Author for correspondence:
Anthony J. Mulholland
e-mail: anthony.mulholland@strath.ac.uk

A dual weighted residual method applied to complex periodic gratings

Natacha H. Lord and Anthony J. Mulholland

Department of Mathematics and Statistics, Strathclyde university, Livingstone Tower, 26 Richmond Street, Glasgow G1 1XH, UK

An extension of the dual weighted residual (DWR) method to the analysis of electromagnetic waves in a periodic diffraction grating is presented. Using the α,0-quasi-periodic transformation, an upper bound for the *a posteriori* error estimate is derived. This is then used to solve adaptively the associated Helmholtz problem. The goal is to achieve an acceptable accuracy in the computed diffraction efficiency while keeping the computational mesh relatively coarse. Numerical results are presented to illustrate the advantage of using DWR over the global *a posteriori* error estimate approach. The application of the method in biomimetic, to address the complex diffraction geometry of the Morpho butterfly wing is also discussed.

1. Introduction

This paper investigates the interaction of electromagnetic waves with periodic diffraction gratings. This type of grating has been used recently in crystalline silicon solar cells [1], gas sensors [2] and medical X-ray imaging [3,4]. They are also used in acousto-optic devices to characterize the composition of drugs, and in spectroscopic sensing for detecting trace gases [5]. Such periodic structures also occur naturally in the Morpho butterfly wing. The iridescent scales of this butterfly display a different optical response to the presence of different vapours [2]. The information about the nature and concentration of the vapours provided by the reflectance spectra of the scales has therefore been proposed as a design for a gas sensor.

Theoretically, electromagnetic wave diffraction is based on solving Maxwell's equations in the diffraction grating region [6,7]. If it is assumed that there is neither charge nor current, then the Maxwell equations reduce to the Helmholtz equation [7]. This boundary

value problem has been investigated in two dimensions, for a periodic grating, using a finite-element method [8]. This work also introduced the concept of the α_0, 0-quasi-periodic transformation which is used in this paper. In order to achieve a reliable and efficient numerical method, it is necessary to know how the system parameters affect the error between the exact and approximate solution. These errors can be classified into three different sources; data errors, model errors and computation errors [9]. As the finite-element method is used here, then the focus will be on the error arising from the discretization (data and modelling). For more details on the finite-element modelling in periodic structure, see [10–12].

An *a priori* error estimate has already been derived in Lord & Mulholland [8] to guarantee that the discrete solution converges to the exact solution provided that the interpolation error is kept small with respect to the wavenumber k. This study found an upper bound on the error between the exact and the approximate solutions in terms of the exact solution and the solution to the dual problem. Hence, this upper bound depends on some stability constants whose dependency on the system parameters is unknown. This limits the numerical approach to using a uniform mesh, which is computationally expensive. In this paper, an *a posteriori* error estimate that arises when we discretize the Helmholtz problem for a periodic grating is derived. In contrast to the *a priori* error estimate, it will provide a computable upper bound so that local refinement of the mesh (h-version) or raising the degree of the polynomial basis (p-version) can be performed. By doing so, this will reduce and control both the numerical error and the computational cost. This process is known as the adaptive computation method [9,13,14]. There are two different approaches to estimate the approximation error (denoted e_h). The first approach consists of looking for upper and lower bounds on e_h (global error estimate) and the second one is to use a quantitative estimate of some local feature of the exact solution called the quantity of interest (strain, displacement, etc.) [9], and to look for an estimate of the error in this quantity of interest (goal-oriented error estimate). In many applications, the interest lies in the error that arises in some specific, real valued physical quantity of interest Q that depends on the solution $U_{\alpha,0}$. The global error does not, however, provide useful bounds for this error in the quantity of interest. Also the sensitivity of the global error to local error sources is not properly represented when global stability constants are used [9,14]. This paper will employ a goal-oriented error estimate so that the error in the quantity of physical interest can be controlled and at the same time the efficiency of computing this quantity can be optimized. In particular, this paper focuses on the dual weighted residual (DWR) method [13]. Here, the dual problem is combined with the direct problem to derive an estimate of the error in the target quantity from each local residual on each of the mesh cells. In this way, the error is controlled locally in computing the target quantity. Given a grating profile, one of the main concerns is to know how much of the incident wave will be reflected and how much will be transmitted. This is achieved by computing the efficiency of the grating. Hence, rather than estimating the energy norm of the error in the solution to the Helmholtz problem, a particular linear functional of this error is estimated. This linear functional denoted by Q is chosen so that the error made by computing the efficiency of the grating is minimized.

There are two types of grating, the perfectly conducting and the dielectric transmitting grating [8]. There are also two fundamental polarizations, the transverse electric (TE) mode and the transverse magnetic (TM) mode [7,15,16]. Hence, there are four cases to investigate: Case 1A/B: perfectly conducting; and Case 2A/B: transmitting dielectric where A (B) denotes the TE (TM) wave. In §2, the mathematical statement of the problem is posed both in its continuous and discrete forms. In §3, the dual problem is introduced, which then facilitates the derivation of the *a posteriori* error estimate in §4. The final section provides some numerical results and applications.

2. Problem statement

In this section, the geometry of the problem, the associated function spaces, the form of the Helmholtz problem under the α_0,0-quasi-periodic transformation and the boundary conditions are detailed.
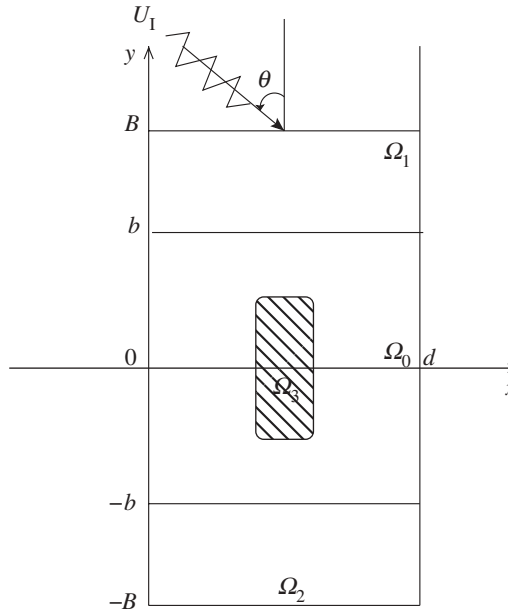
Figure 1. The truncated periodic grating domain Ω. Define the region above the scattering region Ω_1 to be $\{(x,y): 0 \leq x < d,\ b \leq y \leq B\}$ and the substrate Ω_2 to be $\{(x,y): 0 \leq x < d,\ -B \leq y \leq -b\}$.

(a) Boundary value problem

Using the $\alpha_0,0$-quasi-periodic transformation with the finite-element method, it has been shown in Lord & Mulholland [8] that the periodic diffraction grating problem can be solved numerically inside the truncated single vertical strip, $\Omega = [0,d] \times [-B,B]$ as shown in figure 1. Define $\Gamma_+ = \{(x,y): 0 \leq x \leq d, y = B\}$ and $\Gamma_- = \{(x,y): 0 \leq x \leq d, y = -B\}$. Also define the wavenumber k to be

$$k(x,y) = \begin{cases} k_1 \in \mathbb{R}, & \text{for } (x,y) \in \Omega_1, \\ k_0(x,y) \in \mathbb{C}, & \text{for } (x,y) \in \Omega_0, \\ k_2 \in \mathbb{C}, & \text{for } (x,y) \in \ \Omega_2, \end{cases} \tag{2.1}$$

where $\Omega_1 = [0,d] \times [b,B]$, $\Omega_0 = [0,d] \times [-b,b]$ and $\Omega_2 = [0,d] \times [-B,-b]$ with d the grating period, b and B positive real numbers. The incident wave is given by $U_I = \mathrm{e}^{i\alpha_0 x - i\beta_1^0 y}$, where $\alpha_0 = k_1 \sin\theta$, $\beta_1^0 = k_1 \cos\theta$ and θ is the angle of incidence of the wave as shown in figure 1. The scattered and diffracted waves are composed of bounded outgoing waves by demanding that $\Re(k_j) > 0$ and $\Im(k_j) \geq 0$, where $\Re(k_j)$ ($\Im(k_j)$) denotes the real (imaginary) part of k_j for $j \in \{0,1,2\}$.

The notation below will allow us to define the various function spaces arising on the boundaries of the domains in figure 1 which are

$$H_\#^s([0,d]) = \{g \in H^s([0,d]): g(d) = g(0)\},$$

and the function spaces used inside Ω which are

$$H_\#^s(\Omega) = \{f \in H^s(\Omega): f(d,y) = f(0,y), \quad \forall\, y \in [-B,B]\},$$

with $s \in \mathbb{R}$, $H^s(\Omega)$ and $H^s([0,d])$ are Sobolev spaces [17,18].

