IMA Journal of Applied Mathematics (2015) **80**, 508–532
doi:10.1093/imamat/hxt054
Advance Access publication on 23 January 2014

Scattering of time-harmonic electromagnetic plane waves by perfectly conducting diffraction gratings

GUANGHUI HU AND ANDREAS RATHSFELD*
Weierstrass Institute for Applied Analysis and Stochastics, Mohrenstrasse 39, 10117 Berlin, Germany
*Corresponding author: andreas.rathsfeld@wias-berlin.de

[Received on 20 June 2012; revised on 28 October 2013; accepted on 13 December 2013]

Consider the scattering of time-harmonic electromagnetic plane waves by a doubly periodic surface in $\mathbb{R}^3$. The medium above the surface is supposed to be homogeneous and isotropic with a constant dielectric coefficient, while the material below is perfectly conducting. This paper is concerned with the existence of quasiperiodic solutions for any frequency. Based on an equivalent variational formulation established by the mortar technique of Nitsche, we verify the existence of solutions for a broad class of incident waves including plane waves. The only assumption is that the grating profile is a Lipschitz biperiodic surface. Note that the solvability result of the present paper covers the resonance case where Rayleigh frequencies are allowed. Finally, non-uniqueness examples are presented in the resonance case and in the case of TE or TM polarization for classical gratings.

Keywords: electromagnetic scattering; diffraction gratings; variational approach; mortar technique; non-uniqueness.

1. Introduction

Consider a time-harmonic electromagnetic plane wave incident from above to a biperiodic surface $\tilde{\Gamma}$ in $\mathbb{R}^3$. Here a biperiodic (or doubly periodic) surface means a Lipschitz continuous surface, which is Λ_1-periodic in x_1, Λ_2-periodic in x_2 and bounded in x_3 and which divides $\mathbb{R}^3$ into two regions. The dielectric coefficient in the upper region $\tilde{\Omega}$ is supposed to be a fixed positive constant, while the medium below $\tilde{\Gamma}$ is a perfect conductor. Such structures are also called *crossed diffraction gratings* in the engineering and physics literature. This paper is concerned with a new existence result for the scattering problem valid for any fixed frequency $\omega > 0$. Note that gratings with perfectly conducting substrate materials are often used to model metallic profile gratings (cf. e.g. Turunen *et al.*, 2000; Kleemann, 2002, Sections 4.3.2 and 6.2). In particular, for infrared light, perfectly conducting boundary conditions over the interface profile yield very good approximations. These structures and more general gratings, possibly involving dielectric or non-perfectly conducting inhomogeneities or substrates, have many important applications in micro-optics and semiconductor industry (cf. e.g. Bao *et al.*, 2001; Turunen & Wyrowski, 2003).

There are many contributions on the scattering of electromagnetic waves by general inhomogeneous biperiodic diffraction gratings. First rigorous results on existence and uniqueness have been obtained by Chen & Friedman (1991) and Nédélec & Starling (1991) using integral equation methods. Abboud (1993) has introduced a variational formulation in a truncated periodic cell involving a nonlocal boundary (Dirichlet-to-Neumann) operator for a transparent boundary condition. This variational problem is of saddle point type and the existence and uniqueness follow from Fredholm's alternative. In the case of a magnetic permeability constant in $\mathbb{R}^3$, Abboud's arguments have been adapted to isotropic biperiodic inhomogeneous medium by Dobson (1994), Bao (1997), Bao & Dobson (2000), Schmidt (2004) and to anisotropic materials by Schmidt (2003). The corresponding new variational formulations involve the magnetic field only. They are strongly elliptic over the quasiperiodic Sobolev space H^1 and appear to be well adapted for the analytical and numerical treatment of quite general diffractive structures. It has been proved that there always exists a unique quasiperiodic solution of locally finite energy for all frequencies except those in a discrete set accumulating at most at infinity. Moreover, uniqueness for any frequency can be guaranteed if an absorbing (lossy) material is included into the grating (cf. Schmidt, 2003; Hu *et al.*, 2009) or if the non-absorbing (lossless) inhomogeneous material satisfies a non-trap condition (cf. Bonnet-Bendhia & Starling, 1994 in the cases of TE and TM polarizations). Schmidt (2003) additionally shows existence of solutions for plane-wave incidence even if this is not unique. In other words, even Rayleigh frequencies are admitted. Note that Rayleigh frequencies are excluded in all the other above-mentioned references. This is either due to the quasiperiodic fundamental solution of the Helmholtz equation needed in integral equation methods (Chen & Friedman, 1991; Nédélec & Starling, 1991) or to the explicit formula for the Dirichlet-to-Neumann (D-to-N) map of the transparent boundary condition (cf. e.g. Dobson, 1994; Ammari, 1995; Bao & Dobson, 2000 or (2.8)). Both, the fundamental solution and the D-to-N map, are well defined only if Rayleigh frequencies are excluded.

From the above-mentioned results, only those of Abboud (1993) and Ammari (1995) apply to scattering problems with perfectly conducting substrate. Some researchers seem to believe that the quasiperiodic solution in $H_{\mathrm{loc}}(\mathrm{curl}\,,\tilde{\Omega})$ is unique for all frequencies, provided the grating profile is given by the graph of a C^2-smooth periodic function and if Rayleigh frequencies are excluded. However, in this paper we will present a counterexample (cf. Example 2.4) showing that uniqueness does not hold in general. This counterexample is constructed in the TM polarization case, where the perfectly conducting boundary value problem of the curl –curl equation reduces to the Neumann boundary value problem of a two-dimensional scalar Helmholtz equation. This enables us to construct non-uniqueness examples for the Maxwell system, relying on the existence of non-trivial solutions for the reduced homogeneous Neumann problem in Kamotski & Nazarov (2002). Non-uniqueness examples in the resonance or non-graph case are presented as well.

Since a grating profile is a special case of a rough surface, the non-uniqueness examples reported in the present paper can also be viewed as counterexamples to the electromagnetic scattering by perfectly conducting rough surfaces. Concerning the variational approach applied to electromagnetic rough surface scattering problems modelled by the full Maxwell system, we refer to Li *et al.* (2011), where existence and uniqueness is established for an incident magnetic or electric dipole in a lossy medium, and to Haddar & Lechleiter (2011) for the more challenging case of a penetrable dielectric layer. As far as we know, the well-posedness of electromagnetic scattering by perfectly conducting rough surfaces or biperiodic structures in a homogeneous non-absorbing (lossless) medium is still an open problem.

The aim of this paper is to prove the following existence result on the scattering problem: For any fixed wavenumber $k > 0$ and for a broad class of incident waves including plane waves, there always exists a quasiperiodic solution in $H_{\mathrm{loc}}(\mathrm{curl}\,,\tilde{\Omega})$, whenever the grating profile is a Lipschitz biperiodic surface. Moreover, the far-field part of the solution reflected into the upper half space is unique. In comparison to Abboud (1993) and Ammari (1995), this result is a generalization in two directions. On the one hand, no discrete sets of wavenumbers need to be excluded. In particular, wavenumbers corresponding to Rayleigh frequencies are admitted. On the other hand, no restrictive additional condition on the surface is required, neither smallness, nor smoothness, nor any representation as the graph of a function. Note that non-graph gratings are frequently employed in diffractive optics (cf. the binary in e.g. Elschner & Schmidt, 1998).

To prove the existence of quasiperiodic solutions in $H_{\mathrm{loc}}(\mathrm{curl}\,,\tilde{\Omega})$ for any frequency, we need a replacement of the D-to-N map imposed on the artificial boundary $\tilde{\Gamma}_b$ above the grating surface.

Motivated by the variational formulations in Huber *et al.* (2009) and Rathsfeld (2011), we employ the mortar technique combined with Nitsche's method (cf. Nitsche, 1970; Sternberg, 1988). In other words, we couple the electric field E below $\tilde{\Gamma}_b$ with the Rayleigh series expansion of the scattered field E^+ above $\tilde{\Gamma}_b$. For the pair (E, E^+), we define a variational formulation which is equivalent to the boundary value problem for the time-harmonic Maxwell equation. This way the necessary transmission conditions are fulfilled on $\tilde{\Gamma}_b$ so that the coupled field is locally in $H(\mathrm{curl})$. We show the Fredholmness of the operator generated by the corresponding sesquilinear form and prove the existence of quasiperiodic solutions for any frequency. In other words, this paper provides an existence result and, additionally, a theoretical justification of the modified Nitsche method applied to electromagnetic scattering problems for periodic structures. It is expected that the arguments of this paper extend to more general inhomogeneous diffraction gratings as considered in Huber *et al.* (2009) and Rathsfeld (2011). Since the D-to-N map is not involved in the presented variational formulation, the approximation of the transparent boundary operator employed in Bao (1997) can be avoided. The technique for the proof is in many steps analogous to that for the coupling of finite elements and boundary elements (cf. e.g. Hiptmaier, 2012). In particular, the subsequent splitting of the Fourier mode space Y into the sum of two subspaces Y_1 and Y_0 corresponds to the Hodge decomposition of boundary traces. Finally, note that the presented variational approach is a basis for the numerical analysis of an FEM coupled with Fourier mode expansions (cf. Monk, 2003; Huber *et al.*, 2009; Hu & Rathsfeld, 2012).

The remaining part is organized as follows. The boundary value problem (BVP) and the Sobolev spaces are defined in Section 2. Further in this section, the main result on the existence of solutions, Theorem 2.1, is formulated and non-uniqueness examples are presented. In Section 3, we propose a variational formulation based on the method of Nitsche and prove its equivalence to (BVP). The Fredholmness of the operator generated by the corresponding sesquilinear form will be established in Section 4. Finally, applying the Fredholm alternative, we prove the main Theorem 2.1 in Section 5.

2. Mathematical formulations and non-uniqueness examples

Consider the scattering of an electromagnetic plane wave by a perfectly conducting grating profile in an isotropic homogeneous lossless medium. Recall that the symbol $\tilde{\Gamma}$ denotes the grating profile which is (Λ_1, Λ_2)-periodic in (x_1, x_2) and that $\tilde{\Omega}$ denotes the region above $\tilde{\Gamma}$. Suppose that a time-harmonic electromagnetic plane wave E^{in} (time dependence $\mathrm{e}^{-\mathrm{i}\omega t}$) given by

$$E^{\mathrm{in}} := q \exp(\mathrm{i}kx \cdot \hat{\theta}) = q \exp(\mathrm{i}(x' \cdot \alpha - \beta x_3)), \quad \mathrm{i} := \sqrt{-1} \tag{2.1}$$

is incident to the grating from above. Here $k := \omega\sqrt{\varepsilon\mu}$ is the positive wavenumber in terms of the angular frequency ω, the electric permittivity ε and the magnetic permeability μ, which are assumed to be positive constants in $\tilde{\Omega}$. The symbol $\hat{\theta}$ denotes the direction of incidence

$$\hat{\theta} := (\sin\theta_1 \cos\theta_2, \sin\theta_1 \sin\theta_2, -\cos\theta_1)^\top \in \mathbb{S}^2 := \{x = (x_1, x_2, x_3)^\top \in \mathbb{R}^3 : \|x\| = 1\}$$

with the incident angles $\theta_1 \in [0, \pi/2), \theta_2 \in [0, 2\pi)$. Throughout the paper, the symbol $(\cdot)^\top$ denotes the transpose of a row vector in $\mathbb{C}^2$ or $\mathbb{C}^3$. In (2.1), the vector $q = (q_1, q_2, q_3)^\top \in \mathbb{S}^2$ stands for the direction of polarization satisfying $q \perp \hat{\theta}$, and

$$x' := (x_1, x_2)^\top \in \mathbb{R}^2, \quad \alpha = (\alpha_1, \alpha_2)^\top := k(\sin\theta_1 \cos\theta_2, \sin\theta_1 \sin\theta_2)^\top \in \mathbb{R}^2, \quad \beta := k\cos\theta_1.$$

Since the substrate below $\tilde{\Gamma}$ is a perfect conductor, the total electric field E, which can be decomposed as the sum of the incident field E^{in} and the scattered field E^{sc}, satisfies the following boundary condition in a weak sense (cf. e.g. Buffa *et al.*, 2002)

$$\nu \times E = 0 \quad \text{on } \tilde{\Gamma}. \tag{2.2}$$

Here $\nu = (\nu_1, \nu_2, \nu_3)^\top \in \mathbb{S}^2$ is the unit normal on $\tilde{\Gamma}$ pointing into the exterior of $\tilde{\Omega}$. The total electric field E fulfills the reduced time-harmonic curl–curl equation

$$\operatorname{curl}\operatorname{curl} E - k^2 E = 0 \quad \text{in } \tilde{\Omega}. \tag{2.3}$$

