Convergence analysis of planewave expansion methods for Schrödinger operators with discontinuous periodic potentials

R. Norton and R. Scheichl

Bath Institute For Complex Systems

Preprint 6/09 (2009)

http://www.bath.ac.uk/math-sci/BICS

Convergence analysis of planewave expansion methods for Schrödinger operators with discontinuous periodic potentials

Richard Norton[*] Robert Scheichl[†]

April 22, 2009

Abstract

In this paper we consider the problem of computing the spectrum of a Schrödinger operator with discontinuous, periodic potential in two dimensions using Fourier (or planewave expansion) methods. Problems of this kind are currently of great interest in the design of new optical devices to determine band gaps and to compute localised modes in photonic crystal materials. Although Fourier methods may not be every applied mathematician's first choice for this problem because of the discontinuities in the potential, we will show here that, even though (as expected) the convergence is not exponential, the method has several desirable features that make it competitive with other discretisation techniques, such as finite element methods, both with respect to implementation and convergence properties. In particular, we will prove that simple preconditioners for the system matrix are optimal leading to a computational complexity of $\mathcal{O}(N \log N)$ in the number of planewaves N (using the Fast Fourier Transform). Moreover, we derive sharp error estimates that show that the method is essentially third order in the eigenvalues and of order $\frac{3}{2}$ in the eigenfunctions in the H^1-norm and $\frac{5}{2}$ in the L^2-norm. To improve the planewave expansion method in the case of discontinuous potentials, it has been proposed in the physics literature to replace the discontinuous potential with an *effective* potential that is smooth, despite the additional error this incurs. We will here answer the question whether this smoothing is worth it. In fact, our convergence analysis of the modified method provides an optimal choice for the smoothing parameter, but it also shows that the overall rate of convergence is no faster than before and so smoothing does not seem to be worth it. All the theoretical results are confirmed in our numerical experiments.

Key words. spectral approximation, PDE eigenproblems, Fourier methods, error analysis, optimal preconditioners, discontinuous coefficients, smoothing.

AMS subject classifications. 65N15, 65N25, 65N35, 65T40, 78M25

1 Introduction

Photonic crystal materials, i.e. periodic optical nanostructures which consist in the simplest case of a periodic arrangement of glass and air, are currently of great interest for their properties in guiding, bending or slowing down light [7, 9]. The design of these materials relies on mathematical modelling of light propagation, and in particular on being able to find gaps and

[*] Mathematical Institute, University of Oxford, 24-29 St Giles', Oxford OX1 3LB, UK. norton@maths.ox.ac.uk

[†] Dept of Mathematical Sciences, University of Bath, Bath BA2 7AY, UK. R.Scheichl@bath.ac.uk

localised modes in the spectrum of the underlying differential operators. In general this will be the Maxwell operator, but under some simplifying assumptions, e.g. in the case of optical fibres made of photonic crystal materials, this can be simplified to the Schrödinger operator

$$L := -\nabla^2 + V(\mathbf{x}) \tag{1}$$

on the Hilbert space $L^2(\mathbb{R}^2)$ (cf. [2, 16]). The potential $V(\mathbf{x})$ is related to the spatial distribution of the refractive index of the material and will therefore in general be discontinuous, piecewise constant.

It is well known (cf. [6, 10]) that periodic potentials $V(\mathbf{x}) = V_p(\mathbf{x})$ lead in general to a band structure in the essential spectrum of L, and that compact perturbations of these periodic nanostructures allow for localised modes (or L^2-eigenvalues). In this paper we consider the problem of numerically computing the spectrum of L in the case of discontinuous, periodic potentials $V(\mathbf{x}) = V_p(\mathbf{x})$ using Fourier (or planewave expansion) methods. To be able to apply these methods also in the context of periodic potentials with a compact perturbation $V(\mathbf{x}) = V_p(\mathbf{x})+V_c(\mathbf{x})$, e.g. to derive localised modes, we resort to the so-called *supercell method*. We replace $V(\mathbf{x})$ with a periodic potential $V_p^{\text{super}}(\mathbf{x})$ with sufficiently large period cell (the *supercell*). Since the essential spectrum is not affected by the compact perturbation (cf. [6]), it can be computed by discarding $V_c(\mathbf{x})$ and using $V(\mathbf{x}) = V_p(\mathbf{x})$ in (1). Localised modes, on the other hand, can be computed with $V(\mathbf{x}) = V_p^{\text{super}}(\mathbf{x})$, and due to the exponential decay of the localised eigenmodes (cf. [10]) the convergence of this supercell method is exponential in the diameter of the supercell, i.e. any localised mode of the compactly perturbed periodic operator is approximated by a thin band of essential spectrum whose width decreases exponentially with the diameter of the supercell. This follows directly from the convergence analysis for a slightly different problem in [20] (although this seems to be unpublished for the operator L considered here). Henceforth we will only consider periodic potentials $V(\mathbf{x})$, but we note that the local variation within the period cell may be complicated. By applying the so-called Floquet-Bloch transform we can reduce the calculation of the spectrum of L to a family of variational eigenproblems on the period cell of $V(\mathbf{x})$, which we then discretise using Fourier methods. The arising matrix eigenproblems are solved by Krylov subspace iteration.

The motivation for studying Fourier methods for (1) stems from the fact that they are well suited for periodic problems and very popular in nonlinear optics and solid state physics, as they are very simple to implement and can be extremely fast and accurate. However, when applied to problems with piecewise smooth or discontinuous coefficients much of this efficiency and accuracy is lost (see [3, 21] for some numerical studies). Probably due to this fact, there is comparatively little material in the mathematical literature on this problem. The only theoretical paper that we could find is [15], but this is restricted to 1D and considers a slightly different spectral problem with discontinuous coefficients.

Our basic analysis will rely on abstract theory for Galerkin methods applied to variational eigenvalue problems, as presented in [1], and on Fourier approximation results that can to a large extent be found in [19]. The key regularity results are derived from the theory in [12, 13]. We show that certain classes of piecewise smooth periodic functions $V(\mathbf{x})$ are in the Sobolev spaces $H^{1/2-\varepsilon}$ for all $\varepsilon > 0$, which in a straightforward way leads to a convergence of the eigenfunctions of order $\frac{3}{2} - \varepsilon$ in the H^1-norm. The eigenvalues converge (as expected) at twice this rate. By carefully studying the decay of the Fourier coefficients of $V(\mathbf{x})$, we are able to get rid of ε and show that the eigenfunctions actually converge with order $\frac{3}{2}$, which is confirmed in our numerical experiments. Using a standard duality argument we can then also deduce that the eigenfunctions converge with order $\frac{5}{2}$ in the L_2-norm.

To improve the planewave expansion method in the case of discontinuous potentials, it has been proposed in the physics literature [8, 14, 17, 18] to replace the discontinuous potential with an *effective* potential that is smooth. However, this leads to an additional error, and so the question arises whether smoothing is really worth it. Using Strang's lemma and the abstract error analysis in [1] in a non-standard way we show that in fact the overall rate of convergence is indeed no faster than for the standard method and we give an optimal choice for the smoothing parameter.

Even though exponential convergence of the method is as expected not achieved, the method has several desirable features that make it competitive with other discretisation techniques, such as finite element (FE) methods. Firstly, without specially adapted meshes, FEs lead to the same order of convergence with respect to the number of degrees of freedom (if at least quadratic elements are used). For better convergence rates it is necessary to use adaptive FEs. Secondly, the computational complexity of both methods hinges on robust preconditioners that guarantee that the convergence of the iterative eigensolver is independent of the number of degrees of freedom. In the case of FEs it is necessary to use multilevel methods such as multigrid to achieve this, whereas we will show here (rigorously) that for Fourier methods a simple diagonal preconditioner is sufficient, and so we can guarantee a total computational cost of order $\mathcal{O}(N \log N)$ in the number N of Fourier modes for our method (dominated by the application of the Fast Fourier Transform in each iteration).

The paper is organised as follows. We start in §2 by defining the problem and describing the method. In §3 we analyse the regularity of piecewise smooth, periodic functions, which will be crucial for our convergence analysis in §4. In §5 we give details on the numerical implementation and analyse optimal preconditioners, followed by some numerical results in §6. Section 7 is devoted to the issue of smoothing, and we finish the paper in §8 with some numerical results that confirm that smoothing is not worth it.

2 Problem Definition and Planewave Expansion

Let us consider the differential operator

$$L := -\nabla^2 + V(\mathbf{x}) + K, \tag{2}$$

on the Hilbert space $L^2(\mathbb{R}^2)$ with domain $D(L) = H^2(\mathbb{R}^2)$, where $V \in L^\infty(\mathbb{R}^2)$ is periodic on the Bravais lattice $\mathbb{Z}^2$ and $K \geq \|V\|_{L^\infty(\mathbb{R}^2)} + 2\pi^2 + \frac{1}{2}$ is a constant (which ensures that the spectrum of L is positive). We are interested in computing the spectrum $\sigma(L)$ of L in the case when $V(\mathbf{x})$ may be discontinuous and is only piecewise smooth. We choose the period cell for the lattice $\mathbb{Z}^2$ to be $\Omega := (-\frac{1}{2}, \frac{1}{2})^2$ and set $L^2_p := \{f \in L^2_{loc}(\mathbb{R}^2) : f \text{ is periodic on } \mathbb{Z}^2\}$ with the usual $L^2(\Omega)$ inner product. The extension to more general lattices is straight forward.

The operator L is positive definite and self-adjoint, and so the spectrum of L is a subset of the positive real axis. Moreover, it is well-known that the spectrum of L is absolutely continuous (see [10] and references therein) which implies that the spectrum consists of purely essential spectrum. Since the coefficients of L are periodic we can apply the Floquet–Bloch transform to this problem (see e.g. [10, 20] for details). To do this we need to first introduce periodic Sobolev spaces on $\mathbb{R}^2$.

2.1 Periodic Sobolev Spaces

Let $\mathcal{D}(\mathbb{R}^2) := C_0^\infty(\mathbb{R}^2)$ be the usual space of test functions and let $\mathcal{D}'(\mathbb{R}^2)$ be the set of distributions, i.e. the set of sequentially continuous linear functionals on $\mathcal{D}(\mathbb{R}^2)$. A distribution $u \in \mathcal{D}'(\mathbb{R}^2)$ is periodic, if

$$\langle u, \tau_{\mathbf{n}}\phi\rangle = \langle u, \phi\rangle, \qquad \text{for all } \ \phi \in \mathcal{D}(\mathbb{R}^2) \ \text{ and } \mathbf{n} \in \mathbb{Z}^2,$$

where $\tau_{\mathbf{n}}\phi(\mathbf{x}) := \phi(\mathbf{x}+\mathbf{n})$ for all $\mathbf{x} \in \mathbb{R}^2$. The set of all periodic distributions is denoted by $\mathcal{D}'_p(\mathbb{R}^2)$.

To define Fourier expansions of periodic distributions it is useful to introduce a function $\theta : \mathbb{R}^2 \to \mathbb{R}$ such that

$$\theta \in \mathcal{D}(\mathbb{R}^2), \qquad 0 \le \theta \le 1, \qquad \text{and} \qquad \sum\nolimits_{\mathbf{n}\in\mathbb{Z}^2} \tau_{\mathbf{n}}\theta = 1. \tag{3}$$

See [16] for a simple example of a function θ that satisfies (3). Then, for any $u \in \mathcal{D}'_p(\mathbb{R}^2)$ and $\mathbf{n} \in \mathbb{Z}^2$ the Fourier coefficient of u with index $\mathbf{n}$ is defined by

$$[u]_{\mathbf{n}} := \langle u, \psi_{\mathbf{n}}\rangle\,, \qquad \text{where} \qquad \psi_{\mathbf{n}}(\mathbf{x}) := \theta(\mathbf{x})\mathrm{e}^{-i2\pi\mathbf{n}\cdot\mathbf{x}}\,. \tag{4}$$

Since θ is not uniquely defined by (3), it might appear that $[u]_{\mathbf{n}}$ depends on the choice of θ. However, it can easily be shown that this is not the case (cf. [16, 19]).

It follows from (4) and from convergence in $\mathcal{D}'_p(\mathbb{R}^2)$ that every periodic distribution $u \in \mathcal{D}'_p(\mathbb{R}^2)$ can be identified with its Fourier expansion such that

$$u(\mathbf{x}) = \sum\nolimits_{\mathbf{n}\in\mathbb{Z}^2} [u]_{\mathbf{n}}\mathrm{e}^{i2\pi\mathbf{n}\cdot\mathbf{x}}, \qquad \text{for all } \ \mathbf{x} \in \mathbb{R}^2\,, \tag{5}$$

where equality has to be understood in the distributional sense. A proof of this result in 1D can be found in [19], while the obvious extension to $\mathbb{R}^d$, $d \in \mathbb{N}$, is given in [16].

