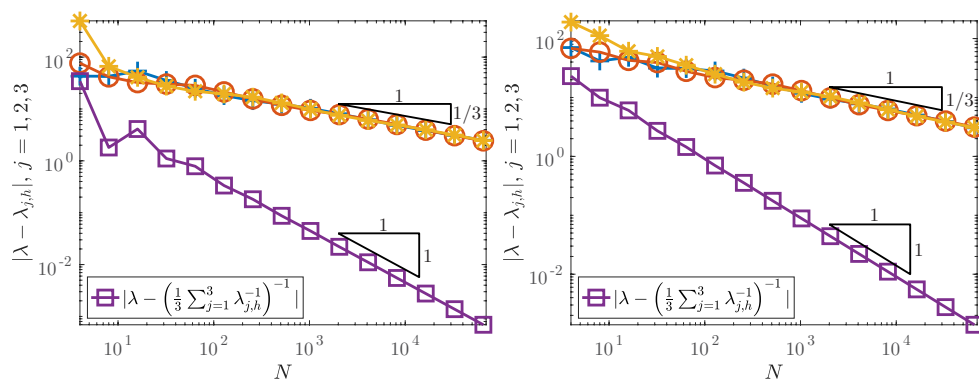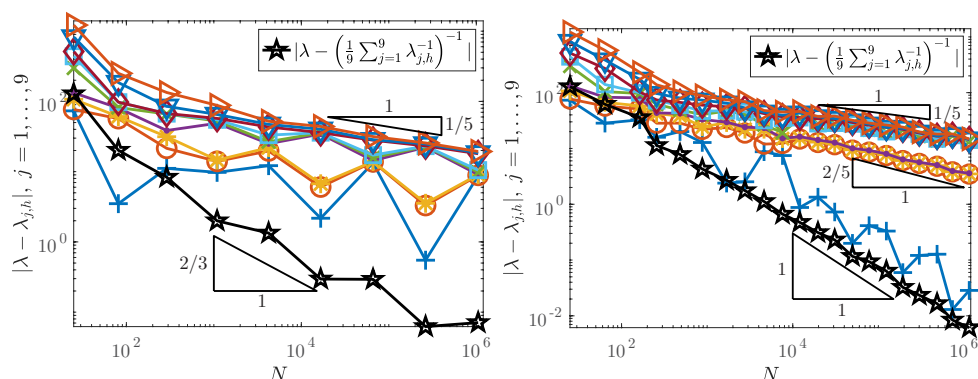

FIG. 5.6. *Example with reduced regularity for $\mathbb{P}_1$ and $\mathbb{P}_2$ FEM in 1d.*

FIG. 5.7. *Example with reduced regularity for* $\mathbb{P}_1$ *FEM on uniform and adaptive meshes in 2d.*

the eigenvalues, which is expected from the defect of λ but is not further decreased by the low regularity. This might be a preasymptotic effect. In the right figure we use an adaptive mesh refinement algorithm [14], [15], [19], [29]. Based on the previous observation, that even very small perturbations lead to a split of the defective eigenvalue into clustered eigenvalues, we measure the error of the defective eigenvalue, as if it was a cluster of eigenvalues, with the a posteriori error estimator

$$
\eta_h^2 := \sum_{j=1}^m \sum_{T \in \mathcal{T}} \Bigg(h_T^2 \|\Delta u_{j,h} + \lambda_{j,h} u_{j,h}\|_{0,T}^2 + \sum_{E \subset \partial T \setminus \partial \Omega} h_E \|[a \nabla u_{j,h} \cdot \mathbf{n}]\|_{0,E}^2
$$
$$
+ \sum_{E \subset \partial T \cap \Gamma_{\mathrm{R}}} h_E \|a \nabla u_{j,h} \cdot \mathbf{n} + c u_{j,h}\|_{0,E}^2
$$
$$
+ h_T^2 \|\Delta u_{j,h}^* + \lambda_{j,h}^* u_{j,h}^*\|_{0,T}^2 + \sum_{E \subset \partial T \setminus \partial \Omega} h_E \|[\overline{a} \nabla u_{j,h}^* \cdot \mathbf{n}]\|_{0,E}^2
$$
$$
+ \sum_{E \subset \partial T \cap \Gamma_{\mathrm{R}}} h_E \|\overline{a} \nabla u_{j,h}^* \cdot \mathbf{n} + \overline{c} u_{j,h}^*\|_{0,E}^2 \Bigg),
$$

where $(\lambda_{j,h}^*, u_{j,h}^*)$ denotes the jth eigenpair of the adjoint eigenvalue problem. Although η_h^2 provides only a valid upper bound for clustered eigenvalues, we observe in Figure 5.7 that adaptive mesh refinement based on η_h^2 leads to optimal convergence of the mean eigenvalue error. By construction, η_h^2 cannot give any a posteriori information about the ascent of the eigenvalue. Nevertheless, this experiment illustrates that an error estimator is in principle able to heuristically detect defective eigenvalues from their reduced convergence rates, although a theoretical foundation has still to be developed.

6. Conclusions. We described a constructive way of deriving benchmark problems with highly defective eigenvalues. We provided the parameters for two such examples. We confirmed in numerical experiments that the Babuška–Osborn theory is sharp and that convergence rates between the two extreme cases do occur in practical computations. Since even for nonsmooth eigenfunctions, the mean eigenvalue error converges faster, one is in principle able to detect defective eigenvalues numerically by tracking the convergence behavior of the eigenvalues and the mean eigenvalue error on uniformly or adaptively refined meshes. If the mean eigenvalue error converges faster, that means that the eigenvalue is defective.