Since the grating profile is biperiodic, we require E to be α-quasiperiodic in the sense that

$$\begin{aligned} E(x_1 + \Lambda_1, x_2, x_3) &= \exp(\mathrm{i}\Lambda_1\alpha_1)E(x_1, x_2, x_3), \quad (x_1, x_2, x_3)^\top \in \tilde{\Omega}, \\ E(x_1, x_2 + \Lambda_2, x_3) &= \exp(\mathrm{i}\Lambda_2\alpha_2)E(x_1, x_2, x_3), \quad (x_1, x_2, x_3)^\top \in \tilde{\Omega}. \end{aligned} \tag{2.4}$$

We also impose a radiation condition in x_3-direction assuming that the scattered field E^{sc} is composed of bounded outgoing plane waves:

$$E^{\text{sc}}(x) = \sum_{n \in \mathbb{Z}^2} E_n \exp(\mathrm{i}(\alpha_n \cdot x' + \beta_n x_3)) \quad \text{for } x_3 > \Gamma_{\max} := \max_{x \in \tilde{\Gamma}}\{x_3\}, \quad E_n \perp (\alpha_n, \beta_n)^\top, \tag{2.5}$$

where $\alpha_n := (\alpha_n^{(1)}, \alpha_n^{(2)})^\top \in \mathbb{R}^2$, with $\alpha_n^{(j)} = \alpha_j + 2\pi n_j/\Lambda_j, j = 1, 2$ for $n = (n_1, n_2)^\top \in \mathbb{Z}^2$, and

$$\beta_n = \beta_n(k, \alpha) := \begin{cases} \sqrt{k^2 - |\alpha_n|^2} & \text{if } |\alpha_n| \leqslant k, \\ \mathrm{i}\sqrt{|\alpha_n|^2 - k^2} & \text{if } |\alpha_n| > k. \end{cases}$$

For the constant coefficient vector $E_n = (E_n^{(1)}, E_n^{(2)}, E_n^{(3)})^\top \in \mathbb{C}^3$, the relation $E_n \perp (\alpha_n, \beta_n)^\top$ means that $E_n^{(1)}\alpha_n^{(1)} + E_n^{(2)}\alpha_n^{(2)} + E_n^{(3)}\beta_n = 0$. The series in (2.5), which is also referred to as the Rayleigh series expansion, is the radiation condition we are going to use in the following sections. The constant vectors E_n are called the Rayleigh coefficients. Since β_n are real-valued only for the finitely many indices n from the set $\{n \in \mathbb{Z}^2 : |\alpha_n| \leqslant k^2\}$, we observe that only a finite number of plane waves in (2.5) propagate into the far field, while the remaining part consists of evanescent (or surface) waves decaying exponentially as $x_3 \to +\infty$. Thus, the sum in (2.5) converges uniformly with all derivatives in the half plane $\{x_3 > a\}$ for any $a > \Gamma_{\max}$.

It is assumed throughout this paper that the grating profile $\tilde{\Gamma}$ is a Lipschitz biperiodic surface in $\mathbb{R}^3$, which is not necessarily the graph of a biperiodic function. Since the unbounded domain $\tilde{\Omega}$ is (Λ_1, Λ_2)-periodic in x' and the incident and scattered fields are both quasiperiodic, we can reduce the scattering problem to a single periodic cell Ω. To this end, we introduce

$$\begin{aligned} &\Gamma := \{(x_1, x_2, x_3)^\top \in \tilde{\Gamma} : 0 < x_j < \Lambda_j,\ j = 1, 2\}, \\ &\Omega := \{(x_1, x_2, x_3)^\top \in \tilde{\Omega} : 0 < x_j < \Lambda_j,\ j = 1, 2\}, \quad \tilde{\Gamma}_b := \{(x_1, x_2, x_3)^\top : x_3 = b\}, \\ &\Gamma_b := \{(x_1, x_2, x_3)^\top \in \tilde{\Gamma}_b : 0 < x_j < \Lambda_j,\ j = 1, 2\}, \quad \Omega_b := \{x \in \Omega : x_3 < b\} \end{aligned}$$

for some $b > \Gamma_{\max}$. We next introduce some scalar and vector valued α-quasiperiodic Sobolev spaces. Let $H^s(\tilde{\Gamma}_b)$ be the complex valued L^2-based Sobolev spaces of order s in $\tilde{\Gamma}_b$. Write

$$\begin{aligned}
H_{\mathrm{loc}}(\mathrm{curl}\,,\tilde{\Omega}) &:= \{G : \chi G, \mathrm{curl}\,(\chi G) \in L^2(\tilde{\Omega})^3,\ \forall \chi \in C_0^\infty(\mathbb{R}^3)\},\\
H^s_{\mathrm{loc}}(\tilde{\Gamma}_b) &:= \{G : \chi G \in H^s(\tilde{\Gamma}_b),\ \forall \chi \in C_0^\infty(\tilde{\Gamma}_b)\},\\
H^s_{t,\mathrm{loc}}(\tilde{\Gamma}_b) &:= \{G \in H^s_{\mathrm{loc}}(\tilde{\Gamma}_b)^3 : e_3 \cdot G = 0\},\ e_3 := (0,0,1)^\top,\\
H^s_{t,\mathrm{loc}}(\mathrm{Div}\,,\tilde{\Gamma}_b) &:= \{G : G \in H^s_{t,\mathrm{loc}}(\tilde{\Gamma}_b),\ \mathrm{Div}\, G \in H^s_{\mathrm{loc}}(\tilde{\Gamma}_b)\},\\
H^s_{t,\mathrm{loc}}(\mathrm{Curl}\,,\tilde{\Gamma}_b) &:= \{G : G \in H^s_{t,\mathrm{loc}}(\tilde{\Gamma}_b),\ \mathrm{Curl}\, G \in H^s_{\mathrm{loc}}(\tilde{\Gamma}_b)\},\\
H(\mathrm{curl}\,,\Omega_b) &:= \{G|_{\Omega_b} : G \in H_{\mathrm{loc}}(\mathrm{curl}\,,\tilde{\Omega}), G \text{ is } \alpha\text{-quasiperiodic}\},\\
H^s_t(\Gamma_b) &:= \{G|_{\Gamma_b} : G \in H^s_{t,\mathrm{loc}}(\tilde{\Gamma}_b), G \text{ is } \alpha\text{-quasiperiodic}\},\\
H^s_t(\mathrm{Div}\,,\Gamma_b) &:= \{G|_{\Gamma_b} : G \in H^s_{t,\mathrm{loc}}(\mathrm{Div}\,,\tilde{\Gamma}_b), G \text{ is } \alpha\text{-quasiperiodic}\},\\
H^s_t(\mathrm{Curl}\,,\Gamma_b) &:= \{G|_{\Gamma_b} : G \in H^s_{t,\mathrm{loc}}(\mathrm{Curl}\,,\tilde{\Gamma}_b), G \text{ is } \alpha\text{-quasiperiodic}\},
\end{aligned}$$

where $\mathrm{Div}\,(\cdot)$ and $\mathrm{Curl}\,(\cdot)$ stand for the surface divergence and the surface scalar curl operators, respectively. Clearly, each $E(x') \in H^s_t(\Gamma_b)$, $s \in \mathbb{R}$ admits the Fourier series expansion

$$E(x') = \sum_{n\in\mathbb{Z}^2} E_n \exp(\mathrm{i}\alpha_n \cdot x'), \quad E_n := (\Lambda_1\Lambda_2)^{-1} \int_0^{\Lambda_1}\int_0^{\Lambda_2} E(x') \exp(-\mathrm{i}\alpha_n \cdot x')\,\mathrm{d}x_1\,\mathrm{d}x_2.$$

Then, the spaces $H^s_t(\Gamma_b)$, $H^s_t(\mathrm{Div}\,,\Gamma_b)$ and $H^s_t(\mathrm{Curl}\,,\Gamma_b)$ can be equipped with the equivalent Sobolev norms $\|E\|_{H^s_t(\Gamma_b)} = (\sum_{n\in\mathbb{Z}^2} |E_n|^2(1+|\alpha_n|^2)^s)^{1/2}$ and

$$\|E\|_{H^s_t(\mathrm{Div}\,,\Gamma_b)} = \left(\sum_{n\in\mathbb{Z}^2}(1+|\alpha_n|^2)^s(|E_n|^2+|E_n\cdot\alpha_n|^2)\right)^{1/2},$$
$$\|E\|_{H^s_t(\mathrm{Curl}\,,\Gamma_b)} = \left(\sum_{n\in\mathbb{Z}^2}(1+|\alpha_n|^2)^s(|E_n|^2+|E_n\times\alpha_n|^2)\right)^{1/2},$$

respectively. Recall that the space dual to $H^s_t(\mathrm{Div}\,,\Gamma_b)$ w.r.t. the L^2-scalar product is $H^s_t(\mathrm{Div}\,,\Gamma_b)' = H^{-s-1}_t(\mathrm{Curl}\,,\Gamma_b)$, and that, for $s = -\frac{1}{2}$,

$$\begin{aligned}
H^{-1/2}_t(\mathrm{Div}\,,\Gamma_b) &= \{e_3 \times E|_{\Gamma_b} : E \in H(\mathrm{curl}\,,\Omega_b)\},\\
H^{-1/2}_t(\mathrm{Curl}\,,\Gamma_b) &= \{(e_3 \times E|_{\Gamma_b}) \times e_3 : E \in H(\mathrm{curl}\,,\Omega_b)\}.
\end{aligned}$$

Further, the corresponding trace mappings from $H(\mathrm{curl}\,,\Omega_b)$ to the tangential spaces $H^{-1/2}_t(\mathrm{Div}\,,\Gamma_b)$ and $H^{-1/2}_t(\mathrm{Curl}\,,\Gamma_b)$ are continuous and surjective (cf. Buffa *et al.*, 2002; Monk, 2003 and the references there). Finally, define the variational space by $X = X_b := \{E : \Omega_b \to \mathbb{C}^3 : E \in H(\mathrm{curl}\,,\Omega_b), \nu \times E|_\Gamma = 0\}$ endowed with the norm $\|E\|_X := \|E\|_{H(\mathrm{curl}\,,\Omega_b)} = (\|E\|^2_{L^2(\Omega_b)^3} + \|\mathrm{curl}\,E\|^2_{L^2(\Omega_b)^3})^{1/2}$.

The boundary value problem for our scattering problem can be stated as follows. Let the grating profile Γ and the number $b > \Gamma_{\max}$ be fixed.

(BVP): Given an incident electric field E^{in}, determine the total field $E = E^{\mathrm{in}} + E^{\mathrm{sc}} \in X$ such that E satisfies the curl – curl equation (2.3) over Ω_b in a distributional sense and that E^{sc} admits a Rayleigh expansion (2.5) valid for any $\Gamma_{\max} < x_3 \leqslant b$.

Note that any E^{sc} satisfying (2.5) in the strip $\Gamma_{\max} < x_3 \leqslant b$ can be extended to the upper half space by (2.5). Below is our main result to (BVP) for a broad class of incident waves.

THEOREM 2.1 Assume that the incident electric wave takes the form:

$$E^{\mathrm{in}}_{\mathrm{gen}} := \sum_{n:\beta_n>0} Q_n \exp(\alpha_n \cdot x' - \beta_n x_3), \tag{2.6}$$

where $Q_n \in \mathbb{C}^3$ satisfies $Q_n \perp (\alpha_n, -\beta_n)^\top$. Then the problem (BVP) admits at least one solution for any $k \in \mathbb{R}^+$. Moreover, the part of the solution reflected into the upper half space is unique, i.e., the Rayleigh coefficients of the plane wave modes propagating into the upper half space (namely, those with $\beta_n > 0$) are unique.

Clearly, E^{in} of (2.1) is of the form (2.6), where $Q_n = q$ for $n = (0,0)^\top$ and $Q_n = (0,0,0)^\top$ else. We do not exclude 'resonances' in Theorem 2.1, i.e., the set $\Upsilon := \{n \in \mathbb{Z}^2 : \beta_n(k,\alpha) = 0\}$ can be nonempty. An incident angular frequency ω with $\Upsilon \neq \emptyset$ is called Rayleigh frequency. Note that the set of Rayleigh frequencies depends on Λ_1 and Λ_2 but not on the shape of Γ.