Definition 2.1. Let $f \in H_\#^{1/2}(\Gamma_\pm)$. We define the Dirichlet to Neumann (DtN) maps, $T_\pm^{\alpha,0}$, where $T_\pm^{\alpha,0} f \in H_\#^{-1/2}(\Gamma_\pm)$, and

$$
\left.
\begin{aligned}
T_\pm^{\alpha,0} f(x) &= \sum_{n \in \mathbb{Z}} \mathrm{i}\beta_j^n f^{(n)}(\pm B)\, \mathrm{e}^{\mathrm{i}(2\pi n/d)x}, \\
f^{(n)}(\pm B) &= \frac{1}{d} \int_0^d f(x, \pm B)\, \mathrm{e}^{-\mathrm{i}(2\pi n/d)x}\, \mathrm{d}x
\end{aligned}
\right\}
\tag{2.2}
$$

where

for $n \in \mathbb{Z}$, $\alpha_n = \alpha_0 + 2\pi n/d$ and $\beta_j^n = \mathrm{e}^{\mathrm{i}z_n/2}(|k_j^2 - \alpha_n^2|)^{1/2}$, such that $z_n = \arg(k_j^2 - \alpha_n^2)$, for $j \in \{1,2\}$ with $k_j^2 \neq \alpha_n^2$. The case where $k_j^2 = \alpha_n^2$ corresponds to the resonance phenomenon and is not considered here [7]. It can be shown that $T_\pm^{\alpha,0}$ is a continuous function from $H_\#^{1/2}(\Gamma_\pm) \to H_\#^{-1/2}(\Gamma_\pm)$ [11, p. 29].

The solution to the scattering problem is α_0-quasi-periodic, that is it can be written as the product of a periodic function with $\mathrm{e}^{\mathrm{i}\alpha_0 x}$. By denoting this periodic function by $U_{\alpha,0}$, the α_0, 0-quasi-periodic transformation can be used [8]. Hence, the scattering problem when an incident wave interacts with a periodic grating can be reformulated as $AU_{\alpha,0} = f$, where

$$
A = \Delta + (k^2 - \alpha_0^2) + 2\mathrm{i}\alpha_0 \partial_x
\tag{2.3}
$$

for Cases 1A, 1B and 2A and

$$
A = \nabla_{\alpha,0} \cdot \left(\frac{1}{k^2} \nabla_{\alpha,0} \right) + I_d
\tag{2.4}
$$

for Case 2B. $\nabla_{\alpha,0}$ is given by $\nabla + \mathrm{i}(\alpha_0, 0)$, I_d is the identity operator, f is some given data, $A: H_\#^1(F) \to H_\#^1(F)$ is a linear operator on $H_\#^1(F)$, and

$$
F = \begin{cases} \Omega \setminus \mathrm{int}\, \Omega_3 & \text{for Case 1} \\ \Omega & \text{for Case 2.} \end{cases}
\tag{2.5}
$$

For all cases, $U_{\alpha,0}$ also satisfies the following boundary conditions:

$$
\left(T_+^{\alpha_0,0} - \frac{\partial}{\partial n} \right) U_{\alpha,0} = 2\mathrm{i}\beta_1^0\, \mathrm{e}^{-\mathrm{i}\beta_1^0 B}, \quad \text{on } \Gamma_+
\tag{2.6}
$$

and

$$
\left(T_-^{\alpha_0,0} - \frac{\partial}{\partial n} \right) U_{\alpha,0} = 0, \quad \text{on } \Gamma_-.
\tag{2.7}
$$

where n represent the outward unit normal and in addition for Case 1

$$
\left.
\begin{aligned}
U_{\alpha,0}|_{\partial \Omega_3} &= 0, \quad \text{for Case 1A} \\
\partial_n U_{\alpha,0}|_{\partial \Omega_3} &= 0, \quad \text{for Case 1B.}
\end{aligned}
\right\}
\tag{2.8}
$$

and

The exact solution of the boundary value problem is denoted by $U_{\alpha,0}$ and $U_{\alpha,0_h} \in X^\alpha$, such that $X^\alpha \subset H_\#^1(F)$ is the corresponding numerical solution.

(b) Continuous problem

The continuous variational formulations given for the four cases are represented as a single problem as follows. Find $U_{\alpha,0} \in H_\#^1(F)$ such that

$$
a(U_{\alpha,0}, v) = (f, v)_{\Gamma_+}
\tag{2.9}
$$

for all $v \in H^1_\#(F)$ with

$$
a(U_{\alpha,0}, v) = \begin{cases}
\left(\dfrac{1}{q}\nabla U_{\alpha,0}, \nabla v\right)_F - \left(\dfrac{1}{q}(k^2 - \alpha_0^2)U_{\alpha,0}, v\right)_F \\
-2i\alpha_0\left(\dfrac{1}{q}\partial_x U_{\alpha,0}, v\right)_F - \left(\dfrac{1}{q}T^{\alpha,0}_\pm U_{\alpha,0}, v\right)_{\Gamma_\pm} \\
+i(\alpha_0 n_x U_{\alpha,0}, v)_{\partial F}, & \text{for Case 1B} \\[2ex]
\left(\dfrac{1}{q}\nabla U_{\alpha,0}, \nabla v\right)_F - \left(\dfrac{1}{q}(k^2 - \alpha_0^2)U_{\alpha,0}, v\right)_F \\
-2i\alpha_0\left(\dfrac{1}{q}\partial_x U_{\alpha,0}, v\right)_F - \left(\dfrac{1}{q}T^{\alpha,0}_\pm U_{\alpha,0}, v\right)_{\Gamma_\pm}, & \text{otherwise}
\end{cases}
\tag{2.10}
$$

and

$$
(f, v)_{\Gamma_+} = -2i\beta_1^0 \int_{\Gamma_+} \frac{1}{q}\, e^{-i\beta_1^0 B}\,\overline{v}, \tag{2.11}
$$

where

$$
q = \begin{cases} k^2 & \text{for Case 2B} \\ 1 & \text{for Cases 1A, 1B and 2A,} \end{cases} \tag{2.12}
$$

such that k is given by equation (2.1) and n_x is the outward unit normal with respect to the x-axis. For Case 2B, because both $U_{\alpha,0}$ and v are periodic, then

$$
\left(\partial_x U_{\alpha,0}, v\right)_F = -\left(U_{\alpha,0}, \partial_x v\right)_F. \tag{2.13}
$$

(c) Discretized problem

If X^α is a finite-dimensional subspace of $H^1_\#(F)$ [17–19], then the discrete problem is to find $U_{\alpha,0_h} \in X^\alpha$ such that

$$
a(U_{\alpha,0_h}, v_h) = (f, v_h)_{\Gamma_+} \tag{2.14}
$$

for all $v_h \in X^\alpha$, with $a(\cdot, \cdot)$ and $(f, \cdot)$ as given, respectively, by equations (2.10) and (2.11).

Where q and F are given, respectively, by equations (2.12) and (2.5). The discretization error is denoted by

$$
e_h = U_{\alpha,0} - U_{\alpha,0_h}. \tag{2.15}
$$

In order to establish an *a posteriori* error estimate, the dual problem is studied in §3.

3. Dual problem

Here, the DWR method is used to provide a goal-oriented error estimate [13]. A quantity of interest is defined and shown to be a continuous linear functional, as this is a necessary condition for the dual problem to be well posed. The dual problem is then formulated and used to study the error in the quantity of interest.