REFERENCES

[1] I. BABUŠKA AND J. OSBORN, *Eigenvalue problems*, in Handbook of Numerical Analysis, Vol. II, Elsevier, New York, NY, 1991, pp. 641–787.

[2] S. C. BRENNER AND L. R. SCOTT, *The Mathematical Theory of Finite Element Methods*, 3rd ed., Springer, New York, NY, 2008.

[3] C. CARSTENSEN, J. GEDICKE, V. MEHRMANN, AND A. MIEDLAR, *An adaptive homotopy approach for non-selfadjoint eigenvalue problems*, Numer. Math., 119 (2011), pp. 557–583.

[4] F. CHATELIN, *La méthode de Galerkin. Ordre de convergence des éléments propres*, C.R. Math. Acad. Sci. Paris, 278 (1974), pp. 1213–1215.

[5] F. CHATELIN, *Spectral Approximation of Linear Operators*, Academic Press, New York, NY, 1983.

[6] P. CIARLET, *The Finite Element Method for Elliptic Problems*, North-Holland, Amsterdam, 1978.

[7] R. DAUTRAY AND J.-L. LIONS, *Mathematical Analysis and Numerical Methods for Science and Technology. Vol. 3, Spectral Theory and Applications*, Springer-Verlag, Berlin, 2000.

[8] J. W. DEMMEL, *Applied Numerical Linear Algebra*, Society for Industrial and Applied Mathematics, Philadelphia, PA, 1997.

[9] N. DUNFORD AND J. SCHWARTZ, *Linear Operators Part II: Spectral Theory*, Wiley-Interscience, New York, NY, 1963.

[10] E. DYAKONOV, *Optimization in Solving Elliptic Problems*, CRC Press, Boca Raton, FL, 1996.

[11] D. J. EWINS, *Modal Testing: Theory and Practice*, Research Studies Press, Letchworth, UK, 1984.

[12] B. FRIEDMAN, *Principles and Techniques of Applied Mathematics*, John Wiley & Sons, New York, NY, 1956.

[13] R. GASSER, *The Finite Element Method for Elliptic Problems with Defective Eigenvalues*, Master's thesis, Institut für Mathematik, Universität Zürich, Zürich, Switzerland 2017, http://www.math.uzh.ch/compmath/index.php?id=dipl.

[14] J. GEDICKE AND C. CARSTENSEN, *A posteriori error estimators for convection-diffusion eigenvalue problems*, Comput. Methods Appl. Mech. Engrg., 268 (2014), pp. 160–177.

[15] S. GIANI, L. GRUBIŠIĆ, A. MIEDLAR, AND J. S. OVALL, *Robust error estimates for approximations of non-self-adjoint eigenvalue problems*, Numer. Math., 133 (2016), pp. 471–495.

[16] P. GRISVARD, *Elliptic Problems in Nonsmooth Domains*, Pitman, Boston, MA, 1985.

[17] W. HACKBUSCH, *Elliptic Differential Equations*, Springer-Verlag, Berlin, 1992.

[18] Z.-F. FU AND J. HE, *Modal Analysis*, Elsevier, New York, NY, 2001.

[19] V. HEUVELINE AND R. RANNACHER, *A posteriori error control for finite approximations of elliptic eigenvalue problems*, Adv. Comput. Math., 15 (2001), pp. 107–138.

[20] J. JACKSON, *Classical Electrodynamics*, John Wiley & Sons, New York, NY, 1998.

[21] T. KATO, *Perturbation Theory for Linear Operators*, Springer-Verlag, Berlin, 1966.

[22] A. KNYAZEV, *Sharp a priori error estimates of the Raleigh-Ritz method without assumptions of fixed sign or compactness*, Math. Notes, 38 (1986), pp. 998–1002.

[23] A. V. KNYAZEV AND J. E. OSBORN, *New a priori FEM error estimates for eigenvalues*, SIAM J. Numer. Anal., 43 (2006), pp. 2647–2667.

[24] M. MACHOVER, *The alternative theorem and the nonselfadjoint generalized Green's function*, J. Differential Equations, 35 (1980), pp. 266–274.

[25] H. G. NATKE, *Einführung in Theorie und Praxis der Zeitreihen-und Modalanalyse: Identifikation schwingungsfähiger elastomechanischer Systeme*, Springer-Verlag, Berlin, 2013.

[26] E. OVTCHINNIKOV, *Cluster robust error estimates for the Rayleigh–Ritz approximation I: Estimates for invariant subspaces*, Linear Algebra Appl., 415 (2006), pp. 167–187.

[27] S. SAUTER, *hp-Finite elements for elliptic eigenvalue problems: Error estimates which are explicit with respect to λ, h, and p*, SIAM J. Numer. Anal., 48 (2010), pp. 95–108.

[28] G. STRANG AND G. FIX, *An Analysis of the Finite Element Method*, Prentice-Hall, Englewood Cliffs, NJ, 1973.

[29] Y. YANG, L. SUN, H. BI, AND H. LI, *A note on the residual type a posteriori error estimates for finite element eigenpairs of nonsymmetric elliptic eigenvalue problems*, Appl. Numer. Math., 82 (2014), pp. 51–67.