REMARK 2.1 It seems to be known that, for all wavenumbers except those from a sequence $k_j \in \mathbb{R}^+$, $k_j \to +\infty$, the problem (BVP) admits a unique solution. To see this, consider the variational formulation

$$\begin{aligned}\int_{\Omega_b} [\operatorname{curl} E \cdot \operatorname{curl} \overline{\varphi} - k^2 E \cdot \overline{\varphi}]\, \mathrm{d}x - \int_{\Gamma_b} \mathscr{R}(e_3 \times E) \cdot (e_3 \times \overline{\varphi})\, \mathrm{d}s \\ = \int_{\Gamma_b} [(\operatorname{curl} E^{\mathrm{in}})_T - \mathscr{R}(e_3 \times E^{\mathrm{in}})] \cdot (e_3 \times \overline{\varphi})\, \mathrm{d}s\end{aligned} \tag{2.7}$$

for all $\varphi \in X$, where $(\cdot)_T := \nu \times (\cdot)|_{\Gamma_b} \times \nu$, and $\mathscr{R} : H_t^{-1/2}(\mathrm{Div}\,, \Gamma_b) \to H_t^{-1/2}(\mathrm{Curl}\,, \Gamma_b)$ is the Dirichlet-to-Neumann map defined by

$$(\mathscr{R}\tilde{E})(x') = -\sum_{n\in\mathbb{Z}^2} \frac{1}{\mathrm{i}\beta_n}[k^2 \tilde{E}_n - (\alpha_n \cdot \tilde{E}_n)\alpha_n] \exp(\mathrm{i}\alpha_n \cdot x'), \tag{2.8}$$

for $\tilde{E}(x') = \sum_{n\in\mathbb{Z}^2} \tilde{E}_n \exp(\mathrm{i}\alpha_n \cdot x') \in H_t^{-1/2}(\mathrm{Div}\,, \Gamma_b)$; cf. Abboud (1993) and Ammari (1995). Note that the operator $\mathscr{R}$ maps $(e_3 \times E^{\mathrm{sc}})|_{\Gamma_b}$ to $(\operatorname{curl} E^{\mathrm{sc}}|_{\Gamma_b})_T$ and that Rayleigh frequencies must be excluded in (2.8). An alternative Dirichlet-to-Neumann operator for the magnetic field is given in Bao & Dobson (2000), Bao (1997) and Dobson (1994).

It is seen from Lemma A.1 that the variational formulation is uniquely solvable for all frequencies $k \in (0, k_0]$ with $k_0 > 0$ being sufficiently small. This combined with the analytic Fredholm theory (cf. e.g. Colton & Kress, 1998, Theorem 8.26 or Gohberg & Krein, 1969, Theorem I.5.1) leads to the existence and uniqueness for all $k \in \mathbb{R}^+ \backslash \mathscr{D}$, where $\mathscr{D} \supseteq \Upsilon$ is a discrete set with the only accumulating point at infinity. Since such a solvability result is contained in many references on diffraction gratings, we skip the details and refer to Schmidt (2003), Elschner & Hu (2010) and Elschner & Schmidt (1998) for the applications of the analytic Fredholm theory in periodic structures. The exceptional set of wavenumbers $\mathscr{D}$ cannot be avoided. Indeed, the examples from below show that uniqueness for (BVP) does not hold in general, even if Γ is a smooth graph and Rayleigh frequencies are excluded.

The proof of Theorem 2.1 will be carried out in Section 5 using an equivalent formulation which covers the resonance case. Next we present some non-uniqueness examples to (BVP) by constructing non-trivial solutions to the homogeneous scattering problem ($E^{\mathrm{in}} = 0$). Suppose that the periodicities Λ_1 and Λ_2 are fixed for the remainder of this paper.

EXAMPLE 2.2 For any fixed Rayleigh frequency ω, there exists a biperiodic surface $\tilde{\Gamma}$ such that the solutions to (BVP) are non-unique. Indeed, the grating profile defined by $\tilde{\Gamma} := \{x_3 = 0\}$ is such an example. Defining the electric field $E^{\mathrm{sc}}(x) = e_3 \sum_{n:\beta_n=0} C_n \exp(\mathrm{i}\alpha_n \cdot x')$ with $C_n \in \mathbb{C}$, the α-quasiperiodic field E^{sc} satisfies the curl − curl equation (2.3), the Rayleigh expansion condition (2.5) and the boundary condition (2.2).

In the following examples, the branch of the square root is chosen such that its imaginary part is non-negative, i.e., $\sqrt{a} = \mathrm{i}\sqrt{-a}$ if $a < 0$.

EXAMPLE 2.3 There exists a non-Rayleigh frequency ω and a non-graph grating profile $\tilde{\Gamma}$ such that the solutions to (BVP) are non-unique. Restrict the search for examples to gratings which remain invariant in x_2-direction. We seek a special solution of the form $E^{\mathrm{sc}}(x) = (0, u^{\mathrm{sc}}(x_1, x_3), 0)^\top$, where the scalar function u^{sc} fulfills

$$
\begin{aligned}
&(\triangle + k^2)u^{\mathrm{sc}} = 0 \quad \text{in } \tilde{\Omega}_0 := \tilde{\Omega} \cap \{x_2 = 0\}, \quad u^{\mathrm{sc}} = 0 \quad \text{on } \tilde{\Gamma} \cap \{x_2 = 0\},\\
&u^{\mathrm{sc}} = \sum_{n_1 \in \mathbb{Z}, n_2 = 0} C_{n_1} \exp(\mathrm{i}[\alpha_n^{(1)} x_1 + (k^2 - |\alpha_n^{(1)}|^2)^{1/2} x_3]), \quad (x_1, x_3)^\top \in \tilde{\Omega}_0, \ x_3 > \Gamma_{\max},
\end{aligned}
$$

with $C_{n_1} \in \mathbb{C}$. Recall that $n = (n_1, n_2)^\top \in \mathbb{Z}^2$ and $\alpha_n^{(j)}$ denotes the jth component of $\alpha_n \in \mathbb{R}^2$, $j = 1, 2$. In fact, the previous Dirichlet boundary value problem is the TE polarization of (BVP). Non-trivial solutions to the above problem do exist for the Λ_1-periodic non-graph grating profile constructed in Gotlib (2000) with $\Lambda_1 = 2\pi$. Thus, the solution E^{sc}, which is independent of x_2 and transversal to the (x_1, x_3)-plane, is an α-quasiperiodic solution to the homogeneous scattering problem (BVP) with $\alpha = (\alpha_1, 0)$.

EXAMPLE 2.4 There exists a non-Rayleigh frequency ω and a grating $\tilde{\Gamma}$ represented as the graph of a profile function such that the solutions to (BVP) are non-unique. Again restrict to gratings invariant in x_2-direction and consider gratings such that $\tilde{\Gamma} \cap \{x_2 = 0\}$ can be represented as a smooth function $x_3 = f(x_1)$ of period $\Lambda_1 = 2\pi$. We seek a special magnetic field H^{sc} of the form $H^{\mathrm{sc}}(x) = 1/(\mathrm{i}k)\,\mathrm{curl}\,E^{\mathrm{sc}}(x) = (0, u^{\mathrm{sc}}(x_1, x_3), 0)^\top$. Since H^{sc} should satisfy the curl − curl equation (2.3) in $\tilde{\Omega}$ and the boundary condition $\nu \times \mathrm{curl}\,H^{\mathrm{sc}} = 0$ on $\tilde{\Gamma}$, we only need to find a non-trivial scalar function u^{sc} such that

$$
\begin{aligned}
&(\triangle + k^2)u^{\mathrm{sc}} = 0 \quad \text{in } x_3 > f(x_1), \quad \frac{\partial u^{\mathrm{sc}}}{\partial \mathbf{n}} = 0 \quad \text{on } x_3 = f(x_1),\\
&u^{\mathrm{sc}} = \sum_{n_1 \in \mathbb{Z}, n_2 = 0} C_{n_1} \exp(\mathrm{i}[\alpha_n^{(1)} x_1 + (k^2 - |\alpha_n^{(1)}|^2)^{1/2} x_3]), \quad C_{n_1} \in \mathbb{C}, \ x_3 > \max_{x_1} f(x_1),
\end{aligned}
$$

where $\mathbf{n} \in \mathbb{R}^2$ denotes the normal to the one-dimensional curve $x_3 = f(x_1)$ in the (x_1, x_3)-plane. This case is just the TM polarization of (BVP). It follows from Kamotski & Nazarov (2002) that exponentially decaying solutions (surface waves) to the above Neumann boundary value problem exist for a broad class of grating profiles that are given by the graphs of smooth functions. Thus, we obtain a TM polarized solution H^{sc}, which is transversal to the (x_1, x_3)-plane, and a non-trivial solution $E^{sc}(x) = -1/(ik)\,\mathrm{curl}\, H^{sc}(x)$ to the homogeneous problem of (BVP) given by

$$E^{sc}(x) = \frac{1}{k} \sum_{n_1 \in \mathbb{Z}, n_2 = 0} C_{n_1} (\sqrt{k^2 - |\alpha_n^{(1)}|^2}, 0, -\alpha_n^{(1)})^\top \exp(i[\alpha_n^{(1)} x_1 + \sqrt{k^2 - |\alpha_n^{(1)}|^2} x_3]).$$

Note that the last two examples in the non-resonance case are obtained only if the grating surface $\tilde{\Gamma}$ remains constant in x_2-direction. Similar non-trivial solutions can be constructed for biperiodic structures only varying in x_1-direction. However, we do not have a corresponding example for the diffraction gratings that vary in two orthogonal directions. It remains an interesting question under what kind of geometric conditions imposed on $\tilde{\Gamma}$ the uniqueness of (BVP) holds. Although there is no uniqueness in the general case, we can prove the existence of solutions to (BVP) for any wavenumber $k \in \mathbb{R}^+$. This will be done in the subsequent sections.

3. An equivalent variational formulation

The goal of this section is to propose a variational formulation equivalent to (BVP). We begin with the fact that any column vector $E_n \in \mathbb{C}^3$ satisfying $(\alpha_n, \beta_n)^\top \perp E_n$ for some $n = (n_1, n_2)^\top \in \mathbb{Z}^2$ can be represented as a linear combination $E_n = C_{n,0} E_{n,0} + C_{n,1} E_{n,1}$, $C_{n,0}, C_{n,1} \in \mathbb{C}$ of the vectors $E_{n,1}, E_{n,2} \in \mathbb{C}^3$, where

$$E_{n,0} := \begin{cases} (-\alpha_n^{(2)}, \alpha_n^{(1)}, 0)^\top / |\alpha_n| \in \mathbb{S}^2 & \text{if } |\alpha_n| \neq 0, \\ (0, 1, 0)^\top & \text{else}, \end{cases} \tag{3.1}$$

$$E_{n,1} := \begin{cases} \dfrac{|\alpha_n|}{h_n} (\alpha_n, \beta_n)^\top \times E_{n,0} = (-\alpha_n^{(1)} \beta_n, -\alpha_n^{(2)} \beta_n, |\alpha_n|^2)^\top / h_n & \text{if } |\alpha_n| \neq 0, \\ (-1, 0, 0)^\top & \text{else}, \end{cases} \tag{3.2}$$

with $h_n := |\alpha_n| \sqrt{|\alpha_n|^2 + |\beta_n|^2}$. Obviously, we have $(\alpha_n, \beta_n)^\top \perp E_{n,l}$, $|E_{n,l}| = 1$ for $l = 0, 1$ and $n \in \mathbb{Z}^2$. We further observe that $E_{n,1} \in \mathbb{S}^2$ if $\beta_n \in \mathbb{R}$, and that $E_{n,1} = e_3$ if $\beta_n = 0$. The above decomposition of E_n allows us to rewrite the Rayleigh expansion (2.5) as

$$E^{sc}(x) = \sum_{n \in \mathbb{Z}^2, l = 1,2} C_{n,l} U_{n,l}(x), \quad U_{n,l} := E_{n,l} \exp(i[\alpha_n \cdot x' + \beta_n x_3]), \quad C_{n,l} \in \mathbb{C} \tag{3.3}$$

for $x_3 > \Gamma_{\max}$ (cf. also Rathsfeld, 2011, Section 2.5). Define the layer Ω_b^+ above Γ_b of height one by $\Omega_b^+ := \{x \in \mathbb{R}^3 : 0 < x_j < \Lambda_j, j = 1, 2, b < x_3 < b + 1\}$ (cf. Fig. 1), and spaces Y_l by

$$Y_l := \{U \in H(\mathrm{curl}\,, \Omega_b^+) : U(x) = \sum_{n \in \mathbb{Z}^2} C_{n,l} U_{n,l}(x), C_{n,l} \in \mathbb{C}\}, \quad l = 0, 1.$$

The spaces Y_0, Y_1 and $Y := Y_0 \oplus Y_1$ are closed subspaces of the Sobolev space $H(\mathrm{curl}\,, \Omega_b^+)$ (cf. the subsequent orthogonality relations (4.7) and (4.8)). Moreover, the splitting $Y = Y_0 \oplus Y_1$ is the Hodge