(a) Quantity of interest

Let $f \in H^{1/2}_\#([0, d])$, and define the map Q, where $Q(f) \in \mathbb{C}$, as

$$
Q(f) = \frac{1}{q} \sum_{|\alpha_n| < |k|} c_n f^{(n)}(\pm B), \tag{3.1}
$$

such that $c_n = 0$ or $c_n = 1$ and the Fourier coefficient $f^{(n)}(\pm B)$ of $f(x, \pm B)$ is given by

$$
f^{(n)}(\pm B) = \frac{1}{d} \int_0^d f(x, \pm B)\, e^{-i(2\pi n/d)x}\, dx.
$$

Relating these to the diffraction efficiency [8, §§ 4.1.4 and 4.2.3] gives

$$\sum_{|\alpha_n|<|k|} |r_1^n| + |t_2^n| = \sum_{|\alpha_n|<|k_1|, n\neq 0} |U_{\alpha,0}^{(n)}(B) + U_{\alpha,0}^{(0)}(B) - e^{-i\beta_1^0 B}|$$

$$+ \sum_{|\alpha_n|<|k_2|} |U_{\alpha,0}^{(n)}(-B) e^{\Im(\beta_2^n)B}|,$$

$$\leq \sup_{|\alpha_n|<|k_2|} e^{\Im(\beta_2^n)B} \sum_{|\alpha_n|<|k_j|} |U_{\alpha,0}^{(n)}(\pm B)| + 1,$$

$$\leq q\, e^{\Im(k_2)B} Q(|U_{\alpha,0}|) + 1,$$

because $\Im(k_2) \geq \Im(\beta_2^n) \geq 0$, for $|\alpha_n| < |k_2|$, where $c_n = 1$ was used in equation (3.1) to keep notation simple. r_1^n and t_2^n are complex scalars called Rayleigh coefficients used to compute the diffraction efficiency of order n [8]. Hence, the quantity of interest is chosen so that it is related to the computation of the diffraction efficiencies corresponding to the propagating waves, given the condition $|\alpha_n| < |k|$. In addition, the constants c_n have been introduced so that a particular efficiency order m can be isolated. Hence, $c_n = 1$ if $m = n$ and $c_n = 0$ otherwise.

Lemma 3.1. *Let $f \in H_\#^{1/2}(\Gamma_\pm)$, $k_{\mathrm{ref}} > 0$ such that $|k| > k_{\mathrm{ref}}$, C a positive constant and let Q be given by equation (3.1), then Q is a continuous linear functional and*

$$|Q(f)| \leq \frac{\sqrt{2}}{k_{\mathrm{ref}}^2} \ln(|k| + C + |\alpha_0| + \sqrt{1 + (|k| + C + |\alpha_0|)^2})^{1/2} \, \|f\|_{H_\#^{1/2}(\Gamma_\pm)}.$$

Proof. The linearity of Q follows from its definition. From Schwarz's inequality

$$|Q(f)| = \left| \sum_{|\alpha_n|<|k|} \frac{1}{q} f^{(n)}(\pm B) \left(1 + \left(\frac{2\pi n}{d}\right)^2\right)^{1/4} \left(1 + \left(\frac{2\pi n}{d}\right)^2\right)^{-1/4} \right|,$$

$$\leq \frac{1}{k_{\mathrm{ref}}^2} \left(\sum_{|\alpha_n|<|k|} |f^{(n)}(\pm B)|^2 \left(1 + \left(\frac{2\pi n}{d}\right)^2\right)^{1/2} \right)^{1/2} \left(\sum_{|\alpha_n|<|k|} \frac{1}{\sqrt{1 + (2\pi n/d)^2}} \right)^{1/2}.$$

As the sum $\sum_{|\alpha_n|<|k|} (1/\sqrt{1 + (2\pi n/d)^2})$ is finite and bounded by a constant. We can estimate this constant as follows:

$$\sum_{|\alpha_n|<|k|} \frac{1}{\sqrt{1 + (2\pi n/d)^2}} \leq \int_{-|k|-C}^{|k|+C} \frac{1}{\sqrt{1 + (\alpha_n - \alpha_0)^2}} \, \mathrm{d}\alpha_n,$$

with C a positive constant, then the change of variable $x = \alpha_n - \alpha_0$ gives

$$\sum_{|\alpha_n|<|k|} \frac{1}{\sqrt{1 + (2\pi n/d)^2}} \leq \int_{-|k|-C-\alpha_0}^{|k|+C-\alpha_0} \frac{1}{\sqrt{1 + x^2}} \, \mathrm{d}x,$$

$$= \int_0^{|k|+C+\alpha_0} \frac{1}{\sqrt{1 + x^2}} \, \mathrm{d}x + \int_0^{|k|+C-\alpha_0} \frac{1}{\sqrt{1 + x^2}} \, \mathrm{d}x$$

Evaluating the integrals gives

$$\sum_{|\alpha_n|<|k|} \frac{1}{\sqrt{1 + (2\pi n/d)^2}} = [\ln(x + \sqrt{1 + x^2})]_0^{|k|+C+\alpha_0} + [\ln(x + \sqrt{1 + x^2})]_0^{|k|+C-\alpha_0}$$

$$\leq 2\ln(|k| + C + |\alpha_0| + \sqrt{1 + (|k| + C + |\alpha_0|)^2}).$$

Hence,

$$|Q(f)| \leq \frac{\sqrt{2}}{k_{\mathrm{ref}}^2} \ln(|k| + C + |\alpha_0| + \sqrt{1 + (|k| + C + |\alpha_0|)^2})^{1/2}$$

$$\times \left(\sum_{\alpha_n \in \mathbb{Z}} |f^{(n)}(\pm B)|^2 \left(1 + \left(\frac{2\pi n}{d} \right)^2 \right)^{1/2} \right)^{1/2},$$

$$\leq \frac{\sqrt{2}}{k_{\mathrm{ref}}^2} \ln(|k| + C + |\alpha_0| + \sqrt{1 + (|k| + C + |\alpha_0|)^2})^{1/2} \, \|f\|_{H_{\#}^{1/2}(\Gamma_{\pm})}$$

from the definition of the $H_{\#}^s$-norm [20, p. 150] for $s = \frac{1}{2}$. ∎

(b) Strong form and variational form of the dual problem

A dual function $z \in H_{\#}^2(F)$ is introduced in order to estimate $Q(U_{\alpha,0}) - Q(U_{\alpha,0_h}) = Q(e_h)$, and it satisfies

$$(Av, z) = a(v, z) = Q(v) \quad \forall v \in H_{\#}^1(F), \tag{3.2}$$

where a is given by equation (2.10) and A is given by equation (2.3) or (2.4). Hence, the strong form of the dual problem given by equation (3.2) is provided by lemma 3.2.

Lemma 3.2. *Let $z \in H_{\#}^2(F)$ be the dual solution corresponding to the dual problem of equation* (2.10) *then z satisfies*

$$J_{\pm} = \begin{cases} -q\nabla \cdot \left(\frac{1}{q} \nabla \right) z + 2i\alpha_0 \partial_x z - (k^2 - \alpha_0^2)z + \partial_n z \\ \quad -T_{\pm}^{*\alpha,0} z + i\alpha_0 n_x z & \text{for Case 1B} \\ -q\nabla \cdot \left(\frac{1}{q} \nabla \right) z + 2i\alpha_0 \partial_x z - (k^2 - \alpha_0^2)z + \partial_n z - T_{\pm}^{*\alpha,0} z & \text{otherwise.} \end{cases} \tag{3.3}$$

The functional $J_{\pm}$ is defined as

$$J_{\pm} = \sum_{|\alpha_n| < |k_j|} \frac{c_n}{d} \, e^{-i(2\pi n/d)x}, \tag{3.4}$$

with $j = 1, 2$, respectively, for J_+ (on Γ_+) and J_- (on Γ_-), and $T_{\pm}^{*\alpha,0}$ is the dual of $T_{\pm}^{\alpha,0}$ [21, p. 476].

Proof. The divergence theorem and applying integration by parts to $a(v, z)$ in equation (3.2) gives

$$\left(\frac{1}{q} \nabla v, \nabla z \right)_F = \left(\nabla v, \frac{1}{q} \nabla z \right)_F = \left(v, -\nabla \cdot \left(\frac{1}{q} \nabla \right) z \right)_F + \left(v, \frac{1}{q} \partial_n z \right)_{\partial F}$$

and

$$\int_F (\partial_{x_i} v) z = -\int_F v(\partial_{x_i} z) + \int_{\partial F} (vz) n_i \, \mathrm{d}s,$$

where n_i is the outward unit normal in the direction of x_i. This leads to

$$\int_F (\partial_x v) \bar{z} = -\int_F v \partial_x \bar{z},$$

from equation (2.13). Hence, for Cases 1A and 2,

$$a(v, z) = \left(v, -\nabla \cdot \left(\frac{1}{q} \nabla z \right) \right)_F + 2i\alpha_0 \left(\frac{1}{q} v, \partial_x z \right)_F + \left(\frac{1}{q} v, \partial_n z \right)_{\partial F}$$

$$- \left(\frac{1}{q} (k^2 - \alpha_0^2) v, z \right)_F - \left(\frac{1}{q} v, T_{\pm}^{*\alpha,0} z \right)_{\Gamma_{\pm}} = Q(v),$$

for all $v \in H^1_\# F$. Let $\phi \in H^{1/2}_\#(\Gamma)$, then

$$T^{*\alpha,0}_\pm \phi(x) = \sum_{n \in \mathbb{Z}} i\beta^n_j \overline{\phi}^{(n)}(\pm B) \, e^{-i(2\pi n/d)x},$$

and $\overline{\phi}^{(n)}(y) = (1/d) \int_0^d \overline{\phi(x,y)} \, e^{i(2\pi n/d)x} \, dx$. Note that

$$Q(v) = \left(\frac{1}{q} v, J_\pm \right), \qquad (3.5)$$

such that $J_\pm$ is given by (3.4). For Case 1B, $(vn_x, z)_{\partial F} = (v, n_x z)_{\partial F}$. ∎

As the exact dual solution z is not known, a weak formulation is used in the numerical implementation to approximate z. From lemma 3.2, the variational formulation corresponding to the dual problem is given as follows.