We can now define periodic Sobolev spaces and their corresponding norms. For $s \in \mathbb{R}$ we define

$$H_p^s := \{u \in \mathcal{D}'_p(\mathbb{R}^2) : \|u\|_{H_p^s} < \infty\}\,, \qquad \text{where}$$

$$\|u\|^2_{H_p^s} := \sum\nolimits_{\mathbf{n}\in\mathbb{Z}^2} |\mathbf{n}|_\star^{2s}|[u]_{\mathbf{n}}|^2 \qquad \text{and} \qquad |\mathbf{n}|_\star := \begin{cases} 1, & \text{if } \ \mathbf{n} = \mathbf{0}, \\ |\mathbf{n}|, & \text{if } \ \mathbf{n} \neq \mathbf{0}. \end{cases}$$

The space H_p^s is complete with respect to this norm and it is a Hilbert space with inner product

$$(u,v)_{H_p^s} := \sum\nolimits_{\mathbf{n}\in\mathbb{Z}^2} |\mathbf{n}|_\star^{2s}[u]_{\mathbf{n}}\overline{[v]_{\mathbf{n}}}\,, \qquad \text{for } \ u, v \in H_p^s.$$

Note that (using (5)) we can identify H_p^0 with L_p^2 and the corresponding inner products are equal. The following Theorem contains some important results about periodic Sobolev spaces which we will require later.

Theorem 1. *Let $s, t \in \mathbb{R}$.*

1. *If $s < t$, then $H_p^t \subset\subset H_p^s$.*

2. *If $s \le t$, $u \in H_p^s$ and $\tau \in [0,1]$, then*

$$\|u\|_{H_p^{\tau s+(1-\tau)t}} \le \|u\|^{\tau}_{H_p^s}\|u\|^{1-\tau}_{H_p^t}.$$

3. *If $s > 1$, then $u \in H_p^s$ is continuous and*

$$\|u\|_{L^\infty} \leq C(s)\|u\|_{H_p^s}\,, \qquad \text{where} \qquad C(s) = \left(\sum\nolimits_{\mathbf{n}\in\mathbb{Z}^2} |\mathbf{n}|_\star^{-2s}\right)^{\frac{1}{2}}.$$

4. *If $t > 1$, $a \in H_p^{\max(|s|,t)}$ and $u \in H_p^s$, then*

$$\begin{aligned} \|au\|_{H_p^s} &\leq C(s)\|a\|_{H_p^{|s|}}\|u\|_{H_p^s}\,, \qquad &\text{if } |s| > 1,\\ \|au\|_{H_p^s} &\leq C(t)\|a\|_{H_p^t}\|u\|_{H_p^s}\,, \qquad &\text{if } |s| \leq 1, \end{aligned}$$

where $C(s)$ and $C(t)$ are constants independent of a and u.

5. *If $s > 0$, then with θ satisfying* (3),

$$\|u\|_{H_p^s} \simeq \|u\|_{H^s(\Omega)} \simeq \|\theta u\|_{H^s(\mathbb{R}^2)}\,,$$

where $\|\cdot\|_{H^s(\Omega)}$ and $\|\cdot\|_{H^s(\mathbb{R}^2)}$ are defined in the usual way (see for example [13]).

Parts 1–4 are standard results for Sobolev spaces adapted to the periodic case. Part 5 shows that periodic Sobolev space norms are equivalent to the usual Sobolev space norms. Detailed proofs for all these results can be found in [16] but they are all heavily based on standard results on Sobolev spaces in [5, 12, 13, 19].

2.2 The Floquet–Bloch Transform

Since we assumed the coefficients of L to be periodic, we can apply the Floquet–Bloch transform to this problem to obtain a family of operators, parametrised by $\boldsymbol{\xi} \in B = [-\pi,\pi]^2$, on the bounded domain $\Omega = (-\frac{1}{2},\frac{1}{2})^2$ with periodic boundary conditions. See [10, 20] and references therein for details. For each $\boldsymbol{\xi} \in B$, we consider

$$L_{\boldsymbol{\xi}} := -(\nabla + i\boldsymbol{\xi})^2 + V(\mathbf{x}) + K$$

on the Hilbert space L_p^2 with domain $D(L_{\boldsymbol{\xi}}) := H_p^2$. Note that $L_{\boldsymbol{\xi}}$ is self-adjoint and compact, and hence the spectrum of $L_{\boldsymbol{\xi}}$ is real and discrete. Moreover, $\lambda(\boldsymbol{\xi}) \in \sigma(L_{\boldsymbol{\xi}})$ considered as a function of $\boldsymbol{\xi}$, is continuous on B. It is often referred to as a *band* in the essential spectrum of L. The key result from Floquet–Bloch theory which we use to find the spectrum $\sigma(L)$ of our original operator L is that

$$\sigma(L) = \bigcup\nolimits_{\boldsymbol{\xi}\in B} \sigma(L_{\boldsymbol{\xi}}). \tag{6}$$

In the physics literature the set B is usually referred to as the *1st Brillouin Zone* and $\boldsymbol{\xi}$ as the *quasi-momentum.*

It follows from (6) that the spectrum of the operator L, which is defined on all of $\mathbb{R}^2$, can be computed by solving a family of eigenproblems on the bounded domain Ω, which is numerically more practical. This is what commonly is done in practice. Hence, for the remainder of this paper we restrict our attention to the problem of approximating the spectrum of $L_{\boldsymbol{\xi}}$ for a fixed $\boldsymbol{\xi} \in B$.

Since the spectrum of $L_{\boldsymbol{\xi}}$ is discrete we need to only consider the eigenproblem

$$L_{\boldsymbol{\xi}} u = \lambda u\,, \qquad \text{for } \mathbf{x} \in \Omega, \tag{7}$$

subject to periodic boundary conditions, or its weak form: Find $\lambda \in \mathbb{R}$ and $0 \neq u \in H_p^1$ such that

$$a(u,v) = \lambda\, b(u,v)\,, \qquad \text{for all } \ v \in H_p^1, \tag{8}$$

where

$$\begin{aligned} a(u,v) &:= \int_\Omega (\nabla + i\boldsymbol{\xi})u \cdot \overline{(\nabla + i\boldsymbol{\xi})v} + V(\mathbf{x})u\overline{v} + Ku\overline{v}\; d\mathbf{x}\,, \\ b(u,v) &:= \int_\Omega u\overline{v}\; d\mathbf{x}\,. \end{aligned}$$

It is a simple calculation to show that $a(\cdot,\cdot)$ is bounded, coercive and Hermitian on H_p^1. Also note that $b(\cdot,\cdot)$ is equal to the usual inner-product of L_p^2.

2.3 The Planewave Expansion Method

The planewave expansion method is a numerical method for solving (7) by expanding u in terms of a finite number of planewaves. In this paper we will represent it as a Galerkin method applied to (8). To do this we first need to define finite dimensional subspaces of H_p^1, formed by taking the span of a finite number of planewaves (or Fourier basis functions). For $G \in \mathbb{N}$, let

$$\mathbb{Z}_G^2 := \left\{ \mathbf{n} \in \mathbb{Z}^2 : |\mathbf{n}| \leq G \right\}\,,$$

and define the following trigonometric function space

$$\mathcal{S}_G := \operatorname{span}\{ \mathrm{e}^{i2\pi\mathbf{n}\cdot\mathbf{x}} : \mathbf{n} \in \mathbb{Z}_G^2 \}\,.$$

The dimension of $\mathcal{S}_G$ is $\mathcal{O}(G^2)$. The set $\{\mathrm{e}^{i2\pi\mathbf{n}\cdot\mathbf{x}} : \mathbf{n} \in \mathbb{Z}_G^2\}$ is an orthogonal basis for $\mathcal{S}_G$ (with respect to the L_p^2 inner product) and we call it a *Fourier basis*. Each member of the Fourier basis is called a *Fourier basis function* or *planewave*.

Applying the Galerkin method to (8) and restricting to $\mathcal{S}_G$ for some $G \in \mathbb{N}$ results in the planewave expansion method: Find $\lambda_G \in \mathbb{R}$ and $0 \neq u_G \in \mathcal{S}_G$ such that

$$a(u_G, v_G) = \lambda_G\, b(u_G, v_G)\,, \qquad \text{for all } \ v_G \in \mathcal{S}_G. \tag{9}$$

This problem (since it is finite dimensional) can be rewritten as a matrix eigenvalue problem. Expanding u_G in terms of the Fourier basis for $\mathcal{S}_G$ yields

$$u_G(\mathbf{x}) = \textstyle\sum_{\mathbf{n}\in\mathbb{Z}_G^2} c_\mathbf{n} \mathrm{e}^{i2\pi\mathbf{n}\cdot\mathbf{x}}\,, \tag{10}$$

where the coefficients $c_\mathbf{n}$ are the Fourier coefficients of u_G, i.e $c_\mathbf{n} = [u_G]_\mathbf{n}$. Therefore, (9) is equivalent to the following $N := \dim(\mathcal{S}_G)$ simultaneous equations:

$$\sum_{\mathbf{n}\in\mathbb{Z}_G^2} c_\mathbf{n} a(\mathrm{e}^{i2\pi\mathbf{n}\cdot\mathbf{x}}, \mathrm{e}^{i2\pi\mathbf{m}\cdot\mathbf{x}}) = \lambda_G \sum_{\mathbf{n}\in\mathbb{Z}_G^2} c_\mathbf{n} b(\mathrm{e}^{i2\pi\mathbf{n}\cdot\mathbf{x}}, \mathrm{e}^{i2\pi\mathbf{m}\cdot\mathbf{x}})\,, \quad \text{for all } \mathbf{m} \in \mathbb{Z}_G^2. \tag{11}$$

Let us define a vector $\mathbf{u}$ of length N that has entries $u_\mathbf{n} = c_\mathbf{n}$ with index $\mathbf{n} \in \mathbb{Z}_G^2$. In practice, we order the entries of $\mathbf{u}$ in ascending order of magnitude of $\mathbf{n}$. Now define a $N \times N$ matrix A with entries $A_{\mathbf{mn}} = a(\mathrm{e}^{i2\pi\mathbf{n}\cdot\mathbf{x}}, \mathrm{e}^{i2\pi\mathbf{m}\cdot\mathbf{x}})$, for $\mathbf{m},\mathbf{n} \in \mathbb{Z}_G^2$. Then

$$\begin{aligned} A_{\mathbf{mn}} &= \int_\Omega (\nabla + i\boldsymbol{\xi})\, \mathrm{e}^{i2\pi\mathbf{m}\cdot\mathbf{x}} \cdot \overline{(\nabla + i\boldsymbol{\xi})\, \mathrm{e}^{i2\pi\mathbf{n}\cdot\mathbf{x}}} + (V(\mathbf{x}) + K)\mathrm{e}^{i2\pi\mathbf{m}\cdot\mathbf{x}}\overline{\mathrm{e}^{i2\pi\mathbf{n}\cdot\mathbf{x}}} dx \\ &= ((\xi + 2\pi\mathbf{m}) \cdot (\xi + 2\pi\mathbf{n}) + K) \int_\Omega \mathrm{e}^{i2\pi(\mathbf{m}-\mathbf{n})\cdot\mathbf{x}} dx + \int_\Omega V(\mathbf{x})\mathrm{e}^{i2\pi(\mathbf{m}-\mathbf{n})\cdot\mathbf{x}} dx \qquad (12) \\ &= (|\xi + 2\pi\mathbf{n}|^2 + K)\delta_{\mathbf{n},\mathbf{n}} + [V]_{\mathbf{n}-\mathbf{m}}\,. \end{aligned}$$

If we use this together with the fact that

$$b(\mathrm{e}^{i2\pi\mathbf{n}\cdot\mathbf{x}}, \mathrm{e}^{i2\pi\mathbf{m}\cdot\mathbf{x}}) = \delta_{\mathbf{n},\mathbf{m}} \,, \qquad\qquad \text{for all} \;\; \mathbf{m}, \mathbf{n} \in \mathbb{Z}^2_G \,,$$

we can write (11) as the matrix eigenvalue problem

$$A\mathbf{u} = \lambda_G \mathbf{u}. \tag{13}$$

From (12) we can see that $A = D + W$ where D is a diagonal matrix with entries $D_{\mathbf{nn}} = |\boldsymbol{\xi} + 2\pi\mathbf{n}|^2 + K$ and W is a dense matrix (in general) with entries $W_{\mathbf{mn}} = [V]_{\mathbf{n}-\mathbf{m}}$. This special form of the matrix A is due to the choice of basis functions for $\mathcal{S}_G$ and the fact that they are eigenfunctions of the Laplacian, which are orthogonal with respect to the L^2_p inner product. The fact that $a(\cdot,\cdot)$ is Hermitian and coercive directly implies that A is Hermitian and positive definite.

To analyse the convergence of the planewave expansion method we need to restrict to a certain class of piecewise smooth periodic potentials $V(\mathbf{x})$ in (2).

3 Two Regularity Classes for Piecewise Smooth Periodic Functions

We define two special classes of functions that we will refer to throughout the paper. We will then study certain examples of piecewise smooth periodic functions and check whether they fall into either of these regularity classes.

Definition 2. *For $f \in \mathcal{D}'_p(\mathbb{R}^2)$ and $n \in \mathbb{N}$ define*

$$F_n(f) := \bigg(\sum_{|g_1|+|g_2|=n} |[f]_{\mathbf{g}}|^2 \bigg)^{\frac{1}{2}}.$$

The two classes of periodic functions are

$$\begin{aligned} \mathcal{X}_p &:= \{f \in H_p^{1/2-\varepsilon} \text{ for any } \varepsilon > 0\} \cap L^\infty(\mathbb{R}^2) \quad \text{and} \\ \mathcal{Y}_p &:= \{f \in \mathcal{D}'_p(\mathbb{R}^2) : F_n(f) \lesssim n^{-1} \text{ for all } n \in \mathbb{N}\} \cap L^\infty(\mathbb{R}^2). \end{aligned}$$

By the definition of $\|\cdot\|_{H^s_p}$ it is clear that $\mathcal{Y}_p \subset \mathcal{X}_p$. The converse is not true in general and $\mathcal{X}_p \neq \mathcal{Y}_p$. In practical terms we may ask what kind of functions are in $\mathcal{X}_p$ and $\mathcal{Y}_p$. To do this let us consider the following piecewise smooth periodic functions.