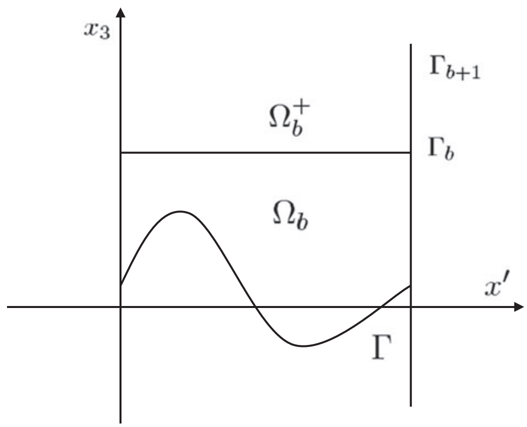

FIG. 1. The geometry of the scattering problem.

decomposition of Y since the traces $E_{n,1}|_{\Gamma_b}$ of the function $E_{n,1}$ are surface gradients and since the surface divergences of the $E_{n,0}|_{\Gamma_b}$ vanish (cf. the definitions (3.1) and (3.2)). Then we see that the function $E^+(x) := E^{\mathrm{sc}}|_{\Omega_b^+}$ belongs to the space Y, and any function in Y can be analytically extended to the whole half-space $\{x_3 > \Gamma_{\max}\}$. Hence, the following problem is equivalent to (BVP):

(BVP′): Given an incident electric field E^{in}, find $(E, E^+) \in \mathbb{H} := X \times Y$ such that E satisfies the curl–curl equation (2.3) in a distributional sense and the transmission conditions

$$e_3 \times (E - E^{\mathrm{in}} - E^+) = 0, \quad e_3 \times \mathrm{curl}\,(E - E^{\mathrm{in}} - E^+) = 0 \quad \text{on } \Gamma_b. \tag{3.4}$$

Motivated by the arguments in Rathsfeld (2011, Section 3.2) and the variational formulation in Huber *et al.* (2009), we propose a new variational formulation that is equivalent to (BVP′). For pairs of fields $(E, E^+), (V, V^+) \in \mathbb{H}$, define the sesquilinear form $a(\cdot,\cdot)$:

$$\begin{aligned} a((E,E^+),(V,V^+)) :=& \int_{\Omega_b} \{\mathrm{curl}\,E \cdot \mathrm{curl}\,\overline{V} - k^2 E \cdot \overline{V}\}\,\mathrm{d}x - \int_{\Gamma_b} \mathrm{curl}\,E^+ \cdot e_3 \times \overline{V}\,\mathrm{d}s \\ &+ \int_{\Gamma_b} e_3 \times (E - E^+) \cdot \mathrm{curl}\,\overline{V}^+\,\mathrm{d}s + \eta \sum_{n\in\Upsilon} \int_{\Gamma_b} e_3 \times (E - E^+) \cdot (e_3 \times \overline{U}_{n,0})\,\mathrm{d}s \\ &\times \overline{\int_{\Gamma_b} e_3 \times (V - V^+) \cdot (e_3 \times \overline{U}_{n,0})\,\mathrm{d}s}, \end{aligned} \tag{3.5}$$

where $\eta > 0$ is a constant factor for mortaring and is normally chosen as a multiple of the reciprocal mesh size (cf. Huber *et al.*, 2009). Our variational formulation is to find $(E, E^+) \in \mathbb{H}$ such that

$$a((E,E^+),(V,V^+)) = -a((0,E^{\mathrm{in}}),(V,V^+)) \quad \forall (V,V^+) \in \mathbb{H}. \tag{3.6}$$

Note that terms like $\int_{\Gamma_b} \mathrm{curl}\,E^+ \cdot e_3 \times \overline{V}\,\mathrm{d}s$ are bounded. Indeed, since E^+ is the solution of the curl–curl equation, we obtain $\mathrm{curl}\,E^+ \in H(\mathrm{curl}\,,\Omega_b^+)$ and $(\mathrm{curl}\,E^+)|_{\Gamma_b} \in H^{-1/2}(\mathrm{Curl}\,,\Gamma_b)$. Further, note that the third term on the right-hand side of (3.5) has the opposite sign than the corresponding term in Huber *et al.* (2009). Moreover, the integrals with factor η in (3.5) correspond to the following term involved in the variational equation established in Huber *et al.* (2009):

$$\eta \int_{\Gamma_b} e_3 \times (E - E^+) \cdot e_3 \times (\overline{V - V^+})\, \mathrm{d}s. \tag{3.7}$$

The expression (3.7) is not meaningful for general $(E, E^+), (V, V^+) \in \mathbb{H}$, since both $e_3 \times (E - E^+)$ and $e_3 \times (\overline{V - V^+})$ belong to $H_t^{-1/2}(\mathrm{Div}\,, \Gamma_b)$. Integrals like $\eta \int_{\Gamma_b} e_3 \times u \cdot e_3 \times \overline{v}\, \mathrm{d}s$ in the mortar approach make sense for finite element methods, where u and v are finite element functions and η tends to zero with the meshsize. The idea employed in Rathsfeld (2011) is to replace the integral (3.7) by the Galerkin approximation

$$\sum_{\substack{n,l:|n|^2<N \\ \beta_n \neq 0 \text{ or } l=0}} \eta \int_{\Gamma_b} e_3 \times (E - E^+) \cdot e_3 \times \overline{U}_{n,l}\, \mathrm{d}s \overline{\int_{\Gamma_b} e_3 \times (V - V^+) \cdot e_3 \times \overline{U}_{n,l}\, \mathrm{d}s} \tag{3.8}$$

$$+ \eta \sum_{n:\beta_n=0} \int_{\Gamma_b} e_3 \times (E - E^+) \cdot \overline{U}_{n,0}\, \mathrm{d}s \overline{\int_{\Gamma_b} e_3 \times (V - V^+) \cdot \overline{U}_{n,0}\, \mathrm{d}s}, \tag{3.9}$$

with a sufficiently large number $N > 0$. It is also mentioned in Rathsfeld (2011) that the summation in (3.8) and (3.9) can even be restricted to all $n \in \mathbb{Z}^2$ with $\beta_n = 0$. In the present paper, we only use the terms of (3.8) with $\beta_n = 0$, which are the last terms in (3.5).

To prove the equivalence of (3.6) and (BVP′), we need two lemmas.

LEMMA 3.1 (i) We have $\mathrm{curl}\, U_{n,l} = \mathrm{i}(-1)^l U_{n,1-l} \sqrt{|\alpha_n|^2 + |\beta_n|^2}^{\,1-2l} k^{2l}$, $l = 0, 1$.

(ii) Setting $\cos\theta_n := \beta_n / \sqrt{|\beta_n|^2 + |\alpha_n|^2}$, there holds

$$e_3 \times U_{n,l}|_{\Gamma_b} = \begin{cases} (-\alpha_n/|\alpha_n|, 0)^\top \exp(\mathrm{i}\alpha_n \cdot x') & \text{if } n \in \Upsilon,\ l = 0, \\ (0, 0, 0)^\top & \text{if } n \in \Upsilon,\ l = 1, \\ (-1)^l[(e_3 \times U_{n,1-l}) \times e_3](\cos\theta_n)^{2l-1} & \text{if } n \notin \Upsilon. \end{cases}$$

(iii) The following set is an L^2-orthogonal basis of the space $H_t^{-1/2}(\Gamma_b)$:

$$\{e_3 \times U_{n,l}|_{\Gamma_b} : n \notin \Upsilon, l = 0, 1\} \cup \{e_3 \times U_{n,0}|_{\Gamma_b} : n \in \Upsilon\} \cup \{U_{n,0}|_{\Gamma_b} : n \in \Upsilon\}.$$

Proof. Lemma 3.1(i) and (ii) can be proved directly using the definitions of $U_{n,l}$ in (3.3). To prove the third assertion, we define the set $\Pi_n := \{e_3 \times E_{n,0}, e_3 \times E_{n,1}\}$ if $\beta_n \neq 0$ and $\Pi_n := \{e_3 \times E_{n,0}, E_{n,0}\}$ if $\beta_n = 0$ with $E_{n,l} \in \mathbb{C}^3$ given in (3.1) and (3.2). Then Lemma 3.1(iii) simply follows from the definition of $H_t^{-1/2}(\Gamma_b)$ and the fact that Π_n spans the set $\{(a_1, a_2, 0)^\top : a_1, a_2 \in \mathbb{C}\}$ for any $n \in \mathbb{Z}^2$. □

In the following sections we make the convention that any sum over l is the sum over $l = 0, 1$.

LEMMA 3.2 For any two absolutely convergent Rayleigh series expansion U and V defined in a neighborhood of Γ_b, there holds

$$\int_{\Gamma_b} (\mathrm{curl}\, U)_T \cdot e_3 \times \overline{V}\, \mathrm{d}s = \int_{\Gamma_b} e_3 \times U \cdot (\mathrm{curl}\, [\overline{V}]_{\mathrm{mo}})_T\, \mathrm{d}s,$$

where $[\cdot]_{\mathrm{mo}}$ is a modification operator defined by

$$\left[\sum_{n\in\mathbb{Z}^2,l} C_{n,l}U_{n,l}\right]_{\mathrm{mo}} := -\sum_{l,n:\beta_n>0} C_{n,l}U_{n,l} + \sum_{l,n:\beta_n\notin\mathbb{R}} C_{n,l}U_{n,l}.$$

Proof. See Rathsfeld (2011, Lemma 3.1). □

We are now going to prove:

LEMMA 3.3 The variational formulation (3.6) and the problem (BVP′) are equivalent.

Proof. (i) Assume that $(E, E^+) \in \mathbb{H}$ is a solution to (BVP′). Applying Green's first vector theorem to the region Ω_b gives

$$0 = \int_{\Omega_b} \{\operatorname{curl}\operatorname{curl} E - k^2 E\} \cdot \overline{V}\,\mathrm{d}x = \int_{\Omega_b} \{\operatorname{curl} E \cdot \operatorname{curl} \overline{V} - k^2 E \cdot \overline{V}\}\,\mathrm{d}x - \int_{\Gamma_b} e_3 \times \overline{V} \cdot \operatorname{curl} E\,\mathrm{d}s$$

for any $V \in X$. Note that the integral over Γ vanishes due to the perfectly conducting boundary condition $\nu \times V = 0$ on Γ, and that the integrals over the vertical parts of $\partial\Omega_b$ cancel because of the α-quasiperiodicity of V and E in Ω_b. This implies that

$$\int_{\Omega_b} \{\operatorname{curl} E \cdot \operatorname{curl} \overline{V} - k^2 E \cdot \overline{V}\}\,\mathrm{d}x = \int_{\Gamma_b} e_3 \times \overline{V} \cdot \operatorname{curl} E\,\mathrm{d}s, \quad \forall V \in X. \tag{3.10}$$

Making use of the identity (3.10) and the transmission conditions in (3.4), we derive from the definition of the sesquilinear form $a(\cdot,\cdot)$ that

$$\begin{aligned} a((E, E^+ + E^{\mathrm{in}}), (V, V^+)) &= \int_{\Gamma_b} \operatorname{curl}(E - E^+ - E^{\mathrm{in}}) \cdot e_3 \times \overline{V}\,\mathrm{d}s + \int_{\Gamma_b} e_3 \times (E - E^+ - E^{\mathrm{in}}) \cdot \operatorname{curl} \overline{V}^+\,\mathrm{d}s \\ &\quad + \eta \sum_{n\in\Upsilon} \int_{\Gamma_b} e_3 \times (E - E^+ - E^{\mathrm{in}}) \cdot (e_3 \times \overline{U}_{n,0})\,\mathrm{d}s \\ &\quad \times \overline{\int_{\Gamma_b} e_3 \times (V - V^+) \cdot (e_3 \times \overline{U}_{n,0})\,\mathrm{d}s} = 0 \end{aligned} \tag{3.11}$$

for any $(V, V^+) \in \mathbb{H}$, i.e., the pair (E, E^+) is a solution to (3.6).

(ii) Suppose that $(E, E^+) \in \mathbb{H}$ is a solution to (3.6). Choose $V \in X$ with a compact support in the interior of Ω_b (i.e. $\mathrm{Supp}(V) \subset \mathrm{Int}\,\Omega_{\mathrm{b}}$) and choose $V^+ \equiv 0$ in Y. Then,

$$0 = a((E, E^+ + E^{\mathrm{in}}), (V, 0)) = \int_{\Omega_b} \{\operatorname{curl} E \cdot \operatorname{curl} \overline{V} - k^2 E \cdot \overline{V}\}\,\mathrm{d}x = \int_{\Omega_b} (\operatorname{curl}\operatorname{curl} E - k^2 E) \cdot \overline{V}\,\mathrm{d}x. \tag{3.12}$$

This implies that $\operatorname{curl}\operatorname{curl} E - k^2 E = 0$ in Ω_b. It remains to prove that only E and E^+ satisfy the transmission conditions in (3.4).