Let $z \in H^1_\#(F)$ (weak solution) then, for any $\psi \in H^1_\#(F)$, for Cases 1A, 2A and 2B

$$a^*(\psi, z) = \left(\frac{1}{q} \nabla \psi, \nabla z \right)_F + 2i\alpha_0 \left(\frac{1}{q} \psi, \partial_x z \right)_F - \left(\frac{1}{q}(k^2 - \alpha_0^2)\psi, z \right)_F$$
$$- \left(\frac{1}{q} \psi, T^{*\alpha,0}_\pm z \right)_{\Gamma_\pm} = \left(\frac{1}{q} \psi, J_\pm \right)_F, \qquad (3.6)$$

and for Case 1B

$$a^*(\psi, z) = \left(\frac{1}{q} \nabla \psi, \nabla z \right)_F + 2i\alpha_0 \left(\frac{1}{q} \psi, \partial_x z \right)_F - \left(\frac{1}{q}(k^2 - \alpha_0^2)\psi, z \right)_F$$
$$+ \frac{i\alpha_0}{q}(\psi, zn_x)_{\partial F} - \left(\frac{1}{q} \psi, T^{*\alpha,0}_\pm z \right)_{\Gamma_\pm} = \left(\frac{1}{q} \psi, J_\pm \right)_F.$$

Having formulated the direct and the dual problem, they can be used to establish a goal-oriented error estimate, using the DWR method, where the quantity of interest is related to the grating efficiency.

4. A posteriori error estimation

Let z be the dual solution associated with the dual problem given by lemma 3.2. Then, for any $\phi \in H^1_\#(F)$,

$$a(\phi, z) = (\phi, A^* z)_F = \left(\frac{1}{q} z, J_\pm \right)_F,$$

such that

$$A^* = \begin{cases} -\nabla \cdot \left(\frac{1}{q} \nabla \right) + \frac{2i\alpha_0}{q} \partial_x - \frac{1}{q}(k^2 - \alpha_0^2) + \frac{1}{q} \partial_n \\ \quad -\frac{1}{q} T^{*\alpha,0}_\pm + i\alpha_0 n_x \qquad\qquad\qquad\qquad\qquad \text{for Case 1B} \\ -\nabla \cdot \left(\frac{1}{q} \nabla \right) + \frac{2i\alpha_0}{q} \partial_x - \frac{1}{q}(k^2 - \alpha_0^2) + \frac{1}{q} \partial_n - \frac{1}{q} T^{*\alpha,0}_\pm \quad \text{otherwise,} \end{cases}$$

where A^* is the dual operator of A defined in equation (2.3) or (2.4), [21, p. 476]. Using equation (3.5) leads to the following problem: find $z \in H^1_\#(F)$ such that

$$a^*(\phi, z) = (\phi, A^* z)_F = Q(\phi), \qquad (4.1)$$

for any $\phi \in H^1_\#(F)$.

Let X^α be a finite-element subspace of order p of $H^1_\#(F)$ and let ζ_h be any regular partition of X^α [17,18,22,23]. Denote by h the maximum mesh size of the triangular elements in this partition. The finite-element approximation of the dual problem given by equation (4.1) is to find $z_h \in X^\alpha$ such that $a(\phi_h, z_h) = (\phi_h, A^* z_h)_F = Q(\phi_h)$, for any $\phi_h \in X^\alpha$.

The estimate of the linear functional of the error $Q(e_h)$ is given by the following theorem.

Theorem 4.1. *Let $U_{\alpha,0_h}$ be the solution to equation (2.14), e_h be given by equation (2.15), $\zeta_h = \{K\}$ be a partition of X^α, and p_K and h_K be the polynomial order and the mesh size associated with the element K. Denote the field equation residual by $R_h(K)$, then*

$$R_h(K) := \begin{cases} \nabla_{\alpha,0}\left(\dfrac{1}{k^2}\nabla_{\alpha,0}U_{\alpha,0_h}\right) + U_{\alpha,0_h} & \text{for Case 2B} \\[2ex] \nabla.\left(\dfrac{1}{q}\nabla U_{\alpha,0_h}\right) - \dfrac{2i\alpha_0}{q}\partial_x U_{\alpha,0_h} - \dfrac{1}{q}(k^2 - \alpha_0^2)U_{\alpha,0_h} & \text{otherwise.} \end{cases} \tag{4.2}$$

The flux residual $r_h(E)$ for Case 2 is given by

$$r_h(E) := \begin{cases} -\dfrac{1}{2q}[\partial_n U_{\alpha,0_h}] & \text{if } E \subset \partial K \setminus \Gamma_\pm, \\[2ex] \dfrac{1}{q}T_+^{\alpha,0}U_{\alpha,0_h} - \dfrac{1}{q}2i\,e^{-i\beta_1^0 B} & \text{if } E \subset \Gamma_+, \\[2ex] \dfrac{1}{q}T_-^{\alpha,0}U_{\alpha,0_h} & \text{if } E \subset \Gamma_-, \end{cases} \tag{4.3}$$

and for Case 1 by

$$r_h(E) := \begin{cases} -\dfrac{1}{2q}[\partial_n U_{\alpha,0_h}] & \text{if } E \subset \partial K \setminus (\Gamma_\pm \cup \partial\Omega_3), \\[2ex] \dfrac{1}{q}T_+^{\alpha,0}U_{\alpha,0_h} - \dfrac{1}{q}2i\,e^{-i\beta_1^0 B} & \text{if } E \subset \Gamma_+, \\[2ex] \dfrac{1}{q}T_-^{\alpha,0}U_{\alpha,0_h} & \text{if } E \subset \Gamma_-, \end{cases} \tag{4.4}$$

with $[\partial_n U_{\alpha,0_h}]$ denoting the jump of the normal derivatives of $U_{\alpha,0_h}$, E an edge of the element K and q as defined by equation (2.12). Then

$$|Q(e_h)| \leq \sum_{K \in \zeta_h} \rho_K w_k, \tag{4.5}$$

where the cell residuals ρ_k and weights w_K are given by

$$\rho_K := \|\, R_h(K)\,\|_{L^2_{\alpha\#}(K)} + (h_K)^{-1/2}\,\|\,r_h(E)\,\|^2_{L^2_{\alpha\#}(E)} \tag{4.6}$$

and

$$w_K := \|\,z - z_h\,\|_{L^2_{\alpha\#}(K)} + (h_K)^{1/2}\,\|\,z - z_h\,\|^2_{L^2_{\alpha\#}(E)} \tag{4.7}$$

with $L^2_{\alpha\#}(\cdot) = H^0_{\alpha\#}(\cdot)$.

Proof. From equations (3.2) and (2.14), $Q(U_{\alpha,0}) - Q(U_{\alpha,0_h}) = a(U_{\alpha,0} - U_{\alpha,0_h}, z)$, and using equations (2.9) and (3.2) $Q(e_h) = (f,z)_{\Gamma_+} - a(U_{\alpha,0_h}, z) = (f,z)_{\Gamma_+} - (AU_{\alpha,0_h}, z)$. Let ζ_h be a partition of the domain F into mesh cells K. From cellwise integration by parts [24, p. 28], [25, p. 12] and by using Galerkin orthogonality [17, p. 58] with equation (2.3) for Cases 1A, 1B and 2A give

$$Q(e_h) = \sum_K (\Delta U_{\alpha,0_h} + 2i\alpha_0\partial_x U_{\alpha,0_h} + (k^2 - \alpha_0^2)U_{\alpha,0_h}, z - \phi)_K$$

$$- (\nabla U_{\alpha,0_h}.n + f, z - \phi)_{\partial K}$$

for all $\phi \in X^\alpha$ and $\bigcup_{K \in \zeta_h} = F$. Using a similar argument for Case 2B gives

$$Q(e_h) = \sum_K \left(\nabla_{\alpha,0}\left(\frac{1}{k^2}\nabla_{\alpha,0}U_{\alpha,0_h}\right) + U_{\alpha,0_h}, z - \phi\right) - (\nabla U_{\alpha,0_h}.n + f, z - \phi)_{\partial K}.$$