Definition 3. *Let $J \in \mathbb{N}$ and let $\{\Omega_j, j = 1, \ldots, J\}$ be a set of disjoint Lipschitz domains such that $\Omega_j \subset\subset \Omega$. (See [13] for a definition of a Lipschitz domain.) Consider a function V that can be expressed as the sum of J periodic functions V_j:*

$$V = V_0 + \sum_{j=1}^{J} V_j \tag{14}$$

such that $\mathrm{supp}(V_j) \cap \Omega \subset \overline{\Omega_j}$, $V_0 \in C^\infty(\mathbb{R}^2)$ *and* $V_j|_{\Omega_j} \in C^\infty(\overline{\Omega_j})$.

Proposition 4. *Let V be defined as in Definition 3. Then $V \in \mathcal{X}_p$.*

Proof. It is obvious that $V \in L^\infty(\mathbb{R}^2)$. To see that $V \in H_p^{1/2-\varepsilon}$ for any $\varepsilon > 0$ let us sketch the proof of [16, Theorem 3.40].

Let $s < \frac{1}{2}$. Using (14) it suffices to show that $V_j \in H_p^s$ for each j. It follows from the definition of a Lipschitz domain that there exists a finite open covering $\{W_k\}_{k=1}^K$ of $\partial\Omega_j$ such that each function f with supp$f \subset W_k \cap \overline{\Omega_j}$ and $f|_{\Omega_j} \in C^\infty(\overline{\Omega_j})$ can be transformed through a uniformly Lipschitz continuous bijective mapping κ such that

$$f \circ \kappa(\mathbf{x}) = \begin{cases} \psi(\mathbf{x}), & \text{for } x_2 \le 0, \\ 0, & \text{for } x_2 > 0, \end{cases}$$

for some $\psi \in C_0^\infty(\mathbb{R}^2)$. Note that $\kappa(W_k \cap \partial\Omega_j) \subset \{x_2 = 0\}$. Also, let W_{K+1} cover the remainder of Ω_j, i.e. $W_{K+1} \supseteq \Omega_j \backslash \bigcup_{k=1}^K W_k$ and define a partition of unity $\{\phi_k\}_{k=1}^{K+1}$ of Ω_j such that $\phi_k \in C^\infty(\mathbb{R}^2)$, supp$\phi_k \subset W_k$ and $\sum \phi_k = 1$ on Ω_j. Finally, define $\theta(\mathbf{x})$ according to (3) with the additional restriction that $\theta(\mathbf{x}) = 1$ for $\mathbf{x} \in \Omega_j$ (which is possible since $\Omega_j \subset\subset \Omega$). Then, using Part 5 of Theorem 1 and the triangle inequality we can write

$$\|V_j\|_{H_p^s} \lesssim \|\theta V_j\|_{H^s(\mathbb{R}^2)} \le \sum_{k=1}^{K+1} \|\phi_k \theta V_j\|_{H^s(\mathbb{R}^2)} \,.$$

Let $f = \phi_k \theta V_j$ and let κ and ψ be defined as above. Then it follows from [13, Exercise 3.22 and Theorem 3.23] that $f \in H^s(\mathbb{R}^2)$. (See [16, Appendix A2] for a proof of [13, Exercise 3.22]). □

Conjecture 5. *Let V be defined as in Definition 3. Then $V \in \mathcal{Y}_p$.*

Unfortunately, we have so far not been able to prove this conjecture for the case of arbitrary Lipschitz domains Ω_j. However, we have managed to prove it for example in the following two special cases.

Proposition 6. *Let V be defined as in Definition 3. In addition, assume that each Ω_j is a convex Lipschitz polygon with finitely many corners. Then $V \in \mathcal{Y}_p \subset \mathcal{X}_p$.*

Proof. Again, it is obvious that $V \in L^\infty(\mathbb{R}^2)$. To see that $F_n \lesssim n^{-1}$ for all $n \in \mathbb{N}$, we proceed as in the proof of Proposition 4 by defining a finite open cover of each Ω_j, $\{W_{jk}\}_{k=1}^{K+1}$ so that each W_{jk} with $k \le K$ covers either a corner or a straight edge of $\partial\Omega$ and $W_{j(K+1)} \cap \partial\Omega_j = \emptyset$. Moreover, we restrict our choice of W_{jk} so that $W_{jk} \subset \Omega$ (possible since $\Omega_j \subset\subset \Omega$). Define a partition of unity $\{\phi_{jk}\}_{k=1}^{K+1}$ for Ω_j such that $\phi_{jk} \in C^\infty(\mathbb{R}^2)$ and supp$\phi_{jk} \subset W_{jk} \subset \Omega$ for each $k = 1, \ldots, K+1$ and $\sum_k \phi_{jk} = 1$ on Ω_j. Define $\tilde{\phi}_{jk}$ as the periodic extension of $\phi_{jk}|_\Omega$ to $\mathbb{R}^2 \backslash \Omega$. ¿From the definition of $F_n(V)$, using (14) and the triangle inequality, we get

$$F_n(V)^2 \le (J+1)^2(K+1)^2 \Bigg(\underbrace{\sum_{|g_1|+|g_2|=n} |[V_0]_{\mathbf{g}}|^2}_{=:I_0(n)} + \sum_{j=1}^{J} \sum_{k=1}^{K+1} \underbrace{\sum_{|g_1|+|g_2|=n} |[\tilde{\phi}_{jk} V_j]_{\mathbf{g}}|^2}_{=:I_{jk}(n)} \Bigg) \,.$$

Given $r \in \mathbb{N}$, since V_0 is in $C^\infty(\mathbb{R}^2)$, we can use integration by parts to show that $|[V_0]_{\mathbf{g}}| \lesssim |\mathbf{g}|^{-r}$ for all $\mathbf{0} \ne \mathbf{g} \in \mathbb{Z}^2$, and hence $I_0(n) \lesssim n^{-2}$, for all $n \in \mathbb{N}$.

It remains to bound $I_{jk}(n)$. Let us fix j and k and consider each $|[\tilde{\phi}_{jk} V_j]_{\mathbf{g}}|$ separately. If $k = K+1$, then $\tilde{\phi}_{j(K+1)} V_j \in C^\infty(\mathbb{R}^2)$ and we can again use integration by parts to show that

$I_{j(K+1)}(n) \lesssim n^{-2}$ for all $n \in \mathbb{N}$. Thus, we can assume that $k \le K$. There are two possibilities: either W_{jk} covers a corner of $\partial\Omega_j$ or W_{jk} covers a straight edge of $\partial\Omega_j$.

If W_{jk} covers a straight edge of $\partial\Omega$, then we can define a map $\kappa := S \circ T$ where S is a rotation and T is a translation so that

$$\phi_{jk}V_j \circ \kappa|_\Omega(\mathbf{x}) = \begin{cases} f(\mathbf{x}), & \text{for } x_2 \le 0, \\ 0, & \text{for } x_2 > 0, \end{cases}$$

for some $f \in C_0^\infty(\mathbb{R}^2)$. For $\mathbf{0} \ne \mathbf{g} \in \mathbb{Z}^2$ and $\mathbf{h} := S^{-1}\mathbf{g}$, if $h_1 \ne 0$ and $h_2 \ne 0$, then

$$\begin{aligned} |[\tilde{\phi}_{jk}V_j]_{\mathbf{g}}| &= \left|\int_\Omega (\phi_{jk}V_j)(\mathbf{x})\mathrm{e}^{-i2\pi\mathbf{g}\cdot\mathbf{x}}d\mathbf{x}\right| = \left|\int_{y_2<0} f(\mathbf{y})\mathrm{e}^{-i2\pi\mathbf{h}\cdot\mathbf{y}}d\mathbf{y}\right| \\ &= \tfrac{1}{4\pi^2|h_1||h_2|}\left|\int_{y_2=0} \tfrac{\partial g}{\partial y_1}\mathrm{e}^{-i2\pi\mathbf{h}\cdot\mathbf{y}}d\mathbf{y} + \int_{y_2<0} \tfrac{\partial^2 g}{\partial y_2 \partial y_1}\mathrm{e}^{-i2\pi\mathbf{h}\cdot\mathbf{y}}d\mathbf{y}\right| \lesssim |h_1|^{-1}|h_2|^{-1}. \end{aligned}$$

If $h_1 = 0$ or $h_2 = 0$ (i.e. when $\mathbf{g}$ is parallel or perpendicular to $\partial\Omega_j$) then we can not carry out both integrations by parts and so

$$|[\tilde{\phi}_{jk}V_j]_{\mathbf{g}}| \lesssim |h_1|_\star^{-1}|h_2|_\star^{-1} \qquad \forall \mathbf{g} \in \mathbb{Z}^2,\ \mathbf{h} = S^{-1}(\mathbf{g})\ . \tag{15}$$

Alternatively, if W_{jk} covers a corner of $\partial\Omega_j$, then we can define a map $\kappa := S \circ T$ (where S is a rotation and T is a translation) so that

$$\phi_{jk}V_j \circ \kappa|_\Omega(\mathbf{x}) = \begin{cases} f(\mathbf{x}), & \mathbf{x} \in \Omega_s\,, \\ 0, & \mathbf{x} \in \mathbb{R}^2\backslash\Omega_s\,, \end{cases}$$

for $c \in \mathbb{R}$, $\Omega_s = \{\mathbf{x} \in \mathbb{R}^2 : x_1 \ge 0 \text{ and } x_2 \le cx_1\}$ and some $f \in C_0^\infty(\mathbb{R}^2)$. Using integration by parts as in (15), for $\mathbf{0} \ne \mathbf{g} \in \mathbb{Z}^2$ and $\mathbf{h} := S^{-1}\mathbf{g}$, we get

$$|[\tilde{\phi}_{jk}V_j]_{\mathbf{g}}| = \left|\int_{\Omega_s} g(\mathbf{y})\mathrm{e}^{-i2\pi\mathbf{h}\cdot\mathbf{y}}d\mathbf{y}\right| \lesssim |h_2|_\star^{-1}|h_1 + ch_2|_\star^{-1}. \tag{16}$$

Note that $h_2 = 0$ and $h_1 + ch_2 = 0$ correspond to $\mathbf{g} \in \mathbb{Z}^2$ that are perpendicular to one of the edges of Ω_j at the corner covered by W_{jk}.

To bound $I_{jk}(n)$ we only look at the case when W_{jk} covers a corner of $\partial\Omega_j$ as the straight edge case is a special case of the corner case with $c = 0$.

In order to simplify the following we define the following four sets of points:

$$\begin{aligned} \mathcal{A}_n &:= \{\mathbf{a} \in \mathbb{Z}^2 : |a_1| + |a_2| = n\}, \\ \mathcal{B}_n &:= \{\mathbf{b} = \eta\mathbf{a} : \mathbf{a} \in \mathcal{A}_n \text{ and } \eta = \tfrac{n}{\sqrt{2}}|\mathbf{a}|^{-1}\}, \\ \mathcal{C}_n &:= \{\mathbf{c} = S^{-1}(\mathbf{b}) : \mathbf{b} \in \mathcal{B}_n\}, \\ \mathcal{D}_n &:= \{\mathbf{d} = \eta\mathbf{c} : \mathbf{c} \in \mathcal{C}_n \text{ and } \eta \text{ is s.t. } |d_2| = d \text{ or } |d_1 + cd_2| = d\sqrt{1+c^2}\}, \end{aligned} \tag{17}$$

where $d = \frac{n}{\sqrt{2}}\min(1 + (\sqrt{1+c^2} \pm c)^2)^{-1/2}$. Note that the vectors in $\mathcal{A}_n$ lie on a rotated square with sides of length $\sqrt{2}n$ centred at the origin; the vectors in $\mathcal{B}_n$ lie on a circle with radius $\frac{n}{\sqrt{2}}$ centred at the origin; the vectors in $\mathcal{C}_n$ also lie on a circle with radius $\frac{n}{\sqrt{2}}$ centred at the origin; and d has been calculated so that the points in $\mathcal{D}_n$ lie on the largest possible rhombus inside a circle of radius $\frac{n}{\sqrt{2}}$ centred at the origin where the sides of the rhombus are perpendicular to either $(0,1)$ or $(1,c)$. Also note that d is the closest distance that a point in $\mathcal{D}_n$ can be to the origin. Define α to be the smallest interior angle of the rhombus. It is possible to define angle preserving bijections between each of these sets. For example, each $\mathbf{b} \in \mathcal{B}_n$ is a scaled vector in $\mathcal{A}_n$, each $\mathbf{c} \in \mathcal{C}_n$ is a rotation of a vector in $\mathcal{B}_n$, and each $\mathbf{d} \in \mathcal{D}_n$ is a scaled vector in $\mathcal{D}_n$. Just as the distance between neighbouring points in $\mathcal{A}_n$ is $\frac{1}{\sqrt{2}}$ we also can bound (from above and below) the distance between neighbouring points in $\mathcal{D}_n$. Let a denote the lower bound. We are now in a position to bound $I_{jk}(n)$. Using the definition of $I_{jk}(n)$ together with (16), where $\mathbf{h} = S^{-1}\mathbf{g}$ as before, and the fact that $\eta \leq 1$ in the definition of $\mathcal{B}_n$ and $\mathcal{D}_n$, we have

$$\begin{aligned} I_{jk}(n) &\lesssim \sum_{\mathbf{g}\in\mathcal{A}_n} \frac{1}{|h_1+ch_2|_\star^2 |h_2|_\star^2} \leq \sum_{\mathbf{h}\in\mathcal{D}_n} \frac{1}{|h_1+ch_2|_\star^2 |h_2|_\star^2} \\ &= \frac{2}{d^2(1+c^2)} \sum_{\substack{\mathbf{h}\in\mathcal{D}_n: \\ |h_1+ch_2|=d\sqrt{1+c^2}}} \frac{1}{|h_2|_\star^2} \lesssim n^{-2}\left(4 + \frac{4}{a^2|\sin\alpha|^2} \sum_{m=1}^{\lceil n/2\rceil} \frac{1}{m^2}\right) \lesssim n^{-2}, \qquad (18) \end{aligned}$$

where we have used symmetry in going from the first to the second line. Therefore $F_n(V) \lesssim n^{-1}$ and $V \in \mathcal{Y}_p$. □

Proposition 7. *Let V be a periodic extension of*

$$V|_\Omega(\mathbf{x}) = \begin{cases} a, & |\mathbf{x}| \leq r < \frac{1}{2}\,, \\ 0, & \textit{otherwise}, \end{cases}$$

to $\mathbb{R}^2$, where a and r are two constants, i.e. the special case of Definition 3 with $J = 1$, $\Omega_1 = B_r(\mathbf{0})$ (ball with radius r), $V_0 \equiv 0$ and $V_1|_{\Omega_1} \equiv a$. Then $V \in \mathcal{Y}_p \subset \mathcal{X}_p$.