Analogously to part (i), multiplying $V \in X$ to the curl–curl equation $\operatorname{curl}\operatorname{curl} E - k^2 E = 0$ in Ω_b and then using integration by parts yields the identity (3.10). Combining this identity with the variational formulation (3.6) gives rise to the equality (3.11) for all $(V, V^+) \in \mathbb{H}$. By Lemma 3.1(ii) and (iii), we consider the Fourier expansion

$$(E - E^+ - E^{\mathrm{in}})_T = \sum_{l,n \notin \Upsilon} C_{n,l} (U_{n,l})_T + \sum_{n \in \Upsilon} [C_{n,0} U_{n,0} + C_{n,1} e_3 \times U_{n,0}] \quad \text{on } \Gamma_b. \tag{3.13}$$

It then suffices to prove that $C_{n,l} = 0$ for all $n \in \mathbb{Z}^2, l = 0, 1$. Indeed, $(E - E^+ - E^{\mathrm{in}})_T = 0$ on Γ_b together with (3.11) for all $V \in X$ would lead to $(\mathrm{curl}\,(E - E^+ - E^{\mathrm{in}})|_{\Gamma_b})_T = 0$.

First we take $V \equiv 0$ and $V^+ = U_{n,1}$ for some $n \in \Upsilon$ in (3.11). Applying Lemma 3.1(i) to $U_{n,1}$ gives the identity $\mathrm{curl}\, V^+ = -\mathrm{i}k U_{n,0}$, and then, using $e_3 \times U_{n,1} = 0$ for $n \in \Upsilon$ (cf. Lemma 3.1(ii)), we derive from (3.11) that

$$\int_{\Gamma_b} e_3 \times (E - E^+ - E^{\mathrm{in}}) \cdot \overline{U}_{n,0} \, \mathrm{d}s = 0 \quad \text{if } n \in \Upsilon. \tag{3.14}$$

Together with (3.13), this implies that $C_{n,1} = 0$ for $n \in \Upsilon$.

Next, inserting (3.14) into (3.11) with $V \equiv 0$ and using Lemma 3.2, we have

$$\begin{aligned} 0 &= \int_{\Gamma_b} e_3 \times (E - E^+ - E^{\mathrm{in}}) \cdot \mathrm{curl}\, \overline{V}^+ \, \mathrm{d}s - \eta \sum_{n \in \Upsilon} \int_{\Gamma_b} (E - E^+ - E^{\mathrm{in}})_T \cdot \overline{U}_{n,0} \, \mathrm{d}s \overline{\int_{\Gamma_b} V_T^+ \cdot \overline{U}_{n,0} \, \mathrm{d}s} \\ &= \int_{\Gamma_b} \mathrm{curl}\, [(E - E^+ - E^{\mathrm{in}}|_{\Gamma_b})]_{\mathrm{mo}} \cdot e_3 \times (\overline{V}^+) \, \mathrm{d}s - \eta \sum_{n \in \Upsilon} \int_{\Gamma_b} (E - E^+ - E^{\mathrm{in}})_T \cdot \overline{U}_{n,0} \, \mathrm{d}s \overline{\int_{\Gamma_b} V_T^+ \cdot \overline{U}_{n,0} \, \mathrm{d}s} \end{aligned} \tag{3.15}$$

for all $V^+ \in Y$, where the quantity

$$[(E - E^+ - E^{\mathrm{in}})|_{\Gamma_b}]_{\mathrm{mo}} := -\sum_{l,n:\beta_n > 0} C_{n,l} U_{n,l} + \sum_{l,n:\beta_n \notin \mathbb{R}} C_{n,l} U_{n,l} \quad \text{on } \Gamma_b$$

is obtained by firstly extending the series expansion (3.13) to a neighborhood of Γ_b and then applying the modification operator $[\cdot]_{\mathrm{mo}}$ of Lemma 3.2. From Lemma 3.1(i) and (ii), it follows that on Γ_b,

$$\begin{aligned} &\{\mathrm{curl}\, [(E - E^+ - E^{\mathrm{in}}|_{\Gamma_b})]_{\mathrm{mo}}\}_T \\ &= \sum_{n,l:\beta_n > 0} \mathrm{i}(-1)^{l+1} k C_{n,l} (U_{n,1-l})_T + \sum_{n,l:\beta_n \notin \mathbb{R}} \mathrm{i}(-1)^l k^{2l} C_{n,l} \sqrt{|\alpha_n|^2 + |\beta_n|^2}^{\,1-2l} (U_{n,1-l})_T \\ &= \sum_{n,l:\beta_n > 0} -\mathrm{i}k C_{n,l} (\cos\theta_n)^{1-2l} e_3 \times U_{n,l} + \sum_{n,l:\beta_n \notin \mathbb{R}} \mathrm{i}k^{2l} C_{n,l} (\beta_n)^{1-2l} e_3 \times U_{n,l}. \end{aligned} \tag{3.16}$$

Inserting (3.16) into (3.15) and choosing $V^+ = U_{n,l}$ for some $n \notin \Upsilon$, we derive $C_{n,l} = 0$. Analogously, the choice $V^+ = U_{n,0}$ for some $n \in \Upsilon$ leads to $C_{n,0} = 0$. This completes the proof. □

REMARK 3.1 In the non-resonance case, i.e. $\Upsilon = \emptyset$, the variational formulations (3.6) and (2.7) are equivalent. In fact, if (E, E^+) is a solution to the problem (3.6), then by Lemma 3.3, the transmission conditions in (3.4) hold. Hence, we obtain

$$0 = a((E, E^+ + E^{\mathrm{in}}), (V, V^+)) = \int_{\Omega_b} \operatorname{curl} E \cdot \operatorname{curl} \overline{V} - k^2 E \cdot \overline{V}\, \mathrm{d}x - \int_{\Gamma_b} \operatorname{curl}(E^+ + E^{\mathrm{in}}) \cdot e_3 \times \overline{V}\, \mathrm{d}s$$

$$= \int_{\Omega_b} \operatorname{curl} E \cdot \operatorname{curl} \overline{V} - k^2 E \cdot \overline{V}\, \mathrm{d}x - \int_{\Gamma_b} \mathscr{R}(e_3 \times E) \cdot e_3 \times \overline{V}\, \mathrm{d}s$$

$$+ \int_{\Gamma_b} [\mathscr{R}(e_3 \times E^{\mathrm{in}}) - \operatorname{curl} E^{\mathrm{in}}] \cdot e_3 \times \overline{V}\, \mathrm{d}s,$$

which is equivalent to the variational formulation (2.7) involving the Dirichlet-to-Neumann map $\mathscr{R}$. Note that in the last step of the previous identity we have used the identity $(\operatorname{curl} E^+)_T = \mathscr{R}(e_3 \times E^+) = \mathscr{R}(e_3 \times E) - \mathscr{R}(e_3 \times E^{\mathrm{in}})$ on Γ_b. On the other hand, supposing that $E \in H(\operatorname{curl}, \Omega_b)$ is a solution to (2.7), we extend the scattered field $E^{\mathrm{sc}} := E - E^{\mathrm{in}}$ from Ω_b to $x_3 > b$ by the Rayleigh expansion (2.5). Assume that the coefficients A_n are given by

$$e_3 \times E^{\mathrm{sc}}|_{\Gamma_b^-} = e_3 \times (E - E^{\mathrm{in}})|_{\Gamma_b^-} = \sum_{n \in \mathbb{Z}^2} A_n\, \mathrm{e}^{\mathrm{i}\alpha_n \cdot x'} \in H_t^{-1/2}(\operatorname{Div}, \Gamma_b), \quad A_n \in \mathbb{C}^3. \tag{3.17}$$

Here and in the following, the symbol $(\cdot)|_{\Gamma_b^-}$ resp. $(\cdot)|_{\Gamma_b^+}$ denotes the trace obtained from below and above Γ_b, respectively. It follows from the variational formulation (2.7) that $e_3 \times \operatorname{curl} E^{\mathrm{sc}} \times e_3|_{\Gamma_b^-} = \mathscr{R}(e_3 \times E^{\mathrm{sc}}|_{\Gamma_b^-})$. The extension of the series in (3.17) to the half space $x_3 > b$ in form of the Rayleigh expansion (2.5) is

$$E^{\mathrm{sc}}(x) = \sum_{n \in \mathbb{Z}^2} [A_n \times e_3 + \beta_n^{-1}(e_3 \times A_n) \cdot \alpha_n e_3]\, \mathrm{e}^{\mathrm{i}\alpha_n \cdot x' + \mathrm{i}\beta_n(x_3 - b)}, \quad x_3 > b.$$

Then, we obtain $e_3 \times E^{\mathrm{sc}}|_{\Gamma_b^-} = e_3 \times E^{\mathrm{sc}}|_{\Gamma_b^+}$ and $e_3 \times \operatorname{curl} E^{\mathrm{sc}} \times ye_3|_{\Gamma_b^+} = \mathscr{R}(e_3 \times E^{\mathrm{sc}}|_{\Gamma_b^+})$. Setting $E^+ = E^{\mathrm{sc}}$ in Ω_b^+, we conclude that (E, E^+) satisfies the transmission conditions (3.4) and thus is a solution of (3.6).

4. Analysis of the variational formulation (3.6)

Since the sesquilinear form $a(\cdot,\cdot)$ defined in Section 3 is bounded on $\mathbb{H}$, it obviously generates a continuous linear operator $A : \mathbb{H} \to \mathbb{H}'$ satisfying

$$a((E, E^+), (V, V^+)) = \langle A(E, E^+), (V, V^+) \rangle_{\Omega_b \times \Omega_b^+}. \tag{4.1}$$

Here $\mathbb{H}'$ denotes the dual of the space $\mathbb{H}$ with respect to the duality $\langle \cdot, \cdot \rangle_{\Omega_b \times \Omega_b^+}$ extending the scalar product in $L^2(\Omega_b)^3 \times L^2(\Omega_b^+)^3$. The aim of this section is to prove the following theorem:

THEOREM 4.1 The operator A defined by (4.1) is a Fredholm operator with index zero.

To prove Theorem 4.1, we need several auxiliary lemmas. We first prove a periodic analogue of the Hodge-decomposition of X, following the argument in Monk (2003, Theorem 4.3). See also Abboud (1993), Ammari (1995), Ammari & Bao (2008) and Hu *et al.* (2010) for other Hodge-decompositions of the Sobolev spaces in periodic structures. Define the subspaces $X_1 := \{\nabla p : p \in H^1(\Omega_b), p = 0 \text{ on } \Gamma\}$ and $X_0 := \{E_0 \in X : \int_{\Omega_b} \nabla p \cdot \overline{E}_0\, \mathrm{d}x = 0 \forall\ \nabla p \in X_1\}$.

LEMMA 4.1 We have $X = X_0 \oplus X_1$ with the subspaces X_0 and X_1 orthogonal in $L^2(\Omega_b)^3$ and $H(\text{curl}, \Omega_b)$. Moreover, $\text{div}\, E_0 = 0$ and $(e_3 \cdot E_0)|_{\Gamma_b} = 0$ for any $E_0 \in X_0$, and X_0 is compactly embedded into $L^2(\Omega_b)^3$.