Using equations (2.6)–(2.8) for Case 1 gives equations (4.2)–(4.4). Then

$$|Q(e_h)| \leq \sum_K |(R_h(K), z - \phi)_K + (r_h(E), z - \phi)_{\partial K}|.$$

From the Cauchy–Schwarz inequality [17, p. 50]

$$|Q(e_h)| \leq \sum_K \parallel R_h(K) \parallel_{L^2_{\alpha\#}(K)} \parallel z - \phi \parallel_{L^2_{\alpha\#}(K)} + \parallel r_h(E) \parallel_{L^2_{\alpha\#}(E)} \parallel z - \phi \parallel_{L^2_{\alpha\#}(E)},$$

$$\leq \sum_K (\parallel R_h(K) \parallel_{L^2_{\alpha\#}(K)} + h_K^{-1/2} \parallel r_h(E) \parallel_{L^2_{\alpha\#}(E)})$$

$$\times (\parallel z - \phi \parallel_{L^2_{\alpha\#}(K)} + h_K^{1/2} \parallel z - \phi \parallel_{L^2_{\alpha\#}(E)}),$$

$$= \sum_K \rho_K w_K.$$

The proof is completed by choosing $\phi = z_h$ and by using equations (4.6) and (4.7). ∎

5. Numerical methods and numerical results

For completeness a pseudo-code algorithm for the implementation of the method proposed in this paper is detailed in §5a. After that, three numerical examples are used to illustrate the computational benefits of this adaptive method.

(a) Algorithm for the adaptive mesh

The operations are done elementwise by looping over all elements of a given triangulation. Hence, each element is mapped to a reference element through an affine transformation.

— step 0. Choose a tolerance TOL;
— step 1. Assemble the mass matrix for $U_{\alpha,0} = \sum U_i \psi_i$ where ψ_i is some basis in the corresponding finite-element subspace X^α.
Using a reference element, ψ_i becomes $\tilde{\psi}_i$. Write $a(\psi_i, \psi_j) = K_{ij} + B_{ij}$ and use equation (2.10) to give

$$K_{ij} = \tilde{b}_1(\nabla\tilde{\psi}_i, \nabla\tilde{\psi}_j)_F + \tilde{b}_2(\partial_x\tilde{\psi}_i, \tilde{\psi}_j)_F + \tilde{b}_3(\tilde{\psi}_i, \partial_x\tilde{\psi}_j)_F + \tilde{b}_4(\partial_y\tilde{\psi}_i, \tilde{\psi}_j)_F$$

$$+ \tilde{b}_5(\tilde{\psi}_i, \tilde{\psi}_j)_F + \tilde{b}_6(n_x\tilde{\psi}_i, \tilde{\psi}_j)_{\partial\Omega_3} \pm \tilde{b}_7(\tilde{\psi}_i, \tilde{\psi}_j)_{\Gamma_\pm}.$$

Numerical integration is used to calculate the surface and element integrals to get K_{ij} [26–28];
— step 2. Assemble the load vector.
Similar to step 1, equation (2.11) becomes $(\tilde{b}_{10}, \tilde{\psi}_j)_{\Gamma_+}$ which is also computed using numerical integration;
— step 3. Assemble the DtN operators.
From step 1, $B_{ij} = b_8(T_+^{\alpha,0M}\tilde{\psi}_i, \tilde{\psi}_j)_{\Gamma_+} + b_9(T_-^{\alpha,0M}\tilde{\psi}_i, \tilde{\psi}_j)_{\Gamma_+}$, where

$$(T_\pm^{\alpha,0M}\tilde{\psi}_i, \tilde{\psi}_j)_{\Gamma_\pm} = d \sum_{m=-M}^{M} i\beta_l^m \tilde{\psi}_i^{(m)}(\pm B)\overline{\tilde{\psi}_j^{(m)}(\pm B)}. \tag{5.1}$$

$\tilde{\psi}_j^{(m)}$ can be computed as follows using equation (2.2):

$$\tilde{\psi}_j^{(m)}(\pm B) = \frac{1}{d} \int_{x_j}^{x_{j+1}} \tilde{\psi}_j(x, \pm B) e^{-i2\pi m/d x} dx, \tag{5.2}$$

where x_j and x_{j+1} delimit the element edge where $\tilde{\psi}$ is different from zero and hence supp $\tilde{\psi}_j(x, \pm B) = [x_j, x_{j+1}]$. Since $\tilde{\psi}_j(x, \pm B)$ is a polynomial of order N then

$$\tilde{\psi}_j(x, \pm B) = \sum_{l=0}^{N} c_l x^l.$$

Hence, equation (5.2) becomes

$$\tilde{\psi}_j^{(m)}(\pm B) = \frac{1}{d} \int_{x_j}^{x_{j+1}} \sum_{l=0}^{N} c_l x^l \, e^{-i2\pi m/ dx} \, dx,$$

$$= \frac{1}{d} \sum_{l=0}^{N} c_l G_l(x_j, x_{j+1}, m, d), \tag{5.3}$$

where $G_l(x_j, x_{j+1}, m, d) = \int_{x_j}^{x_{j+1}} x^l \, e^{-i2\pi m/ dx} \, dx$. It can be shown by integrating by parts and by induction that

$$G_l(x_j, x_{j+1}, l, m, d) = \begin{cases} \dfrac{x^{l+1}}{l+1}, & \text{if } m=0 \\[2mm] \displaystyle\sum_{t=0}^{q} \dfrac{l!}{(l-t)!(2i\pi m/d)^t} \left[x^{l-t} \dfrac{e^{-2i\pi m/d}}{-2i\pi m/d} \right]_{x=x_j}^{x=x_{j+1}} \\[4mm] \quad + \dfrac{l!}{1-(q+1)!} \dfrac{1}{(2i\pi m/d)^{q+1}} G_{l-(q+1)}, & \text{otherwise} \end{cases} \tag{5.4}$$

for $q \leq l-1$. Equations (5.1), (5.3) and (5.4) can then be used to compute B_{ij};
— step 4. Apply the periodicity boundary conditions for all cases and the Dirichlet boundary conditions for Case 1A.
The problem is find $U_{\alpha,0}$ such that $a(U_{\alpha,0}, \psi) = (\hat{f}, \psi)_{\Gamma_+}$. By denoting $(\hat{f}, \tilde{\psi}_i)_{\Gamma_+} = F_i$ and by noting that $a(\tilde{\psi}_i, \tilde{\psi}_j) = K_{ij} + B_{ij}$ the solution $\sum_{j=1}^{N} U_j \phi_j(x,y)$ satisfies

$$(K_{ij} + B_{ij})U_j = F_i \tag{5.5}$$

for $i,j = 1, \ldots N$. For the periodic boundary condition, let N_p denote the set of nodes i such that ϕ_i belongs to the boundary $x = 0$ or $x = d$ and N_O denote the set of nodes j such that ϕ_j belongs to $\partial\Omega_3$ for Case 1A. The techniques described in Ainsworth [29] are used to apply the periodicity constraints and the Dirichlet constraints to the nodes $i \in N_p$ and $j \in N_O$;
— step 5. Solve the system.
Solve system (5.5) and find the coefficients of the approximate solution to the Helmholtz problem;
— step 6. Repeat steps 1–5 to solve the dual problem to find z_h using equation (3.6).
As the variational form of the dual problem as given in equation (3.6) has a form similar to the direct problem, the same technique is used to solve the dual problem;
— step 7. Repeat steps 1–5 to solve the dual problem using equation (3.6) with a higher polynomial order or a finer mesh to give an approximation.
— step 8. Compute the upper bound given by equation (4.5) and denoted this upper bound by

$$I_{\text{err}} = \sum \rho_K w_K. \tag{5.6}$$

Standard techniques can be used (reference element and numerical integration) [26,28,30, 31] to compute equation (4.2). For the flux residual given by equations (4.3) and (4.4), the $T_{\pm}^{\alpha,0^M} \psi_j$ are computed using definition 2.1 and equation (5.3). For the jump derivative $[\nabla \psi_i.n]$, let us denote by K and K^1 two triangles which share an edge E, and let the affine transformation which maps K to the reference element (and ψ_i to $\tilde{\psi}_i$), be described by $Jx + b$ where J is a 2×2 matrix. Let the affine transformation which maps K^1 to the reference element (and ψ_i^1 to $\tilde{\psi}_i^1$), be described by $J^1 x + b$ where J^1 is a 2×2 matrix. We then have

$$[\nabla \psi_i \cdot n] = (J^{-T} \nabla \tilde{\psi}_i + J^{1-T} \nabla \tilde{\psi}_i^1) \cdot n$$

because $n^1 = -n$. Hence, using $y = g(x)$ to represent the curve $\partial K \bigcap \partial K^1$, gives

$$\int_E |[\nabla \psi_i \cdot n]|^2$$

$$= \int_{x=x_0}^{g(x_0)} |\sqrt{1 + g'(x)^2} J^{-T} \nabla \tilde{\psi}_i(x, g(x)) + \sqrt{1 + g'(x)^2} J^{1-T} \nabla \tilde{\psi}_i^1(x, g(x))|^2 \, dx,$$

where the edge E is delimited by $g(x_0)$ and x_0.