Proof. In this case, we can explicitly derive the Fourier coefficients of V, i.e.

$$[V]_\mathbf{g} = \begin{cases} a\pi r^2\,, & \mathbf{g} = \mathbf{0}, \\ \frac{ar}{2\pi|\mathbf{g}|} J_1(\pi|\mathbf{g}|r), & \mathbf{g} \neq \mathbf{0}, \end{cases}$$

where J_1 is the Bessel function of 1st order. Using the fact that there exists a constant b such that $J_1(x) \leq bx^{-1/2}$ for $x \geq 1$ it follows that $[V]_\mathbf{g} \lesssim |\mathbf{g}|^{-3/2}$ and so $F_n(V) \lesssim n^{-1}$. Hence $V \in \mathcal{Y}_p$. □

4 Error Analysis for the Planewave Expansion Method

Throughout this section we will assume that the potential $V(\mathbf{x})$ in (2) is in $\mathcal{X}_p$. We will not require the stronger type of regularity of $\mathcal{Y}_p$ just yet.

4.1 Solution Operator and Regularity of Eigenfunctions

It is useful for the analysis to define the solution operator $\mathrm{T} : L_p^2 \to H_p^1$ corresponding to (8) such that for every $f \in L_p^2$, $\mathrm{T}f \in H_p^1$ is defined by

$$a(\mathrm{T}f, v) = b(f, v), \qquad \text{for all } v \in H_p^1. \qquad (19)$$

Note that T is well-defined and bounded by the Riesz Representation Theorem (since $a(\cdot,\cdot)$ is bounded, Hermitian and coercive). It is self-adjoint with respect to $a(\cdot,\cdot)$, and it follows from

$\mathrm{T} : L_p^2 \to H_p^1$ bounded and $H_p^1 \subset\subset L_p^2$ that $\mathrm{T} : H_p^1 \to H_p^1$ is compact. ¿From the definition of T it follows that (λ, u) is an eigenpair of (8) if and only if $(\frac{1}{\lambda}, u)$ is an eigenpair of T. Using well-known spectral theory results (see for example [13]) we conclude from the fact that T is bounded, compact and self-adjoint, that (8) has real eigenvalues

$$0 < \lambda_1 \leq \lambda_2 \leq \cdots \nearrow +\infty$$

counted up to multiplicity with corresponding eigenfunctions

$$u_1,\ u_2, \ldots$$

that can be chosen so that they are orthogonal to each other with respect to $a(\cdot, \cdot)$ and complete in L_p^2.

T is a smoothing operator and we have the following result.

Lemma 8. *Let $V \in \mathcal{X}_p$ and $u \in H_p^1$. Then*

$$\|\mathrm{T}u\|_{H_p^{5/2-\varepsilon}} \lesssim \|u\|_{H_p^1}\,, \qquad \text{for any } \varepsilon > 0. \tag{20}$$

Proof. Let $V \in \mathcal{X}_p$ and $u \in H_p^1$. From the definition of T we know that $w = \mathrm{T}u$ is a weak solution to an elliptic boundary value problem

$$L_0 w = f\,, \qquad \text{on } \Omega,$$

with $L_0 := -(\nabla + i\boldsymbol{\xi})^2 + K$ and $f := u - V\mathrm{T}u\,,$ subject to periodic boundary conditions. Notice that L_0 has constant coefficients, and hence

$$\|w\|_{H_p^{s+2}} \lesssim \|f\|_{H_p^s}\,, \qquad \text{for any } s \geq 0. \tag{21}$$

This is an adapted version, for periodic boundary conditions, of a result in Lions and Magenes [12]. See [16, Thm. 3.77] for a proof.

Now, since $H_p^0 = L_p^2$ and $V \in \mathcal{X}_p \subset L^\infty(\mathbb{R}^2)$, and since $\mathrm{T} : H_p^1 \to H_p^1$ is bounded, applying (21) we have

$$\|\mathrm{T}u\|_{H_p^2} = \|w\|_{H_p^2} \lesssim \|u\|_{H_p^0} + \|V\|_{L^\infty} \|\mathrm{T}u\|_{H_p^0} \lesssim \|u\|_{H_p^1}\,. \tag{22}$$

Now let $0 \leq s < \frac{1}{2}$. Since $V \in \mathcal{X}_p \subset H_p^s$, we can use (21) and (22) together with Part 4 of Theorem 1 to get $\|\mathrm{T}u\|_{H_p^{s+2}} \lesssim \|u\|_{H_p^s} + \|V\|_{H_p^s} \|\mathrm{T}u\|_{H_p^2} \lesssim \|u\|_{H_p^1}$. □

The following corollary is a trivial consequence of Lemma 8.

Corollary 9. *If u is an eigenfunction of* (8) *with $V \in \mathcal{X}_p$, then*

$$\|u\|_{H_p^{5/2-\varepsilon}} \lesssim \|u\|_{H_p^1}\,, \qquad \text{for any } \varepsilon > 0. \tag{23}$$

4.2 Application of Abstract Theory for the Galerkin Method

Similarly to T we can also define the solution operator $\mathrm{T}_G : L_p^2 \to \mathcal{S}_G$ corresponding to (9). For $G \in \mathbb{N}$ and $f \in L_p^2$ define $\mathrm{T}_G : L_p^2 \to \mathcal{S}_G$ by

$$a(\mathrm{T}_G f, v_G) = b(f, v_G), \qquad \text{for all } v_G \in \mathcal{S}_G. \tag{24}$$

T_G is bounded and self-adjoint with respect to $a(\cdot,\cdot)$. Moreover, since $\mathrm{T}_G = Q_G\mathrm{T}$ with the (bounded) projection Q_G defined by $a(Q_G u - u, v) = 0$, for all $u \in H^1_p$ and $v \in \mathcal{S}_G$, and since $\mathrm{T} : H^1_p \to H^1_p$ is compact, it follows that $\mathrm{T}_G : H^1_p \to H^1_p$ is also compact. Again, λ_G is an eigenvalue of (9) if and only if λ_G^{-1} is an eigenvalue of T_G.

For $s \in \mathbb{R}$ and $G \in \mathbb{N}$ we define an orthogonal projection from H^s_p onto $\mathcal{S}_G$ such that for all $u \in H^s_p$

$$P_G u(\mathbf{x}) := \sum\nolimits_{\mathbf{g}\in\mathbb{Z}^2_G} [u]_{\mathbf{g}} \mathrm{e}^{i2\pi\mathbf{g}\cdot\mathbf{x}} \qquad \text{for all } \mathbf{x} \in \mathbb{R}^2. \tag{25}$$

Lemma 10. *For $s, t \in \mathbb{R}$ with $s \le t$, and $G \in \mathbb{N}$, if $u \in H^t_p$ then*

$$\|u - P_G u\|_{H^s_p} \le G^{s-t} \|u\|_{H^t_p} . \tag{26}$$

Proof. (Adapted from the 1D version in [19].) For $s \le t \in \mathbb{R}$, $u \in H^t_p$ and $G \in \mathbb{N}$,

$$\|u - P_G u\|^2_{H^s_p} = \sum_{|\mathbf{n}|>G} |\mathbf{n}|^{2s} |[u]_{\mathbf{n}}|^2 \le G^{2s-2t} \sum_{|\mathbf{n}|>G} ||\mathbf{n}|^{2t} |[u]_{\mathbf{n}}|^2 \le G^{2s-2t} \|u\|^2_{H^t_p} .$$

□

Corollary 11. *Let $V \in \mathcal{X}_p$. For $u \in H^1_p$ and $\varepsilon > 0$,*

$$\inf_{\chi\in\mathcal{S}_G} \|\mathrm{T}u - \chi\|_{H^1_p} \lesssim G^{-3/2+\varepsilon} \|u\|_{H^1_p} . \tag{27}$$

Moreover, if u is an eigenfunction of (8) and $\varepsilon > 0$, then

$$\inf_{\chi\in\mathcal{S}_G} \|u - \chi\|_{H^1_p} \lesssim G^{-3/2+\varepsilon} \|u\|_{H^1_p} . \tag{28}$$

Proof. Let $\varepsilon > 0$ and $\chi := P_G\mathrm{T}u$. Then it follows from Lemmas 8 and 10 that

$$\inf_{\chi\in\mathcal{S}_G} \|\mathrm{T}u - \chi\|_{H^1_p} \le \|\mathrm{T}u - P_G\mathrm{T}u\|_{H^1_p} \le G^{-3/2+\varepsilon} \|\mathrm{T}u\|_{H^{5/2-\varepsilon}_p} \lesssim G^{-3/2+\varepsilon} \|u\|_{H^1_p} .$$

Inequality (28) is proved analogously using Corollary 9 instead of Lemma 8. □

To use the abstract theory for Galerkin approximations of variational eigenproblems (e.g. in Babuška and Osborn [1]), we need to first prove the following lemma.

Lemma 12. *Let $V \in \mathcal{X}_p$. Then*

$$\|\mathrm{T} - \mathrm{T}_G\|_{H^1_p} \lesssim G^{-3/2+\varepsilon}, \qquad \textit{for any } \varepsilon > 0.$$

Proof. The proof of this result uses Cea's Lemma (see [4, Thm. 2.4.1]) and (27),

$$\|\mathrm{T} - \mathrm{T}_G\|_{H^1_p} = \sup_{u\in H^1_p} \frac{\|\mathrm{T}u - \mathrm{T}_G u\|_{H^1_p}}{\|u\|_{H^1_p}} \lesssim \sup_{u\in H^1_p} \inf_{\chi\in\mathcal{S}_G} \frac{\|\mathrm{T}u - \chi\|_{H^1_p}}{\|u\|_{H^1_p}} \lesssim G^{-3/2+\varepsilon} .$$

□

We also need to define the *gap between two subspaces of a Hilbert space $\mathcal{H}$ with norm $\|\cdot\|_{\mathcal{H}}$*:

$$\delta_{\mathcal{H}}(X, Y) := \sup_{x\in X, \|x\|_{\mathcal{H}}=1} \mathrm{dist}(x, Y) = \sup_{y\in Y, \|y\|_{\mathcal{H}}=1} \mathrm{dist}(y, X) .$$

It can be shown (cf. [16, Appendix] for details) that $\delta_{\mathcal{H}}(\cdot,\cdot)$ obeys a triangle inequality.

We are now ready to state the main theorem for this section.

Theorem 13. *Let $V \in \mathcal{X}_p$ and let λ be an eigenvalue of* (8) *with multiplicity m and corresponding eigenspace M. Then for G sufficiently large and $\varepsilon > 0$ arbitrarily small, there exist m eigenvalues $\lambda_1, \ldots, \lambda_m$ of* (9) *(counted according to their multiplicity) with $\lambda_j = \lambda_j(G)$ and with corresponding eigenspaces $M_1(\lambda_1), \ldots, M_m(\lambda_m) \subset \mathcal{S}_G$ such that*

$$\begin{aligned} \delta_{H^1_p}(M, \mathcal{M}_G) &\lesssim G^{-3/2+\varepsilon}\,, \qquad \textit{where} \qquad \mathcal{M}_G := \bigoplus_{j=1}^{m} M_j(\lambda_j)\,, \quad \textit{and} \\ |\lambda - \lambda_j| &\lesssim G^{-3+2\varepsilon}\,, \qquad \textit{for } j = 1, \ldots, m. \end{aligned}$$

Proof. This result follows directly from [1, Theorems 7.1 & 7.3] applied to T and T$_G$. The details of this are given in [16]. □

This result shows that the planewave expansion method is essentially third order in the eigenvalues and of order $\frac{3}{2} - \varepsilon$ in the eigenfunctions. The numerical results in §6 confirm this. They even suggest that the result might be true with $\varepsilon = 0$. To prove this stronger bound, which we have claimed in the introduction, is more subtle and requires the stronger regularity assumption $V \in \mathcal{Y}_p$. We will come back to this at the end of §7 (cf. Corollary 22).