Proof. Define the sesquilinear form $b(E, V) := \int_{\Omega_b} \{\text{curl}\, E \cdot \text{curl}\, \overline{V} + E \cdot \overline{V}\}\, \text{d}x$ for all $E, V \in X$. Then, for $\nabla p \in X_1$, we obtain $b(\nabla p, \nabla p) = \|\nabla p\|^2_{L^2(\Omega_b)} = \|\nabla p\|^2_X$. Thus, for every $E \in X$, there exists a unique solution $\nabla p \in X_1$ such that

$$b(\nabla p, \nabla \xi) = b(E, \nabla \xi), \quad \forall\, \nabla \xi \in X_1. \tag{4.2}$$

Let $E_0 := E - \nabla p$. Using integration by parts and the quasiperiodicity of E_0 and ξ in Ω_b, (4.2) implies

$$0 = \int_{\Omega_b} E_0 \cdot \nabla \overline{\xi}\, \text{d}x = -\int_{\Omega_b} \overline{\xi}\, \text{div}\, E_0\, \text{d}x + \int_{\Gamma_b} \overline{\xi} e_3 \cdot E_0\, \text{d}s, \quad \forall\, \nabla \xi \in X_1.$$

Consequently, $X = X_1 + X_0$ and $\text{div}\, E_0 = 0$, $(e_3 \cdot E_0)|_{\Gamma_b} = 0$. On the other hand, if $\nabla q \in X_0 \cap X_1$, then the definition of X_0 implies that $\int_{\Omega_b} \nabla p \cdot \nabla \overline{q}\, \text{d}x = 0$. Setting $p = q$, we get $\nabla q = 0$, i.e., $X_0 \cap X_1 = \emptyset$. Finally, the compact imbedding of X_0 into $L^2(\Omega_b)^3$ follows from Monk (2003, Corollary 3.49) (see also Ammari & Bao, 2008, Lemma 3.2). □

By Lemma 4.1 and the definitions of Y_l, we can decompose our space $\mathbb{H}$ into four subspaces. For $(E, E^+), (V, V^+) \in \mathbb{H}$, we may assume that

$$\begin{aligned} &E = \nabla p + E_0,\ E^+ = E_0^+ + E_1^+ \quad \text{where } \nabla p \in X_1,\ E_0 \in X_0,\ E_l^+ \in Y_l,\ l = 1, 2,\\ &V = \nabla \xi + V_0,\ V^+ = V_0^+ + V_1^+ \quad \text{where } \nabla \xi \in X_1,\ V_0 \in X_0,\ V_l^+ \in Y_l,\ l = 1, 2. \end{aligned}$$

To analyse the form a, we define several sesquilinear forms a_j with $j = 1, 2, \ldots, 6$. Let

$$\begin{aligned} a_1(\nabla p, \nabla \xi) &:= k^2 \int_{\Omega_b} \nabla p \cdot \nabla \overline{\xi}\, \text{d}x, \quad \forall\, \nabla p,\ \nabla \xi \in X_1,\\ a_2(E_0, V_0) &:= \int_{\Omega_b} \{\text{curl}\, E_0 \cdot \text{curl}\, \overline{V_0} - k^2 E_0 \cdot \overline{V_0}\}\, \text{d}x, \quad \forall\, E_0,\ V_0 \in X_0,\\ a_3(E_0^+, V_0^+) &:= \int_{\Gamma_b} e_3 \times E_0^+ \cdot \text{curl}\, \overline{V}_0^+\, \text{d}s, \quad \forall\, E_0^+,\ V_0^+ \in Y_0,\\ a_4(E_1^+, V_1^+) &:= \int_{\Gamma_b} e_3 \times E_1^+ \cdot \text{curl}\, \overline{V}_1^+\, \text{d}s, \quad \forall\, E_1^+,\ V_1^+ \in Y_1 \end{aligned}$$

and let

$$a_5((E, E^+), (V, V^+)) := a_5(E, V^+) := \int_{\Gamma_b} e_3 \times E \cdot \text{curl}\, \overline{V}^+\, \text{d}s,$$

$$a_6((E, E^+), (V, V^+)) := \eta \sum_{n \in \Upsilon} \left\{ \int_{\Gamma_b} e_3 \times (E - E^+) \cdot (e_3 \times \overline{U}_{n,0})\, \text{d}s \overline{\int_{\Gamma_b} e_3 \times (V - V^+) \cdot (e_3 \times \overline{U}_{n,0})\, \text{d}s} \right\} \tag{4.3}$$

for any $(E, E^+), (V, V^+) \in \mathbb{H}$.

LEMMA 4.2 For any $\nabla \xi \in X_1$ and $V_0^+ \in Y_0$, we have $a_5(\nabla \xi, V_0^+) = 0$.

Proof. From the definition of Y and Y_0 we conclude that Y_0 is in the subspace of all vector functions $V^+ \in Y$ with $e_3 \cdot V^+ = 0$. Therefore it suffices to prove

$$\int_{\Gamma_b} [e_3 \times \nabla\xi] \cdot \operatorname{curl} \overline{V}^+ \,\mathrm{d}s = k^2 \int_{\Gamma_b} e_3 \cdot \overline{V}^+ \xi \,\mathrm{d}s. \tag{4.4}$$

Note that the right-hand side of (4.4) is a continuous functional of V^+ and ξ. Indeed, from $V^+ \in L^2(\Omega_b^+)$ and $0 = \nabla \cdot V^+ \in L^2(\Omega_b^+)$, we conclude $e_3 \cdot \overline{V}^+|_{\Gamma_b} \in H^{-1/2}(\Gamma_b)$, and $\xi \in H^{1/2}(\Gamma_b)$ follows from $\xi \in H^1(\Omega_b)$. Knowing the continuity, it suffices to prove (4.4) for a dense subset, e.g., for a truncated Rayleigh expansion V^+ and smooth ξ. We conclude

$$\begin{aligned}\int_{\Gamma_b} [e_3 \times \nabla\xi] \cdot \operatorname{curl} \overline{V}^+ \,\mathrm{d}s &= -\int_{\Gamma} [\nu \times \nabla\xi] \cdot \operatorname{curl} \overline{V}^+ \,\mathrm{d}s + \int_{\Omega_b} [\operatorname{curl} \nabla\xi] \cdot \operatorname{curl} \overline{V}^+ \,\mathrm{d}s \\ &\quad - \int_{\Omega_b} [\nabla\xi] \cdot \operatorname{curl}\operatorname{curl} \overline{V}^+ \,\mathrm{d}s = k^2 \int_{\Omega_b} [\nabla\xi] \cdot \overline{V}^+ \,\mathrm{d}s,\end{aligned}$$

where we have used that the tangential derivative $\nu \times \nabla\xi$ of the function ξ with $\xi|_{\Gamma} = 0$ vanishes. Using $\nabla \cdot V^+ = 0$, we continue

$$k^2 \int_{\Omega_b} [\nabla\xi] \cdot \overline{V}^+ \,\mathrm{d}s = k^2 \int_{\Omega_b} \nabla \cdot [\xi \overline{V}^+] \,\mathrm{d}s = k^2 \int_{\Gamma_b} \xi e_3 \cdot \overline{V}^+ \,\mathrm{d}s + k^2 \int_{\Gamma} \xi \nu \cdot \overline{V}^+ \,\mathrm{d}s = k^2 \int_{\Gamma_b} \xi e_3 \cdot \overline{V}^+ \,\mathrm{d}s$$

and the proof is completed. □

Using Lemmas 3.1, 4.1 and 4.2, it follows from the definition of a that (see Table 1)

$$\begin{aligned}
&a((E, E^+), (V, V^+)) \\
&\quad = a((\nabla p + E_0, E_0^+ + E_1^+), (\nabla\xi + V_0, V_0^+ + V_1^+)) \\
&\quad = \int_{\Omega_b} \{\operatorname{curl} E_0 \cdot \operatorname{curl} \overline{V}_0 - k^2 E_0 \cdot \overline{V}_0 - k^2 \nabla p \cdot \nabla\overline{\xi}\} \,\mathrm{d}x - \int_{\Gamma_b} \operatorname{curl} E_0^+ \cdot e_3 \times \overline{V}_0 \,\mathrm{d}x \\
&\qquad - \int_{\Gamma_b} \operatorname{curl} E_1^+ \cdot e_3 \times \overline{V}_0 \,\mathrm{d}x - \int_{\Gamma_b} \operatorname{curl} E_1^+ \cdot e_3 \times \nabla\overline{\xi} \,\mathrm{d}x + \int_{\Gamma_b} e_3 \times E_0 \cdot \operatorname{curl} \overline{V}_0^+ \,\mathrm{d}x \\
&\qquad + \int_{\Gamma_b} e_3 \times E_0 \cdot \operatorname{curl} \overline{V}_1^+ \,\mathrm{d}x + \int_{\Gamma_b} e_3 \times \nabla p \cdot \operatorname{curl} \overline{V}_1^+ \,\mathrm{d}x - \int_{\Gamma_b} e_3 \times E_0^+ \cdot \operatorname{curl} \overline{V}_0^+ \,\mathrm{d}x \\
&\qquad - \int_{\Gamma_b} e_3 \times E_1^+ \cdot \operatorname{curl} \overline{V}_1^+ \,\mathrm{d}x + a_6((E, E^+), (V, V^+)) \\
&\quad = -a_1(\nabla p, \nabla\xi) + a_2(E_0, V_0) - a_3(E_0^+, V_0^+) - a_4(E_1^+, V_1^+) + a_5(E_0, V_0^+) - \overline{a_5(V_0, E_0^+)} + a_5(E_0, V_1^+) \\
&\qquad - \overline{a_5(V_0, E_1^+)} + a_5(\nabla p, V_1^+) - \overline{a_5(\nabla\xi, E_1^+)} + a_6((E, E^+), (V, V^+)).
\end{aligned}$$

DEFINITION 4.2 A bounded sesquilinear form $l(\cdot, \cdot)$ on a Hilbert space X is called strongly elliptic if there exists a compact form $\tilde{l}(\cdot; \cdot)$ and a constant $c > 0$ such that $\operatorname{Re} l(u, u) \geqslant c\|u\|_X^2 - \tilde{l}(u, u)$ for all $u \in X$.

TABLE 1 *The diagram for the sesquilinear form $a - a_6$ over $\mathbb{H} = X \times Y$*

		$\mathbb{H}_0 := X_0 + Y_0$		$\mathbb{H}_1 := X_1 + Y_1$	
		$X_0(V_0)$	$Y_0(V_0^+)$	$X_1(\nabla\xi)$	$Y_1(V_1^+)$
$\mathbb{H}_0$	$X_0(E_0)$	$a_2(E_0, V_0)$	$a_5(E_0, V_0^+)$	0	$a_5(E_0, V_1^+)$
	$Y_0(E_0^+)$	$-a_5(V_0, E_0^+)$	$-a_3(E_0^+, V_0^+)$	0	0
$\mathbb{H}_1$	$X_1(\nabla p)$	0	0	$-a_1(\nabla p, \nabla\xi)$	$a_5(\nabla p, V_1^+)$
	$Y_1(E_1^+)$	$-a_5(V_0, E_1^+)$	0	$-a_5(\nabla\xi, E_1^+)$	$-a_4(E_1^+, V_1^+)$

Obviously, a_1 is coercive on X_1 and by Lemma 4.1 the sesquilinear form a_2 is strongly elliptic over X_0. In addition, a_6 is a compact form over $\mathbb{H}$, since it corresponds to a finite rank operator over $\mathbb{H}$. To derive the Fredholmness of the sesquilinear form a, we have to study a_3, a_4 and a_5.

LEMMA 4.3 There exist compact forms $\tilde{a}_3 : Y_0 \times Y_0 \to \mathbb{C}$ and $\tilde{a}_4 : Y_1 \times Y_1 \to \mathbb{C}$ such that

$$-\mathrm{Re}\, a_3(\cdot,\cdot) \geqslant C\|\cdot\|^2_{H(\mathrm{curl}\,,\Omega_b^+)} - \tilde{a}_3(\cdot,\cdot), \tag{4.5}$$

$$\mathrm{Re}\, a_4(\cdot,\cdot) \geqslant C\|\cdot\|^2_{H(\mathrm{curl}\,,\Omega_b^+)} - \tilde{a}_4(\cdot,\cdot), \tag{4.6}$$

for some constant $C > 0$, i.e., the sesquilinear forms $-a_3$ and a_4 are strongly elliptic over Y_0 and Y_1, respectively.