— step 9. Let N_T be the list of elements K where $\rho(K) w_K$ is in decreasing order and such that $\sum_{N_T} \rho(K) w_K \geq 0.7 I_{\text{err}}$. Refine the mesh in N_T;

— step 10. Check that the mesh is periodic if it is not defined $\Gamma_L = \{y : (0, y) \in \Omega\}$ and $\Gamma_R = \{y : (d, y) \in \Omega\}$. Refine the edge elements on these boundaries until $\Gamma_L = \Gamma_R$ so that the mesh is periodic; and

— repeat steps 1–10 until $I_{\text{err}} \leq \text{TOL}$.

(b) Numerical results

Having briefly described the numerical technique used to solve numerically the Helmholtz problem, the benefits that can be obtained by using an adaptive grid, and in particular where this adaptivity is driven by the DWR method is illustrated below.

Example 5.1. In the following example, the transmitting dielectric lamellar grating as shown in figure 2 is considered. This type of grating, used in modelling multiscale phenomena grating problems, has been studied in Bao *et al.* [32] using a hybrid approach that combines a perfectly matching layer technique and an adaptive finite-element method driven by a global *a posteriori* error estimate and it can be applied in solving optimal design problems. Here, the wavenumbers are fixed as $k_1 = 2\pi$ and $k_2 = (0.22 + 6.71i)2\pi$, the angle of incidence is $\theta = \pi/6$ and the period is $d = 1$. A TM field (Case 2B) is considered as described by equations (2.10) and (2.11), and the reflection efficiency of order zero is computed (R_0^M) [8, sections 4.2.3]. To use the DWR method, equations (3.6) and (4.5) are employed to find the dual solution and the error bounds. The algorithm in §5a is followed with $\text{TOL} = 10^{-4}$. Let d.f. denote the degrees of freedom, that is the total number of nodes needed to define all triangular elements on the domain, and all the polynomial bases on each of these elements. To provide a basis for a relative error, the global method in Bao *et al.* [32] is used with 201205 d.f. which gives $R_0^E = 0.8484815$. The relative error using the global *a posteriori* error estimate in Bao *et al.* [32] and the DWR method can be compared by defining $\epsilon(\text{DWR/global}) = |R_0^M(\text{DWR/global}) - R_0^E|/R_0^E$. Figure 3a shows that the DWR method converges faster than the global *a posteriori* error method studied in Bao *et al.* [32]. Also note that the indicative computed error I_{err} given in equation (5.6) decreases monotonically which shows the convergence of the DWR method. A relative error of 10^{-6} has been chosen as the stopping criteria because the mesh becomes very irregular after this and will lead to a numerical instability [33,34]. In addition, as the focus is on the local error in the quantity of interest, there is a pollution from the neglected global error. In practice, this problem can be avoided by refining the mesh only in the neighbourhood of the high stress gradients; that is, at the singularities in the geometry. However, the areas of high stress gradients cannot be guaranteed to coincide with the areas of interest and this may lead to pollution error. Hence, the goal-oriented error estimation method could be improved by using a balance between the local error and the global error so that the mesh is refined to assure a high level of accuracy of the quantity of interest and at the same time to not underestimate the effect of the global error [14].

Example 5.2. Here, the DWR method is compared with the use of a uniform mesh. As above, the relative error using the uniform mesh and the DWR method is defined by $\epsilon(\text{DWR/uniform}) = |R_0^M(\text{DWR/uniform}) - R_0^E|/R_0^E$. Figure 3b shows that the error associated with the DWR method does not decrease monotonically unlike that associated with the uniform mesh when we increase d.f. However, when d.f. $> 10^4$, the DWR converges faster and requires fewer degrees of freedom than using the uniform mesh.
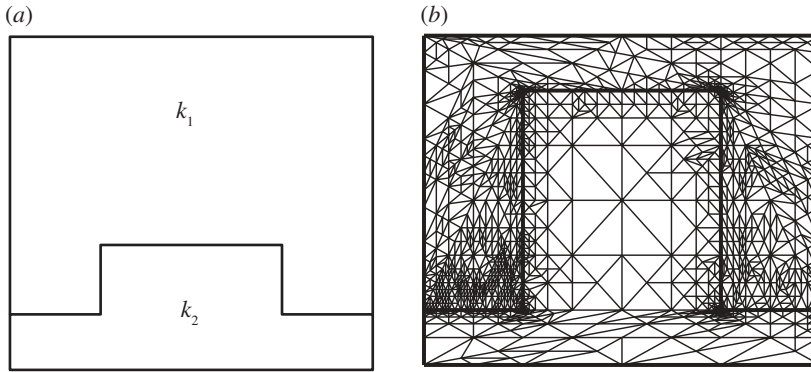
Figure 2. Transmitting dielectric lamellar grating is shown in (*a*). The corresponding adaptive mesh used with DWR is shown in (*b*).
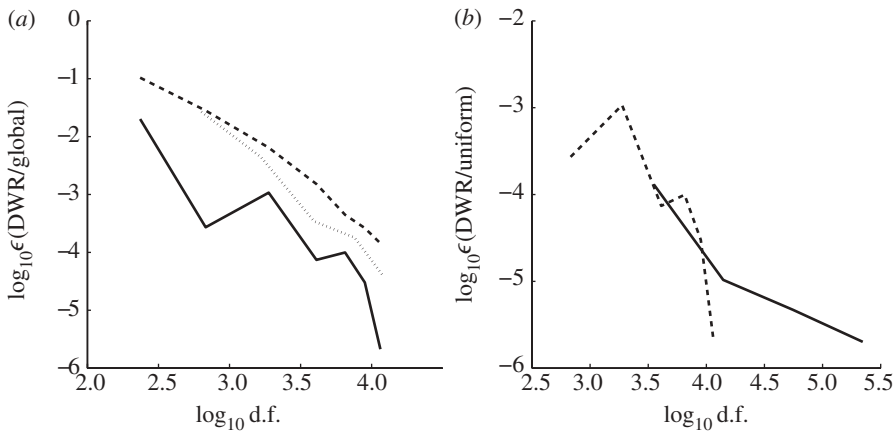

Figure 3. The transmitting dielectric lamellar grating shown in figure 2 in a TM field (Case 2B). (*a*) The relative error in computing the reflection efficiency R_0^M, using the global *a posteriori* error estimate in Bao *et al.* [32] (dashed line), the DWR method (solid line) and the indicative computed error I_{err} (dotted line) given by equation (5.6) which shows an upper bound of the discretization error in solving numerically the Helmholtz problem. (*b*) The relative error in computing the reflection efficiency R_0^M, using uniform mesh (solid line) and the DWR method (dashed line).