A simple extension of this result using a standard duality argument shows that the gap, measured in L^2_p, is $\mathcal{O}(G^{-5/2+\varepsilon})$, and so eigenfunctions essentially converge with order $\frac{5}{2}$ in L^2_p. Again we will come back to this at the end of §7.

5 Implementation and Optimal Preconditioning

In this section we consider how to efficiently solve the matrix eigenvalue problem (13). For the theoretical results in this section we will require that V has the stronger regularity of $\mathcal{Y}_p$.

In practice only the smallest eigenvalues of A are good approximations of eigenvalues of (8). However, in the applications we have in mind in photonic crystals they are the only physically relevant eigenvalues. For this reason we use a Krylov subspace iteration method to calculate only a few of the smallest eigenvalues of A. Moreover, experience tells us that typically, the largest eigenvalues of A^{-1} are more favourably spaced than the smallest eigenvalues of A. Hence, we use the Implicitly Restarted Arnoldi (IRA) method implemented in the ARPACK software package [11] applied to A^{-1} instead of A.

At each iteration of the IRA method we require the action of A^{-1} on a vector, or equivalently, we must solve a linear system. Since A is Hermitian and positive definite we can use the preconditioned conjugate gradient (PCG) method. This method has an advantage over direct solvers for A, because matrix-vector multiplications with A can be computed in $\mathcal{O}(N \log N)$ operations (where N is the dimension of A), while a factorisation of A would require $\mathcal{O}(N^3)$ operations. However, more importantly it turns out that we also have optimal preconditioners for A that guarantee that the number of PCG iterations is independent of N.

To compute the product $A\mathbf{v}$, for $\mathbf{v} \in \mathbb{R}^N$, efficiently we recall that $A = D + W$ where D is diagonal and the entries of W are Fourier coefficients of $V(\mathbf{x})$. Since D is diagonal, it is obvious that $D\mathbf{v}$ can be computed in $\mathcal{O}(N)$. To compute $W\mathbf{v}$ we notice that the entries of $W\mathbf{v}$ are discrete convolutions of the Fourier coefficients of $P_{2G}V$ with the entries of $\mathbf{v}$ and can thus be computed in $\mathcal{O}(N \log N)$ operations using the Fast Fourier Transform. Further details on how $A\mathbf{v}$ is computed can be found in [16].

As a preconditioner for A in the PCG method we first consider simply the diagonal of A, i.e. $P_d := \mathrm{diag}(A)$. Provided that $V \in \mathcal{Y}_p$ and K is sufficiently large, we can prove that this is an optimal preconditioner for (13) in the sense that the condition number of $P_d^{-1}A$ is bounded independently of N. Recall that for every symmetric positive definite matrix T the condition number $\kappa(T) := \lambda_{\max}(T)/\lambda_{\min}(T)$.

Theorem 14. *Let $V \in \mathcal{Y}_p$ and let γ be a constant such that $F_n(V) \leq \gamma n^{-1}$ for all $n \in \mathbb{N}$. For any $C > 1$ (arbitrary), if $K \geq \frac{C+1}{C-1} 2^{11/4}\gamma\sqrt{G} + |[V]_0|$, then*

$$\kappa(P_d^{-1}A) \leq C.$$

Proof. The proof of this result relies on Gershgorin's Circle Theorem, i.e. for any matrix T we have

$$\sigma(T) \subset \bigcup_{i=1}^{N} B(T_{ii}, r_i)\,,$$

where $B(T_{ii}, r_i)$ is an open ball centred at T_{ii} with radius $r_i := \sum_{j\neq i}^{N} |T_{ij}|$.

Let $\mathbf{g} \in \mathbb{Z}_G^2$ and $K \geq \frac{C+1}{C-1} 2^{11/4}\gamma\sqrt{G} + |[V]_0|$ be fixed. The definition of P_d implies that $(P_d^{-1}A)_{\mathbf{gg}} = 1$ and we can bound the radius $r_{\mathbf{g}}$ in the following way:

$$\begin{aligned}
r_{\mathbf{g}} &= \sum\nolimits_{\mathbf{g}\neq\mathbf{g}'\in\mathbb{Z}_G^2} |(P_d^{-1}A)_{\mathbf{gg}'}| \leq \tfrac{1}{|\boldsymbol{\xi}+2\pi\mathbf{g}|^2+K-|[V]_0|} \sum\nolimits_{\mathbf{g}\neq\mathbf{g}'\in\mathbb{Z}_G^2} |[V]_{\mathbf{g}-\mathbf{g}'}| \\
&\leq \tfrac{1}{K-|[V]_0|} \sum_{0<|g_1|+|g_2|\leq 2\sqrt{2}G} |[V]_{\mathbf{g}}| = \tfrac{1}{K-|[V]_0|} \sum_{n=1}^{\lfloor 2\sqrt{2}G\rfloor} \sum_{|g_1|+|g_2|=n} |[V]_{\mathbf{g}}| \\
&\leq \tfrac{1}{K-|[V]_0|} \sum_{n=1}^{\lfloor 2\sqrt{2}G\rfloor} (4n)^{1/2} F_n(V) \leq \tfrac{2\gamma}{K-|[V]_0|} \sum_{n=1}^{\lfloor 2\sqrt{2}G\rfloor} n^{-1/2} \\
&\leq \tfrac{2\gamma}{K-|[V]_0|} \left(1 + \int_1^{2\sqrt{2}G} x^{-1/2} dx\right) \leq \frac{2^{11/4}\gamma\sqrt{G}}{K-|[V]_0|} \leq \frac{C-1}{C+1}\,.
\end{aligned}$$

Applying Gershgorin's Circle Theorem we get

$$\sigma(P_d^{-1}A) \subset \left[1 - \tfrac{C-1}{C+1}, 1 + \tfrac{C-1}{C+1}\right]$$

and thus $\kappa(P_d^{-1}A) = \frac{\lambda_{\max}}{\lambda_{\min}} \leq C$. □

The number of PCG iterations is $\mathcal{O}(\kappa(P_d^{-1}A)^{1/2})$, and thus we can reduce it arbitrarily by increasing the value of K, and it does not grow when we increase the number N of planewaves. The numerical results in Table 1 confirm this. However, there is a certain trade-off. The convergence of IRA depends on the relative gap between the eigenvalues of A^{-1} which decreases when K gets larger leading to a larger number of IRA iterations as we can also see in Table 1. Moreover, the number of IRA iterations shows a dependence on N for larger K.

Although P_d is asymptotically optimal, in practice even better performance can be achieved by choosing the preconditioner to be

$$P_b = \begin{bmatrix} B_1 & 0 \\ 0 & B_2 \end{bmatrix}\,, \tag{29}$$

where B_1 is a $N_b \times N_b$ dense matrix with entries

$$(B_1)_{\mathbf{gg}'} = A_{\mathbf{gg}'} \qquad \mathbf{g}, \mathbf{g}' \in \mathbb{Z}_{G_b}^2$$

Table 1: Comparing different preconditioners for matrix A in (13) in the case of Problem 2 (see §6) with $K = \|V\|_{L^\infty} + \pi^2 + \frac{1}{2} \approx 172.4$ (left table, P_b with $N_b = 512$) and with $K = 5000$ (right table).

		IRA restarts			PCG iterations		
G	N	I	P_d	P_b	I	P_d	P_b
15	709	7	7		50	38	
31	3001	7	7		99	38	
63	12453	7	7	7	204	39	18
127	50617	7	7	7	410	39	18

IRA restarts	PCG iterations
P_d $(K = 5000)$	P_d $(K = 5000)$
22	8
41	8
65	8
96	8

with $N_b := \dim \mathbb{Z}^2_{G_b}$, and B_2 is a $(N - N_b) \times (N - N_b)$ diagonal matrix with entries

$$(B_2)_{\mathbf{gg}} = A_{\mathbf{gg}} \qquad \mathbf{g} \in \mathbb{Z}^2_{G_b},\ |\mathbf{g}| > G_b.$$

In practice we choose $N_b \leq 512$. Note that P_d is a special case of P_b with $G_b = 0$ $(N_b = 1)$. The application of P_b^{-1} requires $\mathcal{O}(N + N_b^2)$ operations, if the Cholesky factorisation of B_1 is pre-computed ($\mathcal{O}(N_b^3)$ operations).

Corollary 15. *Let $V \in \mathcal{Y}_p$. For any $C > 1$, there exists a K sufficiently large such that*

$$\kappa(P_b^{-1}A) \leq C\,.$$

Proof. This follows directly from the fact that

$$\kappa(P_b^{-1}A) = \kappa(P_b^{-1}P_d P_d^{-1}A) \leq \kappa(P_b^{-1}P_d)\kappa(P_d^{-1}A)$$

and

$$\kappa(P_b^{-1}P_d) = \kappa(P_d^{-1}P_b) = \kappa(\mathrm{diag}(B_1)^{-1}B_1)$$

by applying Theorem 14 to $P_d^{-1}A$ and to $\mathrm{diag}(B_1)^{-1}B_1$. □

This shows that the preconditioner P_b is also optimal. Moreover, it seems to be even more efficient as we can see from the numerical results in Table 1, more than halving the number of PCG iterations at little extra cost. In practice we therefore always choose P_b as the preconditioner with a suitably chosen N_b.

To conclude, we have shown that we can solve (13) for a few of the smallest eigenvalues in $\mathcal{O}(N \log N)$ operations using Krylov subspace iteration. Each iteration of the eigensolver requires the solve of a linear system using PCG where the preconditioner is chosen to be P_d or P_b. The memory requirements for storing A and for solving (13) are $\mathcal{O}(N)$.

6 Numerical Experiments

In this section we apply the planewave expansion method to two problems from photonic crystal applications to verify our theoretical results and check if our error bounds are sharp. We refer to the problems as Problem 1 and Problem 2. In both problems $V(\mathbf{x})$ is piecewise constant and takes the values $V_{air} = -10.4$ and $V_{glass} = -162.0$. The structure of V for these two model problems is given in Fig. 1. The period cells for Problems 1 and 2 are $\Omega = (-\frac{1}{2}, \frac{1}{2})^2$ and

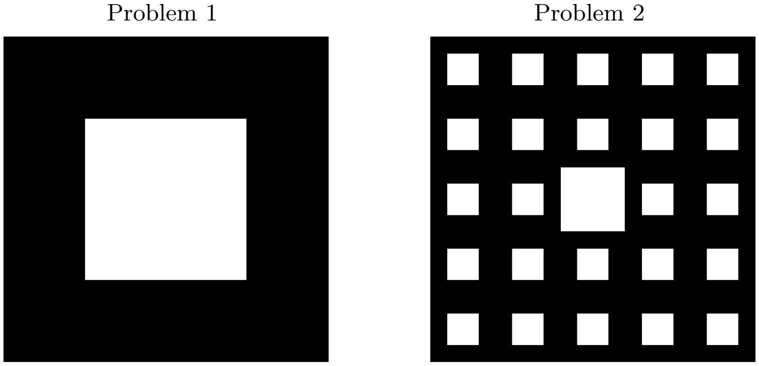

Figure 1: The period cell for $V(\mathbf{x})$ in Problems 1 and 2. $V = -162.0$ in the black regions and $V = -10.4$ in the white regions.

$\Omega = (-\frac{5}{2}, \frac{5}{2})^2$, respectively. In this way, Problem 2 represents a perturbed version of Problem 1, as it would arise by using the supercell method to approximate Problem 1 with a compact perturbation.

In Fig. 2 we plot the relative errors in the eigenvalues and the errors in the normalised eigenfunctions (measured in the H^1_p norm) for $\boldsymbol{\xi} = (0,0)$, (π,π) and $(\frac{\pi}{5}, \frac{\pi}{5})$ for some physically relevant eigenpairs. The reference solution for both problems is computed by solving (9) for a sufficiently large G, i.e. $G = 2^{10} - 1$. (Note that this corresponds to $N = 3.3 \times 10^6$ planewaves and requires 2D FFTs on arrays of size 4096×4096.) The plots confirm the results in Theorem 13. They suggest, in fact, that the eigenfunction errors decay with $\mathcal{O}(G^{-3/2})$ and the eigenvalue errors decay with $\mathcal{O}(G^{-3})$, which would correspond to choosing $\varepsilon = 0$ in Theorem 13. We will come back to this below (cf. Corollary 22).

7 Modified Planewave Expansion Method - Smoothed Potentials

Note that in this section we need to assume the stronger regularity $V \in \mathcal{Y}_p$ for the potential.