Proof. Recall that the functions $U_{n,l}$ defined in (3.3) are basis functions of the space Y_l, $l = 1, 2$. It is easy to check that

$$\begin{aligned}\int_{\Omega_b^+} U_{n,l}\cdot \overline{U}_{n',l'}\,\mathrm{d}x &= \delta_{n,n'}\delta_{l,l'}\Lambda_1\Lambda_2\int_b^{b+1}\exp(\mathrm{i}\beta_n x_3)\exp(-\mathrm{i}\overline{\beta}_n x_3)\,\mathrm{d}x_3\\ &= \begin{cases}\delta_{n,n'}\delta_{l,l'}\Lambda_1\Lambda_2 & \text{if } \beta_n\in\mathbb{R},\\ \delta_{n,n'}\delta_{l,l'}\,\mathrm{e}^{-2|\beta_n|b}(1-\mathrm{e}^{-2|\beta_n|})(2|\beta_n|)^{-1}\Lambda_1\Lambda_2 & \text{if } \beta_n\notin\mathbb{R},\end{cases}\end{aligned} \tag{4.7}$$

and that, by using Lemma 3.1,

$$\int_{\Omega_b^+}\mathrm{curl}\,U_{n,l}\cdot\mathrm{curl}\,\overline{U}_{n',l'}\,\mathrm{d}x = \delta_{n,n'}\delta_{l,l'}k^{4l}\sqrt{|\alpha_n|^2+|\beta_n|^2}^{\,2-4l}\int_{\Omega_b^+}U_{n,1-l}\cdot\overline{U}_{n,1-l}\,\mathrm{d}x. \tag{4.8}$$

Therefore, we can represent the $H(\mathrm{curl}\,,\Omega_b^+)$-norm of $U_{n,l}, l = 0, 1$ as

$$\|U_{n,0}\|^2_{H(\mathrm{curl}\,,\Omega_b^+)} = \begin{cases}(1+k^2)\Lambda_1\Lambda_2 & \text{if } \beta_n\in\mathbb{R},\\ \mathrm{e}^{-2|\beta_n|b}(1-\mathrm{e}^{-2|\beta_n|})(1+2|\beta_n|^2+k^2)(2|\beta_n|)^{-1}\Lambda_1\Lambda_2 & \text{if } \beta_n\notin\mathbb{R},\end{cases}$$

$$\|U_{n,1}\|^2_{H(\mathrm{curl}\,,\Omega_b^+)} = \begin{cases}(1+k^2)\Lambda_1\Lambda_2 & \text{if } \beta_n\in\mathbb{R},\\ \mathrm{e}^{-2|\beta_n|b}(1-\mathrm{e}^{-2|\beta_n|})\left(1+\dfrac{k^4}{2|\beta_n|^2+k^2}\right)(2|\beta_n|)^{-1}\Lambda_1\Lambda_2 & \text{if } \beta_n\notin\mathbb{R}.\end{cases}$$

On the other hand, using simple calculations, we have, for $n \notin \Upsilon$,

$$\int_{\Gamma_b} e_3 \times U_{n,l} \cdot \operatorname{curl} \overline{U}_{n,l}\, \mathrm{d}s = (-\mathrm{i})(\cos\theta_n)^{2l-1} k^{2l} \sqrt{|\alpha_n|^2 + |\beta_n|^2}^{\,1-2l} \int_{\Gamma_b} |e_3 \times U_{n,1-l}|^2\, \mathrm{d}s,$$

by Lemma 3.1(i) and (ii). Furthermore,

$$\int_{\Gamma_b} |e_3 \times U_{n,1-l}|^2\, \mathrm{d}s = \begin{cases} |\mathrm{e}^{\mathrm{i}\beta_n b}|^2 \Lambda_1 \Lambda_2 & \text{if } l = 1, \\ |\mathrm{e}^{\mathrm{i}\beta_n b}|^2 \Lambda_1 \Lambda_2 |\cos\theta_n|^2 & \text{if } l = 0, \end{cases}$$

by the definitions of $U_{n,l}$ given in (3.3). Combining the previous two equalities yields

$$\operatorname{Re} \int_{\Gamma_b} e_3 \times U_{n,1} \cdot \operatorname{curl} \overline{U}_{n,1}\, \mathrm{d}s = \begin{cases} 0 & \text{if } \beta_n \in \mathbb{R}, \\ \dfrac{|\beta_n| k^2}{2|\beta_n|^2 + k^2}\, \mathrm{e}^{-2|\beta_n| b} \Lambda_1 \Lambda_2 & \text{if } \beta_n \notin \mathbb{R}, \end{cases}$$

$$\operatorname{Re} \int_{\Gamma_b} e_3 \times U_{n,0} \cdot \operatorname{curl} \overline{U}_{n,0}\, \mathrm{d}s = \begin{cases} 0 & \text{if } \beta_n \in \mathbb{R}, \\ -|\beta_n|\, \mathrm{e}^{-2|\beta_n| b} \Lambda_1 \Lambda_2 & \text{if } \beta_n \notin \mathbb{R}. \end{cases}$$

Since $|\beta_n| \sim \sqrt{1 + |n|^2}$ as $|n| \to +\infty$, there holds

$$-\operatorname{Re} \int_{\Gamma_b} e_3 \times U_{n,0} \cdot \operatorname{curl} \overline{U}_{n,0}\, \mathrm{d}s \geqslant C \|U_{n,0}\|^2_{H(\operatorname{curl}, \Omega_b)}, \tag{4.9}$$

$$\operatorname{Re} \int_{\Gamma_b} e_3 \times U_{n,1} \cdot \operatorname{curl} \overline{U}_{n,1}\, \mathrm{d}s \geqslant C \|U_{n,1}\|^2_{H(\operatorname{curl}, \Omega_b)}, \tag{4.10}$$

whenever $\beta_n \notin \mathbb{R}$, with $C > 0$ being a constant independent of l and n. Therefore, given the field $E_0^+ = \sum_{n \in \mathbb{Z}^2} C_{n,0} U_{n,0} \in Y_0$, we deduce from (4.9) that

$$\begin{aligned} -\operatorname{Re} a_3(E_0^+, E_0^+) &= -\sum_{n \in \mathbb{Z}^2} |C_{n,0}|^2 \operatorname{Re} \int_{\Gamma_b} e_3 \times U_{n,0} \cdot \operatorname{curl} \overline{U}_{n,0}\, \mathrm{d}s \\ &\geqslant C \sum_{\beta_n \notin \mathbb{R}} |C_{n,0}|^2 \|U_{n,0}\|^2_{H(\operatorname{curl}, \Omega_b^+)} = C \|E_0^+\|^2_{H(\operatorname{curl}, \Omega_b^+)} - \tilde{a}_3(E_0^+, E_0^+), \\ \tilde{a}_3(E_0^+, E_0^+) &:= C \sum_{\beta_n \in \mathbb{R}} |C_{n,0}|^2 \|U_{n,0}\|^2_{H(\operatorname{curl}, \Omega_b^+)}. \end{aligned}$$

Since the set $\{n \in \mathbb{Z}^2 : \beta_n \in \mathbb{R}\}$ consists of a finite number of indices, the form $\tilde{a}_3(\cdot, \cdot) : Y_0 \times Y_0 \to \mathbb{R}$ is compact. Thus the sesquilinear form $-a_3$ is strongly elliptic over Y_0. The proof for a_4 can be carried out analogously by employing (4.10). □

REMARK 4.1 For the component-wise gradient $\nabla U_{n,l}$, the definition of $U_{n,l}$ leads to

$$\int_{\Omega_b^+} |\nabla U_{n,l}|^2\, \mathrm{d}x = (|\alpha_n|^2 + |\beta_n|^2) \int_{\Omega_b^+} |U_{n,l}|^2\, \mathrm{d}x. \tag{4.11}$$

Thus, comparing (4.11) with (4.8) leads to $\|U_{n,0}\|^2_{H^1(\Omega_b^+)^3} = \|U_{n,0}\|^2_{H(\mathrm{curl}\,,\Omega_b^+)}$. This implies that the H^1- and $H(\mathrm{curl}\,)$-norm of the elements from Y_0 are identical, i.e., $\|E_0^+\|_{H^1(\Omega_b^+)^3} = \|E_0^+\|_{H(\mathrm{curl}\,,\Omega_b^+)}$, if $E_0^+ \in Y_0$. However, this is not true for the space Y_1.

We turn to the properties of a_5 defined in (4.3).

LEMMA 4.4 The sesquilinear form a_5 is compact over $X_0 \times Y_1$.

Proof. For $V_1^+ \in Y_1 \subset Y$, define the operator $J(V_1^+) := \mathrm{curl}\, V_1^+$. Obviously, we get $\|J(V_1^+)\|^2_{L^2(\Omega_b^+)^3} \leqslant \|V_1^+\|^2_{H(\mathrm{curl}\,,\Omega_b^+)}$, and $\mathrm{curl}\, J(V_1^+) - k^2 V_1^+ = 0$ implies $\|\mathrm{curl}\, J(V_1^+)\|^2_{L^2(\Omega_b^+)^3} \leqslant k^2 \|V_1^+\|^2_{H(\mathrm{curl}\,,\Omega_b^+)}$. Hence, by Lemma 3.1(i), J is a bounded linear map from Y_1 into Y_0. In view of the equivalence of the norms $\|JV_1^+\|_{H(\mathrm{curl}\,,\Omega_b^+)}$ and $\|JV_1^+\|_{H^1(\Omega_b^+)^3}$ (see Remark 4.1), we see that J is also bounded from the subspace Y_1 of $H(\mathrm{curl}\,, \Omega_b^+)$ into the subspace Y_0 of $H^1(\Omega_b^+)^3$, with the trace $J(V_1^+)|_{\Gamma_b} \in H^{1/2}(\Gamma_b)^3$. Thus, there exists an extension W of $(\mathrm{curl}\, V_1^+)|_{\Gamma_b}$ from $H^{1/2}(\Gamma_b)^3$ into $H^1(\Omega_b)^3$ such that $W = \mathrm{curl}\, V_1^+$ on Γ_b and $\nu \times W = 0$ on Γ. Using integration by parts,

$$a_5(E_0, V_1^+) = \int_{\Gamma_b} e_3 \times E_0 \cdot \overline{J(V^+)}\,\mathrm{d}s = \int_{\Gamma_b} e_3 \times E_0 \cdot \overline{W}\,\mathrm{d}s = \int_{\Omega_b} \{\mathrm{curl}\, E_0 \cdot \overline{W} - E_0 \cdot \mathrm{curl}\, \overline{W}\}\,\mathrm{d}x.$$

From the compact embedding of $W \in H^1(\Omega_b)^3$ into $L^2(\Omega_b)^3$ and that of $E_0 \in X_0$ into $L^2(\Omega_b)^3$, it follows that the sesquilinear form $a_5(E_0, V_1^+)$ is compact over $X_0 \times Y_1$. □

Combining Lemmas 4.1, 4.2, 4.3 and 4.4, we are now in a position to prove the Fredholm property of the variational formulation (3.6).

Proof of Theorem 4.1. It suffices to verify that the sesquilinear form $a - a_6$ is Fredholm over $\mathbb{H}$ with index zero. To do this, we define the spaces $\mathbb{H}_j = X_j \oplus Y_j$ for $j = 0, 1$, so that we can rewrite $\mathbb{H} = X \times Y = \mathbb{H}_0 \times \mathbb{H}_1$. Define the sesquilinear forms:

$$b_0((E_0, E_0^+), (V_0, V_0^+)) := a_2(E_0, V_0) - a_3(E_0^+, V_0^+) + a_5(E_0, V_0^+) - \overline{a_5(V_0, E_0^+)},$$
$$b_1((\nabla p, E_1^+), (\nabla \xi, V_1^+)) := -a_1(\nabla p, \nabla \xi) - a_4(E_1^+, V_1^+) + a_5(\nabla p, V_1^+) - \overline{a_5(\nabla \xi, E_1^+)},$$

for all $(E_0, E_0^+), (V_0, V_0^+) \in \mathbb{H}_0$ and for all $(\nabla p, E_1^+), (\nabla \xi, V_1^+) \in \mathbb{H}_1$, respectively. Now split the form in Table 1 in blocks corresponding to the splitting $\mathbb{H} = \mathbb{H}_0 \times \mathbb{H}_1$. Then the restriction to $\mathbb{H}_0$ is the form b_0 with the strongly elliptic quadratic form $\mathrm{Re}\; b_0((E_0, E_0^+), (E_0, E_0^+)) = a_2(E_0, E_0) - a_3(E_0^+, E_0^+)$. The restriction to $\mathbb{H}_1$ is the form b_1, and the sesquilinear form $-b_1$ has the strongly elliptic quadratic form $-\mathrm{Re}\; b_1((\nabla p, E_1^+), (\nabla p, E_1^+)) = a_1(\nabla p, \nabla p) + a_4(E_1^+, E_1^+)$. Consequently, the diagonal blocks of the 2×2 splitting correspond to Fredholm operators with index zero. On the other hand, the full form in Table 1 differs from the diagonal block matrix only by compact terms. Hence the form a generates a Fredholm operator with index zero. □

5. Proof of Theorem 2.1

Since the problem (BVP′) and (3.6) are equivalent (see Lemma 3.3), to prove Theorem 2.1 we only need to show the existence of solutions to (3.6) with E^{in} replaced by $E^{\mathrm{in}}_{\mathrm{gen}}$ given in (2.6). Consider the homogeneous adjoint problem of the variational formulation (3.6): find $(V, V^+) \in \mathbb{H}$ such that

$$a((W, W^+), (V, V^+)) = 0 \tag{5.1}$$

for all $(W, W^+) \in \mathbb{H}$. By Fredholm's alternative, it suffices to verify $a((0, E^{\mathrm{in}}_{\mathrm{gen}}), (V, V^+)) = 0$ for any solution (V, V^+) to (5.1). The following lemma describes properties of the solution (V, V^+), which will be used later for proving Theorem 2.1.