Example 5.3. One of the main reasons for using a finite-element approach is of course its ability to tackle any prescribed geometry. The examples discussed so far have dealt with simple geometries and this of course has enabled a comparison with other results in the literature. In order to develop more sophisticated grating structures that can help push this technology forward, it is necessary to be able to investigate complex geometries. A recent experimental exploration of what could be achieved by freeing up these geometrical constraints has been inspired by a naturally occurring periodic diffraction grating. When scattered by light, the Morpho butterfly wing produces colour and this allows a dynamical control of light flow and wavelength interaction [35] (figure 4). The scattering of the Morpho butterfly wing has inspired applications in biomimetics such as gas sensors, electronic display screens and paints for cars [2,36–38]. This transmitting dielectric grating type has been studied experimentally in Vukusic *et al.* [35] using a focused laser technique to examine the total reflectivity and transmissivity. In this paper, an adaptive finite-element method driven by the DWR is investigated to study the Morpho butterfly using the experimentally measured geometry. Other numerical attempts have been conducted to study the complex geometry of the Morpho butterfly wings [39] but they have
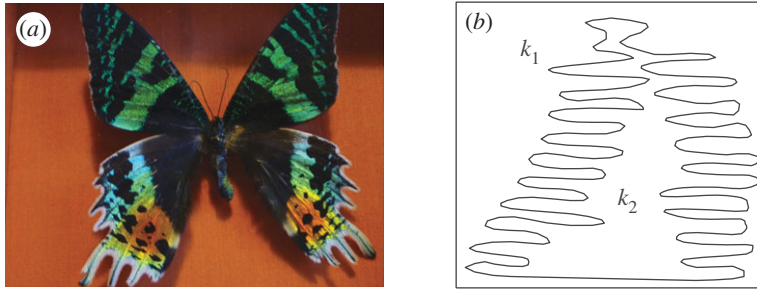
Figure 4. Morpho butterfly wings in (*a*). Reproduction of the image of the Morpho didius butterfly wing over one period *d* is shown in (*b*). (Online version in colour.)
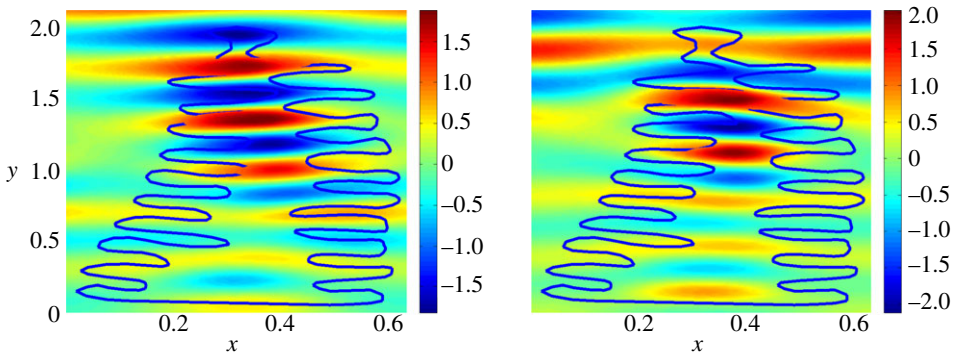

Figure 5. A transmitting dielectric Morpho didius butterfly wing grating (Case 2B) as shown in figure 4*b*. The spatial distribution of the magnetic field is shown for (*a*) the imaginary part and (*b*) the real part. (Online version in colour.)

had to compromise on the geometry. In those studies, the butterfly geometry was approximated by a series of rectangular blocks stacked periodically in the vertical direction. In addition, a poor agreement of the computed transmitted energy with the experimental result was reported for larger wavelengths [39]. In this paper, one period of the butterfly wing image of the Morpho didius in Vukusic *et al.* [35] was digitized, the coordinates of the grating interface extracted and a Savitzky Golay filter [40, pp. 183, 644–645] used to smooth the resulting geometry (figure 4*b*).

The wavenumber inside the butterfly wing was set as $k_2 = (1.57 + 0.06i)2\pi/0.455$, the refractive index is $n_2 = (1.57 + 0.06i)$ and outside the butterfly wing as $k_1 = 2\pi/0.455$; these belong to the range of values given in Vukusic *et al.* [35]. A TM field (Case 2B) was imposed, as described by equations (2.10) and (2.11), and the DWR method was used. Therefore, equations (3.6) and (4.5) were used to find the dual solution and the error bounds. The degrees of freedom used was d.f. $= 11557$, with $M = 15$ Fourier terms and the degree of the polynomial basis $p = 3$. A TOL of 0.001 was used in the algorithm in §5*a* to reproduce the imaginary and real parts of the magnetic field (figure 5). The final indicator computed error I_{err} as given in equation (5.6) shows that the discretization error in solving the Helmholtz problem was kept under $I_{err} = 3 \times 10^{-4}$. The spectra of the transmission and reflection coefficients, in air (solid line) and in isopropyl alcohol (dashed line), corresponding to one layer (a single period is shown in figure 4*b*), are shown, respectively, in figure 6*a*,*b*. A good quantitative comparison with the experimental results (dots) in Vukusic *et al.* [35] corresponding to single iridescent ground scales in air is obtained. The reduced absolute reflection in isopropyl alcohol compared to that in air, as reported in Vukusic *et al.* [35], is also observed. Unfortunately, there were no measured values given for this latter case in order for a quantitative comparison with our theoretical predictions to be made.
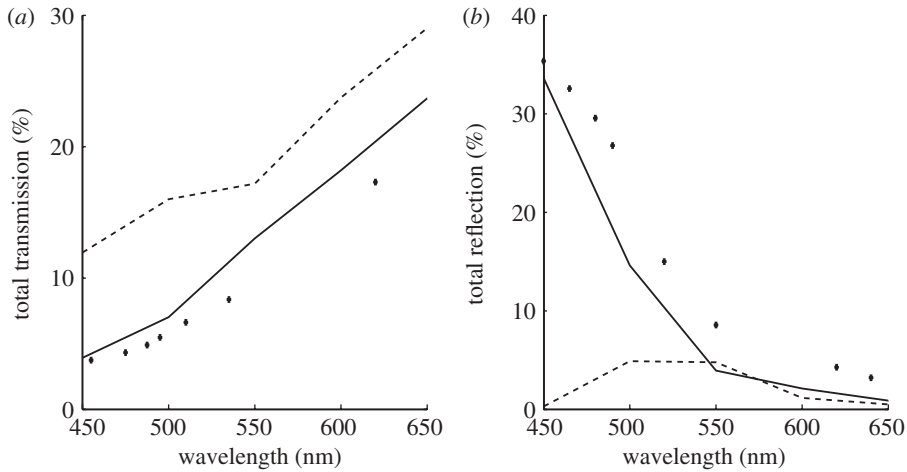
Figure 6. Quantitative comparison of the normal incidence wavelength-dependent (*a*) transmissivity and (*b*) reflectivity for *Morpho didius* (Case 2B). The full line represents data using the DWR method with a single layer as shown in figure 4*b* in air. The dashed line is similarly calculated but now the surrounding medium is isopropyl alcohol. The dots represent the experimental data in air from [35] corresponding to a single iridescent ground scale.

6. Summary and conclusion

This paper considers the use of finite elements to numerically solve the Helmholtz equation in periodic diffraction gratings. There are two ways of establishing an *a posteriori* error estimate which are the global error estimate and the goal-oriented error estimate. This paper considers the latter approach and uses a DWR method. A quantity of interest Q, was defined and shown to be a linear continuous functional. This afforded the formulation of the dual problem and, when combined with the direct problem, was used to establish an upper bound for the error estimate in theorem 4.1. The evaluation of the error in the functional Q represents the primary output from the model, namely the diffraction efficiency.

In the study of Lord & Mulholland [8], the *a priori* error estimate was studied to guarantee that the discrete solution will converge to the exact solution provided that we keep the interpolation error small with respect to the wavenumber k. However, as the upper bounds provided in the *a priori* error estimate depend on the exact solution, they cannot be computed. Hence, numerical implementations are limited to using uniform meshes which are computationally expensive. By contrast, for the *a posteriori* error estimate, the upper bounds given by equation (4.5) can be evaluated. The discrete dual problem z_h is found in the same way as that used to find the discrete direct solution $U_{\alpha,0_h}$ in equation (3.6). For the exact solution of the dual problem z, because it is not known analytically, it can be approximated either by solving the dual problem numerically in a very fine mesh or by increasing the polynomial order using the same mesh as z_h. Hence, the error in the targeted quantity $Q(e_h)$ can be estimated from the local contribution of each error indicator ρ_K and w_K as defined in equations (4.6) and (4.7). In fact, these error indicators are the cell residuals ρ_K multiplied by the weights w_K taken from the computed solution. The ρ_K in turn consists of the field equation residual ($R_h(K)$) and the flux residual ($r_h(E)$) which indicate the smoothness of the discrete solution. The weights w_K capture the influence of the cell residuals on the targeted error $Q(e_h)$ since, differentiating $Q(e_h)$ with respect to ρ_K, gives w_K.

This error estimate affords an automatic mesh adaptation based on the local error indicators ρ_K and w_K. The DWR method uses these error indicators to maximize the accuracy of the computed diffraction efficiency by choosing a TOL and demanding that $\sum_K \rho_K w_K < \text{TOL}$ as the problem is solved. This can be used to optimize the computational efficiency. So, rather than having a uniform mesh, where all elements of the mesh are refined at each step, the refinement is

only performed where the error indicators are large; keeping the coarse mesh where the error indicators are small.

A pseudo-code implementation of this adaptive finite-element method driven by the DWR for solving the periodic grating Helmholtz problem was also described. The merits of using such method were then illustrated for particular examples. In §5*b*, it was shown that the DWR method achieved an acceptable accuracy, converged faster and required fewer degrees of freedom than the global *a posteriori* error estimate proposed in Bao *et al.* [32].