A standard approach in the physics literature to "improve" the convergence rate of the planewave method for Schrödinger operators with discontinuous potentials (e.g. in photonic crystals) is to smooth the discontinuous potential [8, 14, 17, 18]. A typical approach (cf. [17, 18]) is to replace V with a smooth function $\widetilde{V}$ defined by

$$\widetilde{V}(\mathbf{x}) := (\mathcal{G} * V)(\mathbf{x}) = \int_{\mathbb{R}^2} \mathcal{G}(\mathbf{x} - \mathbf{y}) V(\mathbf{y}) d\mathbf{y} \ ,$$

where $\mathcal{G}(\mathbf{x})$ is the normalised Gaussian

$$\mathcal{G}(\mathbf{x}) := \tfrac{1}{2\pi\Delta^2} \exp(-\tfrac{|\mathbf{x}|^2}{2\Delta^2}) \,,$$

for $0 < \Delta < 1$. The parameter Δ determines the width of the Gaussian function and in papers where this method is used it is referred to as FWHM (Full-Width-Half-Maximum). As $\Delta \to 0$

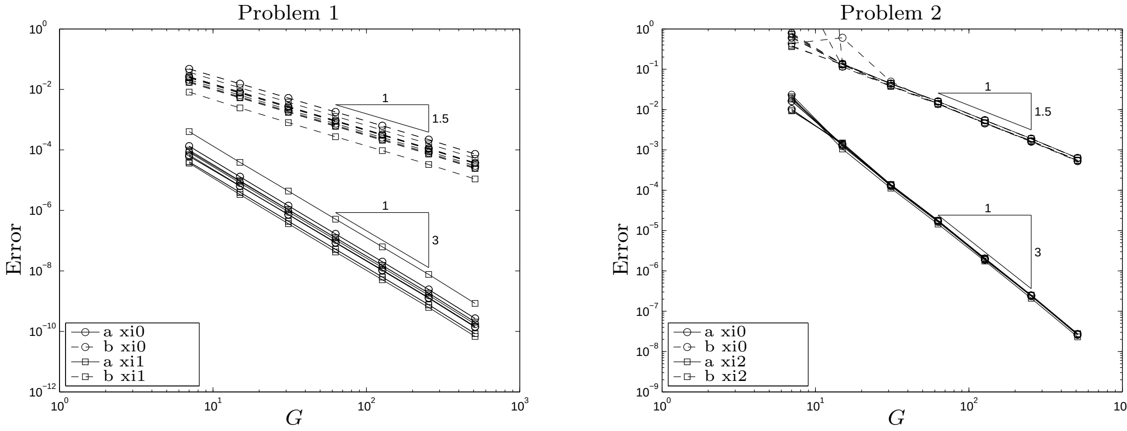

Figure 2: Eigenvalue error (a) and eigenfunction error in the H_p^1 norm (b) plotted against G for selected eigenpairs in Problem 1 (1st-5th eigenpairs) and in Problem 2 (23rd-27th eigenpairs), where $\boldsymbol{\xi} = (0, 0)$ (xi0), (π, π) (xi1), and $(\frac{\pi}{5}, \frac{\pi}{5})$ (xi2).

the Gaussian $\mathcal{G}(\mathbf{x})$ approaches the Dirac delta distribution and $\widetilde{V} \to V$ in the distributional sense.

Now we define the smoothed (variational) problem. For fixed $\boldsymbol{\xi} \in B$, find $\widetilde{\lambda} \in \mathbb{R}$ and $0 \neq u \in H_p^1$ such that

$$\tilde{a}(u, v) = \widetilde{\lambda} b(u, v), \qquad \text{for all } v \in H_p^1 \ , \tag{30}$$

where

$$\tilde{a}(u, v) := \int_\Omega (\nabla + i\xi)\, u \cdot \overline{(\nabla + i\xi)\, v} + \widetilde{V} u \overline{v} + K u \overline{v} dx$$

and where $b(\cdot, \cdot)$ is the same as in (8).

The bilinear form $\tilde{a}(\cdot, \cdot)$ has the same properties as $a(\cdot, \cdot)$, i.e. it is bounded, coercive and Hermitian on H_p^1. Therefore, it defines an inner product on H_p^1 with an induced norm that is equivalent to $\| \cdot \|_{H_p^1}$.

The idea is now to approximate the solution to (30) again with the planewave expansion method: For $G \in \mathbb{N}$, find $\widetilde{\lambda}_G \in \mathbb{C}$ and $0 \neq u_G \in \mathcal{S}_G$ such that

$$\tilde{a}(u_G, v_G) = \widetilde{\lambda}_G b(u_G, v_G), \qquad \text{for all } v_G \in \mathcal{S}_G. \tag{31}$$

Associated with both (30) and (31) are corresponding solution operators $\widetilde{\mathrm{T}}$ and $\widetilde{\mathrm{T}}_G$ which are defined as in (19) and (24). As before, $\widetilde{\mathrm{T}} : H_p^1 \to H_p^1$ and $\widetilde{\mathrm{T}}_G : H_p^1 \to \mathcal{S}_G \subset H_p^1$ are bounded, compact, positive and self-adjoint with respect to $\tilde{a}(\cdot, \cdot)$. However, in general they are *not* self-adjoint with respect to $a(\cdot, \cdot)$! Also as before, $(\widetilde{\lambda}, u)$ is an eigenpair of (30) if and only if $(\widetilde{\lambda}^{-1}, u)$ is an eigenpair of $\widetilde{\mathrm{T}}$. From the properties of $\widetilde{\mathrm{T}}$ we then deduce that (30) has a countable set of positive real eigenvalues with corresponding eigenfunctions that can be chosen so that they are orthogonal with respect to $\tilde{a}(\cdot, \cdot)$ and complete in L_p^2.

The implementation for this Galerkin method is the same as for the standard planewave expansion method without smoothing but the error analysis will have to be refined to estimate the dependence of the error on the amount of smoothing employed.

To justify the use of smoothing we would like to obtain error bounds that show that the error of (31) decreases at a faster rate (with respect to G) than the error of (9). We certainly expect to see that the error of (31) approximating the solution to (30) decreases at a faster rate with respect to G because the smooth problem will have eigenfunctions with more regularity but there is also an additional error due to the smoothing. We need to examine how big the additional error from smoothing is and whether or not it outweighs the benefits of smoothing.

In the following two lemmas we prove some properties of $\widetilde{V}$ and of its Fourier coefficients that will be useful for the error analysis and for the implementation of this variation of the planewave expansion method. The first lemma is a standard result that is used in e.g. [17, 18] (for a proof see [16]).

Lemma 16. *The Fourier coefficients of $\widetilde{V}(\mathbf{x})$ are related to those of $V(\mathbf{x})$ by*

$$[\widetilde{V}]_{\mathbf{g}} = \mathrm{e}^{-2\pi^2|\mathbf{g}|^2\Delta^2}[V]_{\mathbf{g}}\ , \qquad \forall \mathbf{g}\in\mathbb{Z}^2\,.$$

Lemma 17. *Let $V\in\mathcal{Y}_p$, $s\in\mathbb{R}$ and $\Delta\in(0,1)$. Then*

$$\|\widetilde{V}\|_{H_p^s} \lesssim \begin{cases} \Delta^{-s+1/2}, & \textit{for } s>\frac{1}{2},\\ (1+\log(\Delta^{-1}))^{1/2}, & \textit{for } s=\frac{1}{2},\\ 1, & \textit{for } s<\frac{1}{2}, \end{cases} \qquad \textit{and} \tag{32}$$

$$\|V-\widetilde{V}\|_{H_p^s} \lesssim \Delta^{-s+1/2}\ , \qquad \textit{for } -\tfrac{3}{2}<s<\tfrac{1}{2}\,. \tag{33}$$

Proof. Since $V\in\mathcal{Y}_p$ there exists a constant γ such that $F_n(V)\le\gamma n^{-1}$ for all $n\in\mathbb{N}$. Using this together with the definition of H_p^s and Lemma 16 we have

$$\begin{aligned}\|\widetilde{V}\|^2_{H_p^s} &= \textstyle\sum_{\mathbf{g}\in\mathbb{Z}^2}|\mathbf{g}|_\star^{2s}|[\widetilde{V}]_{\mathbf{g}}|^2 = \sum_{\mathbf{g}\in\mathbb{Z}^2}|\mathbf{g}|_\star^{2s}\mathrm{e}^{-4\pi^2\Delta^2|\mathbf{g}|^2}|[V]_{\mathbf{g}}|^2\\ &= |[V]_{\mathbf{0}}|^2+\sum_{n=1}^{\infty}\sum_{|g_1|+|g_2|=n}|\mathbf{g}|^{2s}\mathrm{e}^{-4\pi^2\Delta^2|\mathbf{g}|^2}|[V]_{\mathbf{g}}|^2\\ &\lesssim 1+\sum_{n=1}^{\infty}n^{2s}\mathrm{e}^{-4\pi^2\Delta^2n^2}F_n^2\lesssim 1+\sum_{n=1}^{\infty}n^{2s-2}\mathrm{e}^{-4\pi^2\Delta^2n^2}\ .\end{aligned} \tag{34}$$

Now, to prove (32) let us first consider $s>1/2$. Let $f(t)=t^{2s-2}\mathrm{e}^{-2\pi^2\Delta^2t^2}$ and let $t\ge0$. Then $f(t)$ has a single maximum at $t_0=\sqrt{2\max(s-1,0)}/2\pi\Delta$, and is monotonically increasing on the interval $[0,t_0]$ and monotonically decreasing on $[t_0,\infty)$. Moreover, $f(t_0)\lesssim\Delta^{2\max(1-s,0)}$. Therefore,

$$\begin{aligned}\sum_{n=1}^{\infty}n^{2s-2}\quad \mathrm{e}^{-2\pi^2\Delta^2n^2} &= \sum_{n=1}^{\lfloor t_0\rfloor-1}f(n)+f(\lfloor t_0\rfloor)+f(\lceil t_0\rceil)+\sum_{n=\lceil t_0\rceil+1}^{\infty}f(n)\le\\ &\le\int_1^{\lfloor t_0\rfloor}f(x)dx+2f(t_0)+\int_{\lceil t_0\rceil}^{\infty}f(x)dx\lesssim\Delta^{2-2s}+\int_0^{\infty}x^{2s-2}\mathrm{e}^{-2\pi^2\Delta^2x^2}dx\ .\end{aligned} \tag{35}$$

Putting (35) into (34) and substituting $y=x\,\Delta$ we get

$$\|\widetilde{V}\|^2_{H_p^s}\lesssim 1+\Delta^{2-2s}+\frac{1}{\Delta^{2s-1}}\int_0^{\infty}y^{2s-2}\mathrm{e}^{-2\pi^2y^2}dy\lesssim\Delta^{1-2s}\ .$$

If $s = \frac{1}{2}$, then $f(t) = t^{-1}\mathrm{e}^{2\pi^2\Delta^2 t^2} \leq 1$ is monotonically decreasing and following on from (34), again with $y = x\,\Delta$, we obtain similarly that

$$\begin{aligned}
\|\widetilde{V}\|^2_{H_p^{1/2}} &\lesssim 2 + \sum_{n=2}^{\infty} f(n) \leq 2 + \int_1^{\infty} f(x)dx = 2 + \int_{\Delta}^{\infty} y^{-1}\mathrm{e}^{-2\pi^2 y^2}dy \\
&\leq 2 + \int_{\Delta}^{1} y^{-1}dy + \int_1^{\infty} \mathrm{e}^{-2\pi^2 y}dy = 2 + \log(\Delta^{-1}) + \tfrac{1}{2\pi^2} \lesssim 1 + \log(\Delta^{-1})\,.
\end{aligned}$$

If $s < 1/2$, then Lemma 16 and the assumption that $V \in \mathcal{Y}_p \subset H_p^s$ imply that

$$\|\widetilde{V}\|^2_{H_p^s} = \sum\nolimits_{\mathbf{g}\in\mathbb{Z}^2} |\mathbf{g}|_\star^{2s}\mathrm{e}^{-4\pi^2\Delta^2|\mathbf{g}|^2}|[V]_{\mathbf{g}}|^2 \leq \|V\|^2_{H_p^s} \lesssim 1\,.$$

To prove (33) we proceed as in (34) to obtain

$$\|V - \widetilde{V}\|^2_{H_p^s} \lesssim \sum_{n=1}^{\infty} n^{2s-2}\left(1 - \mathrm{e}^{-2\pi^2\Delta^2 n^2}\right)^2\,. \tag{36}$$

To bound the right-hand-side of (36) we need to consider the function $f(t) = 1 - \mathrm{e}^{-t^2}$. By expanding e^{-t^2} in the usual way it can be shown that $f(t) = t^2 - (\frac{t^4}{2!} - \frac{t^6}{3!}) - \cdots \leq t^2$, for $|t| \leq 3$. Otherwise, $f(t) \leq 1$. Applying these bounds separately to the individual terms on the right-hand-side of (36) we get

$$\|V - \widetilde{V}\|^2_{H_p^s} \lesssim \sum_{n=1}^{\infty} n^{2s-2} f(\sqrt{2}\pi\Delta n)^2 \leq 4\pi^4\Delta^4 \sum_{n=1}^{\lfloor\frac{1}{\pi\Delta}\rfloor} n^{2s+2} + \sum_{n=\lceil\frac{1}{\pi\Delta}\rceil}^{\infty} n^{2s-2}\,, \tag{37}$$

which can be bounded in the same way as in (35) by using appropriate integrals. □

The key result that we need for the error analysis is the following lemma. It shows that the regularity of eigenfunctions of (30) is much greater than the regularity of the eigenfunctions of (8) (cf. Corollary 9). In fact, we see that the eigenfunctions of the smooth problem are infinitely differentiable. However, we also crucially manage to extract how the bounds in different Sobolev norms depend on Δ.