Finally, the geometrical freedom that the finite-element method allows was explored to study a naturally occurring, periodic diffraction grating in the form of a butterfly wing. Its diffraction properties have previously been experimentally measured but this is the first attempt to mirror those results using a finite-element approach. The TM field was successfully calculated and showed the complex interaction between the scatterer and the field. Good quantitative comparison of the experimental results in Vukusic *et al.* [35] with the numerical results presented using here the DWR method for the reflectivity and transmissivity spectrum were also obtained.

There are still some open questions regarding this approach of numerically solving these diffraction problems. It would be interesting to investigate numerically the sensitivity of the grating efficiency to small changes in the geometry of the grating profile. This would be important when considering the inverse problem where the goal is to achieve a desired grating efficiency spectrum by constructing the grating profile that would give rise to this spectrum. It would also be useful to derive an analytical relation to justify the choice of parameters such as the number of the Fourier terms, the size of truncated domain B, the order of the polynomial basis and the mesh size h. It could be shown for example that the accuracy of the numerical results is not significantly affected by the truncation of the DtN operators.

References

1. Peters M, Rüdiger M, Bläsi B, Platzer W. 2010 Electro-optical simulation of diffraction in solar cells. *Opt. Exp.* **18**, A584–A593. (doi:10.1364/OE.18.00A584)
2. Potyrailo RA, Ghiradella H, Vertiatchikh A, Dovidenko JRCK, Olson E. 2007 Morpho butterfly wing scales demonstrate highly selective vapour response. *Nat. Photonics* **1**, 123–128. (doi:10.1038/nphoton.2007.2)
3. Jensen TH, Bech M, Bunk O, Donath T, David C, Feidenhans R, Pfeiffer F. 2010 Directional x-ray dark-field imaging. *Phys. Med. Biol.* **55**, 3317–3323. (doi:10.1088/0031-9155/55/12/004)
4. Zhu P, Zhang K, Wang Z, Liu Y, Liu X, Wu Z, McDonald SA, Marone F, Stampanoni M. 2010 Low-dose, simple, and fast grating-based x-ray phase-contrast imaging. *Proc. Natl Acad. Sci. USA* **107**, 13 756–13 581. (doi:10.1073/pnas.1003198107)
5. Savage N. 2010 Acousto-optic devices. *Nat. Photonics* **4**, 728–729. (doi:10.1038/nphoton.2010.229)
6. Loewen EG, Popov E. 1997 *Diffraction of gratings and applications*. New York, NY: Marcel Dekker.
7. Petit R. 1980 *Electromagnetic theory of gratings*. Berlin, Germany: Springer.
8. Lord N, Mulholland AJ. 2011 Analysis of the α,0-quasi periodic transformation for a periodic diffraction grating. Research Report no. 19, Department of Mathematics and Statistics, University of Strathclyde. See http://www.mathstat.strath.ac.uk/research/reports/2011.
9. Eriksson K, Estep D, Hansbo P, Johnson C. 1995 Introduction to adaptive methods for differential equations. In *Acta numerica*, pp. 105–158. Cambridge, UK: Cambridge University Press.
10. Jin J, Lou Z. 2005 *Finite element modeling of periodic structures*, ch. 3, pp. 129–168. Republic of Singapore: Imperial College Press.
11. Lord N. 2012 Analysis of electromagnetic waves in a periodic diffraction grating using *a priori* error estimates and a dual weighted residual method. Thesis, Department of Mathematics and Statistics, University of Strathclyde. See http://www.mathstat.strath.ac.uk/research/phd_mphil_theses.
12. Popov E. 2012 *Gratings: theory and numeric applications*. France: Institut Fresnel, AMU, CNRS.
13. Bangerth W, Rannacher R. 2003 *Adaptive finite elements methods for differential equations*. Lectures in Mathematics ETH Zürich. Basel, Switzerland: Birkhäuser.

14. Grätsch T, Bathe KJ. 2005 A posteriori error estimation techniques in practical finite element analysis. *Comput. Struct.* **83**, 235–265. (doi:10.1016/j.compstruc.2004.08.011)

15. Maystre D. 1984 Rigorous vector theories of diffraction gratings. In *Progress in optics*, vol. 21 (ed. E. Wolf), pp. 1–67. Amsterdam, The Netherlands: North Holland.

16. Granet G. 1999 Reformulation of the lamellar grating problem through the concept of adaptive spatial resolution. *J. Opt. Soc. Am. A* **16**, 2510–2516. (doi:10.1364/JOSAA.16.002510)

17. Brenner SC, Scott LR. 2002 *The mathematical theory of finite element methods of Texts in Applied Mathematics*, vol. 15. New York, NY: Springer.

18. Ciarlet PG. 1978 *The finite element method for elliptic equations*. Amsterdam, The Netherlands: North Holland.

19. Kreyszig E. 1989 *Introductory functional analysis with applications*. New York, NY: Wiley.

20. Robinson JC. 2001 *Infinite-dimensional dynamical systems*. Cambridge, UK: Cambridge University Press.

21. Ito K. 1987 *Encyclopedic dictionary of mathematics*, 2nd edn. Cambridge, UK: MIT Press.

22. Braess D. 1997 *Finite elements*. Cambridge, UK: Cambridge University Press.

23. Oden JT, Ainsworth M. 2000 *A posteriori error estimation in finite element analysis*. New York, NY: John Wiley and Sons.

24. DeVore R, Iserles A, Süli E. 2001 *Foundations of computational mathematics*. Cambridge, UK: Cambridge University Press.

25. Stein E, Wunderlich W, Ramm E, Rank E, Rannacher R, Schweizerhof K, Wendland W, Wittum G, Wriggers P. 2003 *Error-controlled adaptive finite elements in solid mechanics*. West Sussex, UK: Wiley.

26. Krommer A, Ueberhuber C. 1994 *Numerical integration on advanced computer system*. Berlin, Germany: Springer.

27. Milne W. 1949 *Numerical calculus: approximations, interpolation, finite differences, numerical integration and curve fitting*. Princeton, NJ: Princeton university press.

28. Davis P, Rabinowitz P. 1984 *Method of numerical integration*. New York, NY: New York Academic Press.

29. Ainsworth M. 2001 Essential boundary conditions and multi-point constraints in finite element analysis. *Comp. Meth. Appl. Mech. Eng.* **190**, 6323–6339. (doi:10.1016/S0045-7825(01)00236-5)

30. Silvester P, Ferrari R. 1996 *Finite elements for electrical engineers*, 3rd edn. Cambridge, UK: Cambridge University Press.

31. Kattan P. 2008 *Matlab guide to finite elements*. Berlin, Germany: Springer.

32. Bao G, Chen Z, Wu H. 2005 Adaptive finite element method for diffraction gratings. *J. Opt. Soc. Am. A* **22**, 1106–1114. (doi:10.1364/JOSAA.22.001106)

33. Aksoylu B, Holst M. 2006 Optimality for multilevel preconditioners for local mesh refinement in three dimensions. *SIAM J. Numer. Anal.* **44**, 1005–1025. (doi:10.1137/S0036142902406119)

34. Borgioli A, Cwik T. 1998 Numerical differentiation for adaptively refined finite element meshes. In *Antennas and propagation society International Symposium, IEEE, Atlanta, GA, 21–26 June 1998*, vol. 1, pp. 256–259. New York, NY: IEEE.

35. Vukusic P, Sambles JR, Lawrence CR, Wootton RJ. 1999 Quantified interference and diffraction in single Morpho butterfly scales. *Proc. R. Soc. Lond. B* **266**, 1403–1411. (doi:10.1098/rspb.1999.0794)

36. Yoshioka S, Kinoshita S. 2004 Wavelength-selective and anisotropic light-diffusing scale on the wing of the Morpho butterfly. *Proc. R. Soc. Lond. B* **271**, 581–587. (doi:10.1098/rspb.2003.2618)

37. Vukusic P. 2006 Structural colour in Lepidoptera. *Curr. Biol.* **16**, R621– R623. (doi:10.1016/j.cub.2006.07.040)

38. Prum R, Quinn T, Torres R. 2006 Anatomically diverse butterfly scales all produce structural colours by coherent scattering. *J. Exp. Biol.* **209**, 748–765. (doi:10.1242/jeb.02051)

39. Gralak B, Tayeb G, Enoch S. 2001 Morpho butterfly wings modeled with lamellar grating theory. *Opt. Expr.* **9**, 567–578. (doi:10.1364/OE.9.000567)

40. Press W, Flannery B, Teukolsky S, Vetterling W. 1992 *Numerical recipes*. Cambridge, UK: Cambridge University Press.