Lemma 18. *Let $V \in \mathcal{Y}_p$, $\Delta \in (0,1)$ and let u be an eigenfunction of* (30)*. Then $u \in C^\infty(\mathbb{R}^2)$ and*

$$\|u\|_{H_p^s} \lesssim \zeta(\Delta)\,\|u\|_{H_p^1}\,, \quad \textit{where} \;\; \zeta(\Delta) := \begin{cases} 1, & \textit{for } s < \frac{5}{2}, \\ (1 + \log(\Delta^{-1}))^{1/2}, & \textit{for } s = \frac{5}{2}, \\ \Delta^{-s+5/2}, & \textit{for } s > \frac{5}{2}. \end{cases}$$

Proof. Let $\widetilde{\lambda}$ be the eigenvalue of (30) that corresponds to the eigenfunction u. It follows from (30) that u is the weak solution of the elliptic boundary value problem

$$L_1 u = f\,, \qquad \text{on } \Omega\,,$$

with $L_1 := (\nabla + i\boldsymbol{\xi})^2 + \widetilde{V} + K$ and $f := \widetilde{\lambda}u$, subject to periodic boundary conditions. L_1 is elliptic with periodic C^∞–coefficients, and so we can use standard regularity results for elliptic boundary value problems and the fact that f is a multiple of u to “boot-strap” our way to $u \in H_p^s$ for any $s \in \mathbb{R}$. It then follows from Part 3 of Theorem 1 that $u \in C^\infty(\mathbb{R}^2)$.

To obtain bounds on $\|u\|_{H_p^s}$ we consider a different boundary value problem, i.e.

$$L_2 u = f\,, \qquad \text{on } \Omega\,, \tag{38}$$

with $L_2 := -(\nabla + i\boldsymbol{\xi})^2 + K$ and $f := \widetilde{\lambda} u - \widetilde{V} u + K u\,,$ subject to periodic boundary conditions. L_2 is (uniformly) elliptic and it has constant coefficients. The eigenfunction u is again the unique weak solution to (38).

First, let $s = 2$. Since $\widetilde{V}$ is continuous it follows from [12, pages 188-189] (adapted for periodic boundary conditions, cf. [16]) that

$$\|u\|_{H_p^2} \lesssim \|f\|_{L_p^2} \le |\widetilde{\lambda}|\|u\|_{L_p^2} + \|\widetilde{V}\|_{L^\infty}\|u\|_{L_p^2} + K\|u\|_{L_p^2} \lesssim \|u\|_{H_p^1}\,. \tag{39}$$

Now consider $2 \le s < \frac{5}{2}$. Using again the theory in [12] together with Part 4 of Theorem 1 we have

$$\|u\|_{H_p^s} \lesssim \|f\|_{H_p^{s-2}} \lesssim \|u\|_{H_p^{s-2}} + \|\widetilde{V}\|_{H_p^{s-2}}\|u\|_{H_p^2} \lesssim \|u\|_{H_p^1}\,, \tag{40}$$

where in the last step we have used (39) and Lemma 17. Note that this implies that (40) holds for all $s < \frac{5}{2}$ by the definition of $\|\cdot\|_{H_p^s}$.

Now consider $\frac{5}{2} \le s < \frac{9}{2}$. As above, using Lemma 17 we have

$$\|u\|_{H_p^s} \lesssim \begin{cases} (1+\log(\Delta^{-1}))^{1/2}\|u\|_{H_p^1}\,, & s = \frac{5}{2}\,, \\ \Delta^{-s+5/2}\|u\|_{H_p^1}\,, & \frac{5}{2} < s < \frac{9}{2}\,. \end{cases} \tag{41}$$

We now use induction to prove that $\|u\|_{H_p^s} \lesssim \Delta^{-s+5/2}\|u\|_{H_p^1}$ for $s \in \mathbb{N}$, $s \ge 4$. This is not trivial, if we want to obtain a sharp result. We have already proved the case $s = 4$ in (41). Our inductive hypothesis is to assume for some $k \in \mathbb{N}$, $k \ge 4$, that

$$\|u\|_{H_p^n} \lesssim \Delta^{-n+5/2}\|u\|_{H_p^1}\,, \qquad \text{for all } \; n \in \mathbb{N} \;\text{ s.t. }\; 3 \le n \le k\,. \tag{42}$$

It follows from (42) and the theory in [12] that

$$\|u\|_{H_p^{k+1}} \lesssim \|f\|_{H_p^{k-1}} \lesssim \|u\|_{H_p^{k-1}} + \|\widetilde{V}u\|_{H_p^{k-1}} \lesssim \Delta^{-k+7/2}\|u\|_{H_p^1} + \|\widetilde{V}u\|_{H_p^{k-1}}\,.$$

The key is now to bound $\|\widetilde{V}u\|_{H_p^{k-1}}$ in a clever way. We do not use Part 3 of Theorem 1 because the bound would not be sharp enough. Instead, Let α and β define non-negative multi-indices. We write $\beta \le \alpha$, to mean $\beta_i \le \alpha_i$ for all i and $|\alpha| := \sum_i \alpha_i$. Using Parts 3 and 5 of Theorem 1 together with Lemma 17 and (42) we get

$$\begin{aligned}
\|\widetilde{V}u\|^2_{H^{k-1}_p} &= \|\widetilde{V}u\|^2_{H^{k-1}(\Omega)} = \sum\nolimits_{|\alpha|\le k-1} \|D^\alpha(\widetilde{V}u)\|^2_{L^2(\Omega)} \\
&\lesssim \sum_{|\alpha|\le k-1} \sum_{j=0}^{|\alpha|} \sum_{|\beta|=j,\beta\le\alpha} \|(D^\beta \widetilde{V})(D^{\alpha-\beta}u)\|^2_{L^2(\Omega)} \\
&\le \sum_{|\alpha|\le k-1} \left(\|\widetilde{V}\|^2_{L^\infty} \|D^\alpha u\|^2_{L^2(\Omega)} + \sum_{j=1}^{|\alpha|} \sum_{|\beta|=j,\beta\le\alpha} \|D^\beta \widetilde{V}\|^2_{L^2(\Omega)} \|D^{\alpha-\beta}u\|^2_{L^\infty} \right) \\
&\lesssim \|\widetilde{V}\|^2_{L^\infty} \|u\|^2_{H^{k-1}(\Omega)} + \sum_{j=1}^{k-1} \|\widetilde{V}\|^2_{H^j(\Omega)} \|u\|^2_{H^{k-j+1}(\Omega)} \\
&\lesssim \|\widetilde{V}\|^2_{H^2_p} \|u\|^2_{H^{k-1}(\Omega)} + \sum_{j=1}^{k-2} \|\widetilde{V}\|^2_{H^j(\Omega)} \|u\|^2_{H^{k-j+1}(\Omega)} + \|\widetilde{V}\|^2_{H^{k-1}(\Omega)} \|u\|^2_{H^2(\Omega)} \\
&\lesssim \underbrace{\left(\Delta^{-2k+4} + \sum_{j=1}^{k-2} \Delta^{-2j+1} \Delta^{-2(k-j+1)+5} + \Delta^{-2k+3} \right)}_{\lesssim \Delta^{-2k+3}} \|u\|_{H^1_p} \, .
\end{aligned}$$

The result now follows by induction using Theorem 1, Part 2. □

The following corollary is a direct consequence of Lemma 18. Its proof is similar to the proof of Corollary 11.

Corollary 19. *Let $V \in \mathcal{Y}_p$ and $\Delta \in (0,1)$. Then, for any eigenfunction u of (30), we have*

$$\inf_{\chi\in\mathcal{S}_G} \|u-\chi\|_{H^1_p} \lesssim \Delta^{-s} G^{-3/2-s} \|u\|_{H^1_p} \, , \qquad \text{for any } s>0, \qquad \text{and} \tag{43}$$

$$\inf_{\chi\in\mathcal{S}_G} \|u-\chi\|_{H^1_p} \lesssim C^*(G,\Delta)\, G^{-3/2} \|u\|_{H^1_p} \, , \tag{44}$$

where $C^*(G,\Delta) := \min\{G^\varepsilon, (1+\log(\Delta^{-1}))^{1/2}\}$ for any $\varepsilon > 0$.

The first bound in Corollary 19 shows that (by taking s as large as we like) we can get polynomial decay of the approximation error of arbitrary degree with respect to G. However, the fast decay with respect to G is penalised when Δ is small. Nevertheless, the approximation error cannot become arbitrarily bad when Δ goes to zero due to the second bound.

With these improved regularity and approximation error results we can bound the error of the Galerkin approximation to the smooth problem (30) as in Theorem 13 by applying the abstract theory in [1]. This will form one part of the error bound for the smoothed planewave expansion method in Theorem 21 below.

The second part of the error bound contains the contribution from the smoothing error. We bound this error by comparing the two (infinite dimensional) problems (8) and (30), using again the theory in [1, Theorems 7.1 & 7.3] (in a non-standard way). To do this we must first show that the family of solution operators $\{\widetilde{\mathrm{T}}\}_{\Delta>0}$ (parametrised by Δ) satisfies $\widetilde{\mathrm{T}} \to \mathrm{T}$ as $\Delta \to 0$, and establish certain other bounds related to the convergence of $\widetilde{\mathrm{T}}$ to T. Recall that $\widetilde{\mathrm{T}}$ is not self-adjoint with respect to the inner product $a(\cdot,\cdot)$.

Lemma 20. *Let $V \in \mathcal{Y}_p$ and $\Delta \in (0,1)$. Then*

1. $\widetilde{\mathrm{T}} \to \mathrm{T}$ *as* $\Delta \to 0$ *and*
$$\|\mathrm{T} - \widetilde{\mathrm{T}}\|_{H^1_p} \lesssim \Delta^{3/2} \; ;$$
2. *the adjoint* $\widetilde{\mathrm{T}}^*$ *of* $\widetilde{\mathrm{T}}$ *with respect to* $a(\cdot,\cdot)$ *satisfies*
$$\|\mathrm{T} - \widetilde{\mathrm{T}}^*\|_{H^1_p} \lesssim \Delta^{3/2} \; ; \qquad \textit{and}$$
3. $$|a((\mathrm{T} - \widetilde{\mathrm{T}})u, v)| \lesssim \Delta^{3/2}\|u\|_{H^1_p}\|v\|_{H^1_p} \, , \qquad \textit{for any} \;\; u, v \in H^1_p.$$

Proof. Part 1. The proof for this result relies on an infinite dimensional version of Strang's First Lemma which we couldn't find in the literature (see [4] for a reference in the finite dimensional case).

Let $f \in H^1_p$. Then using Part 3 of Theorem 1 together with Strang's Lemma and choosing $v := Tf \in H^1_p$ we have

$$\begin{aligned} \|\mathrm{T}f - \widetilde{\mathrm{T}}f\|_{H^1_p} &\lesssim \inf_{v \in H^1_p} \left\{ \|\mathrm{T}f - v\|_{H^1_p} + \sup_{w \in H^1_p} \frac{|a(v,w) - \tilde{a}(v,w)|}{\|w\|_{H^1_p}} \right\} \\ &\leq \sup_{w \in H^1_p} \frac{|a(\mathrm{T}f,w) - \tilde{a}(\mathrm{T}f,w)|}{\|w\|_{H^1_p}} \leq \sup_{w \in H^1_p} \frac{\|\mathrm{T}f\|_{L^\infty} \int_\Omega |(V - \widetilde{V})\overline{w}| dx}{\|w\|_{H^1_p}} \\ &\lesssim \|\mathrm{T}f\|_{H^2_p} \|V - \widetilde{V}\|_{H^{-1}_p} \lesssim \Delta^{3/2}\|f\|_{H^1_p} \, , \end{aligned}$$

where in the last step we have used Lemmas 8 and 17.

Part 2. Let $f \in H^1_p$. Since $a(\cdot,\cdot)$ is bounded and coercive on H^1_p we have

$$\begin{aligned} \|(\mathrm{T} - \widetilde{\mathrm{T}}^*)f\|^2_{H^1_p} &\lesssim a((\mathrm{T} - \widetilde{\mathrm{T}}^*)f, (\mathrm{T} - \widetilde{\mathrm{T}}^*)f) = a((\mathrm{T} - \widetilde{\mathrm{T}})(\mathrm{T} - \widetilde{\mathrm{T}}^*)f, f) \\ &\leq \|\mathrm{T} - \widetilde{\mathrm{T}}\|_{H^1_p} \|(\mathrm{T} - \widetilde{\mathrm{T}}^*)f\|_{H^1_p} \|f\|_{H^1_p}. \end{aligned}$$

Dividing through by $\|(\mathrm{T} - \widetilde{\mathrm{T}}^*)f\|_{H^1_p}$ and applying Part 1 we obtain the result.

Part 3 follows directly from Part 1 using the fact that $a(\cdot,\cdot)$ is bounded. □

As above we can use the results in this lemma to apply the abstract theory in [1] to the operators $\widetilde{\mathrm{T}}$ and T to obtain bounds on the errors between the spectra of the two infinite dimensional problems (30) and (8). Putting these bounds together with the bounds on the Galerkin error for the smoothed problem (cf. Corollary 19) and applying the triangle inequality we obtain the main result of this section.

Theorem 21. *Let $V \in \mathcal{Y}_p$ and let λ be an eigenvalue of* (8) *with multiplicity m and corresponding eigenspace M. Then, for sufficiently large G and small $\Delta > 0$, there exist m eigenvalues $\widetilde{\lambda}_1, \ldots, \widetilde{\lambda}_m$ of* (31) *(counted according to multiplicity) with $\widetilde{\lambda}_j = \widetilde{\lambda}_j(G, \Delta)$ and with corresponding eigenspaces $\widetilde{M}_1(\widetilde{\lambda}_1), \ldots, \widetilde{M}_m(\widetilde{\lambda}_m) \subset \mathcal{S}_G$, such that for any $s > 0$, we have*

$$\begin{aligned} \delta_{H^1_p}(M, \widetilde{\mathcal{M}}_{G,\Delta}) &\lesssim \Delta^{3/2} + \Delta^{-s}\, G^{-3/2-s} \, , \;\; \textit{where} \;\; \widetilde{\mathcal{M}}_{G,\Delta} := \bigoplus_{j=1}^m \widetilde{M}_j(\widetilde{\lambda}_j) \, , \quad \textit{and} \\ |\lambda - \widetilde{\lambda}_j| &\lesssim \Delta^{3/2} + \Delta^{-2s}\, G^{-3-2s} \, , \quad \textit{for} \quad j = 1, \ldots, m \, . \end{aligned}$$

The second terms in these bounds can be replaced by $C^*(G,\Delta)G^{-3/2}$ and $C^*(G,\Delta)^2 G^{-3}$, respectively, with $C^*(G,\Delta)$ as defined in Corollary 19.

As expected eigenpairs of (30) converge to eigenpairs of (8) as $\Delta \to 0$ and the convergence with respect to G is potentially faster. The eigenvalue bound does not decrease at twice the rate of the eigenfunction error bound as $\Delta \to 0$ (as in Theorem 13). The reason lies in the fact that $\widetilde{T}$ is not self-adjoint with respect to $a(\cdot,\cdot)$ (cf. [1]). The numerical results in §8 confirm this, but we will see that our bound on the eigenvalue error is not completely sharp. Tracing the slackness back through our error analysis, we notice that Part 3 of Lemma 20 is not sharp.

To obtain the best amount of smoothing, we can now set $\Delta = \mathcal{O}(G^r)$ and choose $r \in \mathbb{R}$ to balance the two terms in the error bounds. However, we see that *at best* we can only achieve eigenfunction errors that are $\mathcal{O}(G^{-3/2})$ in the H_p^1-norm (choosing $r \leq -1$) and eigenvalue errors that are $\mathcal{O}(G^{-3} \log G)$ (choosing $r \leq -2$). Even though this appears to be a very slight improvement on the bounds for the basic planewave expansion method in Theorem 13, the numerical results in the next section will show that the modified planewave expansion method always performs worse. The reason for the better error bound lies in the fact that we have assumed the stronger regularity $V(\mathbf{x}) \in \mathcal{Y}_p$ here. In fact, it turns out that we can use the error analysis in this section also to improve the error bound for the eigenfunctions in the basic planewave expansion method (without smoothing, cf. Theorem 13).

Corollary 22. *Let $V \in \mathcal{Y}_p$ and let M and $\mathcal{M}_G$ be eigenspaces defined as in Theorem 13. Then, for sufficiently large G,*

$$\delta_{H_p^1}(M, \mathcal{M}_G) \lesssim G^{-3/2} \qquad \textit{and} \qquad \delta_{L_p^2}(M, \mathcal{M}_G) \lesssim G^{-5/2}.$$

Proof. First we realise that replacing V in (9) with the smooth function $P_{2G}V$ (cf. (25)) does not actually change the matrix eigenproblem (13). Then we proceed as shown for $\widetilde{V}$ in this section. In particular, we note that $[P_{2G}V]_{\mathbf{g}} = 0$, if $|\mathbf{g}| > 2G$ (cf. Lemma 16), which allows us to obtain the bounds

$$\|P_{2G}\|_{H_p^s} \lesssim \begin{cases} G^{s-1/2}\,, & \text{for } s > \frac{1}{2}, \\ (1+\log G)^{1/2}\,, & \text{for } s = \frac{1}{2}, \\ 1, & \text{for } s < \frac{1}{2}, \end{cases} \qquad \text{and} \qquad \|V - P_{2G}\|_{H_p^{-1}} \lesssim G^{-3/2}\,,$$

in the same way as for $\widetilde{V}$ in Lemma 17. The rest of the proof is identical.

The bound on $\delta_{L_p^2}(M, \mathcal{M}_G)$ follows by a standard duality argument for $\|T - T_G\|_{L_p^2}$ and the fact that both T and T_G are compact operators from L_p^2 to L_p^2 (cf. [1]). □

Unfortunately we were not able to derive an improved eigenvalue error bound. The problem lies again in the lack of a sharper bound for Part 3 of Lemma 20.

In conclusion we can say that smoothing $V(\mathbf{x})$ does not seem to have an advantage over the basic planewave expansion method where V is unmodified, and the numerical experiments in the next section support this conclusion.

We have only one qualifying remark to make. We have assumed throughout that we can calculate the entries of A in (13) exactly using an explicit formula for the Fourier coefficients of V. If this is not the case (as is commonly the case in applications), then smoothing may be of some benefit to reduce the additional error introduced by approximating the Fourier coefficients of V. We hope to shed some light on this issue in future work.

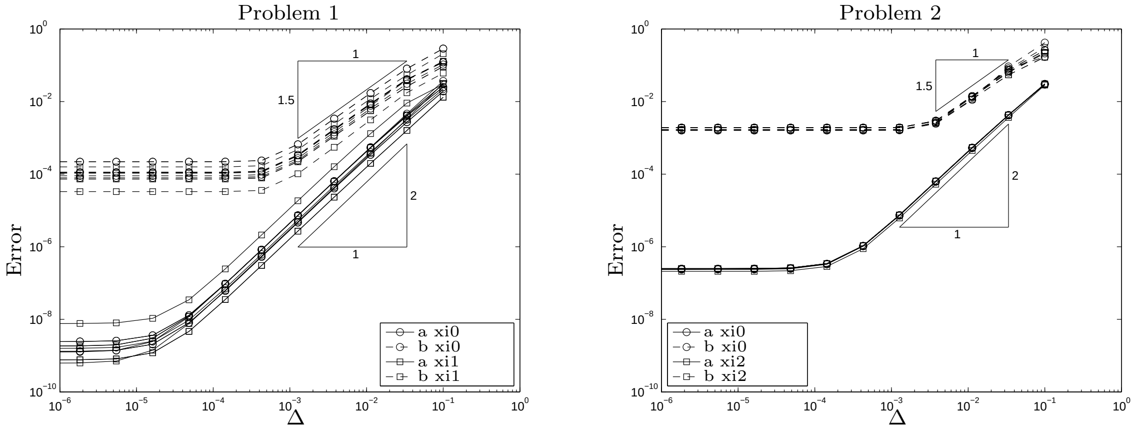

Figure 3: Eigenvalue error (a) and eigenfunction error in the H_p^1 norm (b) plotted against Δ for selected eigenpairs in Problem 1 (1st-5th eigenpairs) and in Problem 2 (23rd-27th eigenpairs), where $\boldsymbol{\xi} = \mathbf{0}$ (xi0), (π, π) (xi1), and $(\frac{\pi}{5}, \frac{\pi}{5})$ (xi2).

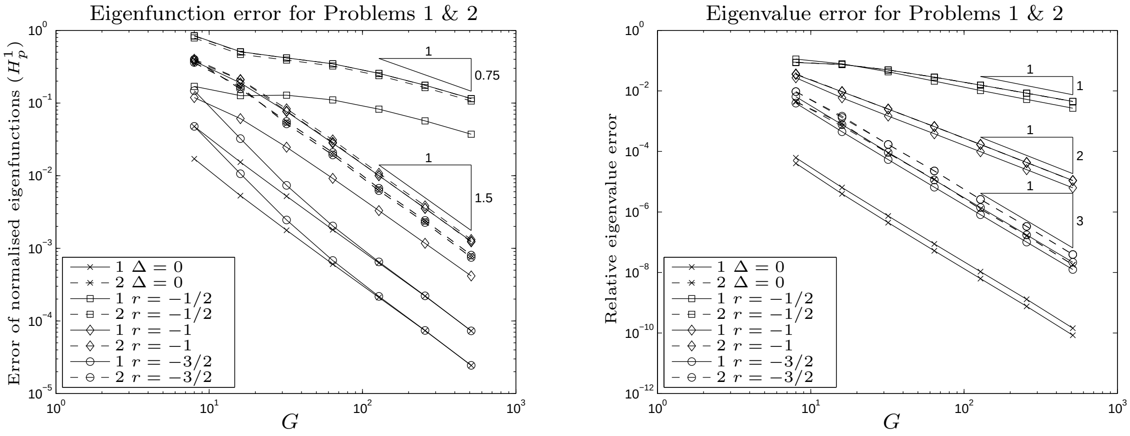

Figure 4: Error for the 1st eigenpair in Problems 1 and 2 plotted against G using the modified planewave expansion method with $\Delta = G^r$ for different $r \in \mathbb{R}$. (Lines for different $\boldsymbol{\xi}$ have been plotted but not labelled as they are indistinguishable.)

8 Numerical Experiments with Smoothing

We now perform some numerical experiments on Problems 1 and 2 from §6 to check whether the error bounds in Theorem 21 are sharp. In Fig. 3 we have plotted eigenfunction and eigenvalue errors against Δ for fixed (large) $G = 2^8 - 1$. The plot suggests that the first term of the eigenvalue error bound in Theorem 21 could be improved to Δ^2. Despite this, our general conclusion that smoothing has no advantage is still correct. Choosing $\Delta = \mathcal{O}(G^r)$ and balancing the error terms again, we find that the best possible eigenvalue error bound is still of order $\mathcal{O}(G^{-3} \log G)$.

Finally to confirm the claim numerically, in Fig. 4 we have set $\Delta = G^r$ for different choices of $r \in \mathbb{R}$ and plotted eigenvalue and eigenfunction errors against G. For comparison we include the case $\Delta = 0$, i.e. the basic planewave expansion method with V unmodified. We see that the error of the basic method is always smallest.

References

[1] I. Babuška & J. Osborn, *Eigenvalue Problems*, Handbook of Numerical Analysis, Vol. 2, Elsevier North-Holland, Amsterdam and London, 1991, pp. 641-787.

[2] T.A. Birks, D.M. Bird et. al., *Scaling laws and vector effects in bandgap-guiding fibres*, Optics Express, 12 (2004), pp. 69-74.

[3] Y. Cao, Z. Hou & Y. Liu, *Convergence problem of plane-wave expansion method for photonic crystals*, Physics Letters A, 327 (2004), pp. 247-253.

[4] P.G. Ciarlet, *The Finite Element Method for Elliptic Problems*, North-Holland, Amsterdam and New York, 1978.

[5] J. Elschner, *Singular Ordinary Differential Operators and Pseudodifferential Equations*, in Lecture notes in mathematics, Volume 1128, Springer, Berlin, 1985.

[6] P.D. Hislop & I.M. Sigal, *Introduction to spectral theory with applications to Schrödinger operators*, Springer, New York, 1996.

[7] J.D. Joannopoulos, S.G. Johnson, J.N. Winn & R.D. Meade, *Photonic Crystals Molding the Flow of Light 2nd edition*, Princeton University Press, Princeton, NJ, 2008.

[8] S.G. Johnson & J.D. Joannopoulos, *Block-iterative frequency-domain methods for Maxwell's equations in a planewave basis*, Optics Express, 8 (2000), pp. 173-190.

[9] J.C. Knight, *Photonic crystal fibres*, Nature, 424 (2003), pp. 847-851.

[10] P. Kuchment, *The mathematics of photonic crystals*, Chapter 7 in Mathematical Modelling in Optical Science, Frontiers in Applied Mathematics, SIAM, 22, (2001), pp. 207-272.

[11] R.B. Lehoucq, D.C. Sorensen & C. Yang, *ARPACK Users' Guide*, SIAM, 1998.

[12] J.L. Lions & E. Magenes, *Non-Homogeneous Boundary Value Problems and Applications Vol. 1*, Springer-Verlag, Berlin, 1972.

[13] W. McLean, *Strongly Elliptic Systems and Boundary Integral Equations*, Cambridge University Press, Cambridge, 2000.

[14] R.D. Meade, A.M. Rappe, K.D. Brommer, J.D. Joannopoulos & O.L. Alerhand, *Accurate theoretical analysis of photonic band-gap materials*, Physical Review B, 48 (1993), pp. 8434-8437.

[15] M.S. Min & D. Gottlieb, *On the convergence of the Fourier approximation for eigenvalues and eigenfunctions of discontinuous problems*, SIAM J. Num. Anal., 40 (2003), pp. 2254-2269.

[16] R.A. Norton, *Numerical Computation of Band Gaps in Photonic Crystal Fibres*, Ph.D. thesis, University of Bath, 2008.

[17] G.J. Pearce, T.D. Hedley & D.M. Bird, *Adaptive curvilinear coordinates in a plane-wave solution of Maxwell's equations in photonic crystals*, Physical Review B, 71 (2005), pp. 195108(10).

[18] G.J. Pearce, *Plane-wave methods for modelling photonic crystal fibre*, PhD Thesis, University of Bath, 2006.

[19] J. Saranen & G. Vainikko, *Periodic Integral and Pseudodifferential Equations with Numerical Approximation*, Springer, Berlin and London, 2002.

[20] S. Soussi, *Convergence of the supercell method for defect modes calculations in photonic crystals*, SIAM J. Numer. Anal., 43 (2005), pp. 1175-1201.

[21] H.S. Sözüer & J.W. Haus, Photonic bands: convergence problems with the planewave method, Physics Review B, 45 (1992), pp. 13962-13973.