LEMMA 5.1 Assume that $(V, V^+) \in \mathbb{H}$ is a solution to the homogeneous adjoint problem (5.1). Then

$$V_T|_{\Gamma_b},\ (\mathrm{curl}\, V^+)_T|_{\Gamma_b} \in \mathrm{Span}\{\{(U_{n,l})_T|_{\Gamma_b} : \beta_n \notin \mathbb{R}, l = 1, 2\} \cup \{U_{n,0}|_{\Gamma_b} : \beta_n = 0\}\}. \tag{5.2}$$

Proof. Analogous to the proof of (3.12), one can prove that $\mathrm{curl}\,\mathrm{curl}\, V - k^2 V = 0$ holds in Ω_b, leading to the identity (3.10) with (V, E) replaced by (W, V). By the definition of $a(\cdot, \cdot)$,

$$\begin{aligned} 0 &= a((W, W^+), (V, V^+)) \\ &= \int_{\Gamma_b} \{e_3 \times W \cdot \mathrm{curl}\, \overline{V} - \mathrm{curl}\, W^+ \cdot e_3 \times \overline{V}\}\, \mathrm{d}s + \int_{\Gamma_b} e_3 \times (W - W^+) \cdot \mathrm{curl}\, \overline{V}^+\, \mathrm{d}s \\ &\quad + \eta \sum_{n \in \Upsilon} \int_{\Gamma_b} e_3 \times (W - W^+) \cdot e_3 \times \overline{U}_{n,0}\, \mathrm{d}s \overline{\int_{\Gamma_b} e_3 \times (V - V^+) \cdot e_3 \times \overline{U}_{n,0}\, \mathrm{d}s} \end{aligned} \tag{5.3}$$

for all $(W, W^+) \in \mathbb{H}$. In the following we will prove (5.2) choosing different test functions $(W, W^+) \in \mathbb{H}$.

(i) Choose $W \equiv 0, W^+ = U_{n,0}$ for some $n \in \Upsilon$ in (5.3). Since $(\mathrm{curl}\, U_{n,0}|_{\Gamma_b})_T = 0$ on Γ_b (cf. Lemma 3.1), simple calculations lead to $\int_{\Gamma_b} [\mathrm{curl}\, V^+ + \eta \Lambda_1 \Lambda_2 e_3 \times (V - V^+)] \cdot e_3 \times \overline{U}_{n,0}\, \mathrm{d}s = 0$. However, one can verify, using Lemma 3.1(i) and (ii), that $\int_{\Gamma_b} \{\mathrm{curl}\, V^+ \cdot e_3 \times \overline{U}_{n,0}\}\, \mathrm{d}s = 0$ for $V^+ \in Y$. Hence,

$$\int_{\Gamma_b} e_3 \times (V - V^+) \cdot e_3 \times \overline{U}_{n,0}\, \mathrm{d}s = 0 \quad \text{if } n \in \Upsilon. \tag{5.4}$$

(ii) Choose $W \equiv 0$ and $W^+ = U_{n,1}$ for some $n \in \Upsilon$ in (5.3). Making use of $e_3 \times U_{n,1} = 0$ for $n \in \Upsilon$, we derive from (5.3) that $\int_{\Gamma_b} \{\mathrm{curl}\, U_{n,1} \cdot e_3 \times \overline{V}\}\, \mathrm{d}s = 0$. This together with Lemma 3.1(i) gives the relation

$$\int_{\Gamma_b} \{e_3 \times \overline{V} \cdot U_{n,0}\}\, \mathrm{d}s = 0 \quad \text{if } n \in \Upsilon. \tag{5.5}$$

(iii) Inserting (5.4) and (5.5) into (5.3) with $W \equiv 0$ and taking into account Lemma 3.2, we obtain $\int_{\Gamma_b} \{\mathrm{curl}\, [V]_{\mathrm{mo}} + \mathrm{curl}\, V^+\} \cdot e_3 \times W^+\, \mathrm{d}s = 0$ for all $W^+ \in Y$. By Lemma 3.1(iii), the last identity implies that

$$\{(\mathrm{curl}\, [V]_{\mathrm{mo}})_T + (\mathrm{curl}\, V^+)_T\}|_{\Gamma_b} \in \mathrm{Span}\{U_{n,0} : n \in \Upsilon\}. \tag{5.6}$$

Since $V^+ \in Y$, we have $\mathrm{curl}\, V^+ \in H(\mathrm{curl}\,, \Omega_b^+)$ and thus the trace $(\mathrm{curl}\, V^+|_{\Gamma_b})_T$ belongs to $H_t^{-1/2}(\mathrm{Curl}\,, \Gamma_b)$. Using Lemma 3.1(iii), we may assume that on Γ_b

$$(\mathrm{curl}\, V^+)_T = \sum_{n:\beta_n = 0} \{B_{n,0} U_{n,0} + B_{n,1}\, e_3 \times U_{n,0}\} + \sum_{l,n:\beta_n \neq 0} B_{n,l}\, e_3 \times U_{n,l}$$

with $B_{n,l} \in \mathbb{C}$. Combining the previous two formulas, we deduce from the definition of the modification operator $[\cdot]_{\mathrm{mo}}$ in Lemma 3.2 that $(\mathrm{curl}\,[V]_{\mathrm{mo}})_T + \sum_{l,n:\beta_n \neq 0} B_{n,l}\, e_3 \times U_{n,l} = 0$ on Γ_b and that $B_{n,1} = 0$ for $\beta_n = 0$. Therefore,

$$(\mathrm{curl}\, V^+)_T = \sum_{n:\beta_n=0} B_{n,0} U_{n,0} - (\mathrm{curl}\,[V]_{\mathrm{mo}})_T \quad \text{on } \Gamma_b. \tag{5.7}$$

(iv) Inserting (5.4) and (5.5) into (5.1) with $W^+ = 0$ and $W = V$, we find (cf. (3.5) with $E^+ \equiv 0$ and $E = V$)

$$0 = \mathrm{Im}\,\overline{a((V,0),(V,V^+))} = \mathrm{Im} \int_{\Gamma_b} e_3 \times \overline{V} \cdot (\mathrm{curl}\, V^+)_T \,\mathrm{d}s, \tag{5.8}$$

where the function $(\mathrm{curl}\, V^+)_T|_{\Gamma_b}$ is given in (5.7). According to Lemma 3.1(iii), we may represent $e_3 \times V|_{\Gamma_b}$ as

$$e_3 \times V = \sum_{l,n:\beta_n \neq 0} C_{n,l} e_3 \times U_{n,l} + \sum_{n:\beta_n=0} \{C_{n,0} e_3 \times U_{n,0} + C_{n,1} U_{n,0}\}, \quad C_{n,l} \in \mathbb{C},$$

on Γ_b. However, by (5.5) there holds $C_{n,1} = 0$ for $n \in \Upsilon$. Thus, applying Lemma 3.1 gives

$$e_3 \times V = \sum_{l,n:\beta_n \neq 0} C_{n,l} (-1)^l (U_{n,1-l})_T (\cos\theta_n)^{2l-1} + \sum_{n:\beta_n=0} C_{n,0} e_3 \times U_{n,0} \quad \text{on } \Gamma_b, \tag{5.9}$$

$$\begin{aligned} (\mathrm{curl}\,[V]_{\mathrm{mo}})_T &= -\sum_{l,n:\beta_n>0} C_{n,l} (\mathrm{curl}\, U_{n,l})_T + \sum_{l,n:\beta_n \notin \mathbb{R}} C_{n,l} (\mathrm{curl}\, U_{n,l})_T \\ &= -\sum_{l,n:\beta_n>0} \mathrm{i}(-1)^l k C_{n,l} (U_{n,1-l})_T + \sum_{l,n:\beta_n \notin \mathbb{R}} \mathrm{i}(-1)^l \sqrt{|\alpha_n|^2 + |\beta_n|^2}^{\,1-2l} k^{2l} C_{n,l} (U_{n,1-l})_T \end{aligned} \tag{5.10}$$

on Γ_b. Inserting the above identity (5.10) into (5.7) and using (5.9), we derive from (5.8) that

$$\begin{aligned} 0 &= \mathrm{Im} \left\{ -\mathrm{i}k \sum_{l,n:\beta_n>0} |C_{n,l}|^2 \|(U_{n,1-l})_T\|^2_{L^2(\Gamma_b)} (\cos\theta_n)^{2l-1} \right\} \\ &\quad + \mathrm{Im} \left\{ -\sum_{l,n:\beta_n \notin \mathbb{R}} |C_{n,l}|^2 (\mathrm{i}k)^{2l} \|(U_{n,1-l})_T\|^2_{L^2(\Gamma_b)} |\beta_n|^{1-2l} [|\alpha_n|^2 + |\beta_n|^2]^{1-2l} \right\} \\ &= -k \sum_{l,n:\beta_n>0} |C_{n,l}|^2 \|(U_{n,1-l})_T\|^2_{L^2(\Gamma_b)} (\cos\theta_n)^{2l-1}, \end{aligned}$$

which, together with the definition of $\cos\theta_n$ in Lemma 3.1, leads to

$$C_{n,l} = 0 \quad \text{for all } \beta_n > 0, \;\; l = 1, 2. \tag{5.11}$$

Finally, combining (5.11) and (5.9) we have proved (5.2) for $V_T|_{\Gamma_b}$, and combining (5.11), (5.10) and (5.7) leads to the desired result for $(\mathrm{curl}\, V^+|_{\Gamma_b})_T$. □

We proceed to prove Theorem 2.1, i.e., to show the existence of a solution $(E, E^+) \in \mathbb{H}$ to the variational problem (3.6) for the incident wave $E^{\mathrm{in}}_{\mathrm{gen}}$. Assume $(V, V^+) \in \mathbb{H}$ satisfies $a((W, W^+), (V, V^+)) = 0$ for all $(W, W^+) \in \mathbb{H}$. Using Lemmas 5.1 and 3.1, it is easy to check that

$$a((0, E^{\mathrm{in}}_{\mathrm{gen}}), (V, V^+)) = -\int_{\Gamma_b} \{\mathrm{curl}\, E^{\mathrm{in}}_{\mathrm{gen}} \cdot e_3 \times \overline{V} + e_3 \times E^{\mathrm{in}}_{\mathrm{gen}} \cdot \mathrm{curl}\, \overline{V}^+\}\, \mathrm{d}s = 0. \tag{5.12}$$

This means that each solution to the homogeneous adjoint problem (5.1) is orthogonal to the right-hand side of the variational problem (3.6) in the sense of (5.12). According to Theorem 4.1, the Fredholm alternative applied to the variational problem (3.6) yields the existence of the solution $(E, E^+) \in \mathbb{H}$ to problem (3.6) for the incident plane waves $E^{\mathrm{in}}_{\mathrm{gen}}$ defined in (2.6).

Formula (5.2) implies that solution V^+ takes the form $V^+(x) = \sum_{\beta_n \notin \mathbb{R}} C_{n,l} U_{n,l}(x) + \sum_{\beta_n = 0} C_{n,l} U_{n,l}(x)$ for $x \in \Omega_b^+$. Thus $V^+ \in Y$ and the coefficients of the propagating modes for $\beta_n > 0$ vanish. By analogous arguments, this assertion even remains valid for the solution (V, V^+) to the homogeneous variational problem $a((V, V^+), (W, W^+)) = 0$ for all $(W, W^+) \in \mathbb{H}$. In other words, the coefficient $C_{n,l}$ of the difference of two solutions of (BVP) are zero if $\beta_n > 0$. The proof of Theorem 2.1 is thus completed.

Acknowledgement

The authors would like to thank their colleague J. Elschner for pointing out the research articles (Gotlib, 2000; Kamotski & Nazarov, 2002) which stimulated this paper.

Funding

This research was supported by the German Research Foundation (DFG) under Grant No. EL 584/1-2 (to G.H.).

References

Abboud, T. (1993) Formulation variationnelle des équations de Maxwell dans un réseau bipériodique de $\mathbb{R}^3$. *C. R. Acad. Sci. Pairs*, **317**, 245–248.

Ammari, H. (1995) Uniqueness theorems for an inverse problem in a doubly periodic structure. *Inv. Probl.*, **11**, 823–833.

Ammari, H. & Bao, G. (2008) Coupling of finite element and boundary element methods for the scattering by periodic chiral structures. *J. Comput. Math.*, **3**, 261–283.

Bao, G. (1997) Variational approximation of Maxwell's equation in biperiodic structures. *SIAM J. Appl. Math.*, **57**, 364–381.

Bao, G., Cowsar, L. & Masters, W. (eds) (2001) *Mathematical Modeling in Optical Science*. *Frontiers in Applied Mathematics*. SIAM Series. Philadelphia: SIAM.

Bao, G. & Dobson, D. C. (2000) On the scattering by biperiodic structures. *Proc. Amer. Math. Soc.*, **128**, 2715–2723.

Bonnet-Bendhia, A. S. & Starling, F. (1994) Guided waves by electromagnetic gratings and non-uniqueness examples for the diffraction problem. *Math. Meth. Appl. Sci.*, **17**, 305–338.

Buffa, A., Costabel, M. & Sheen, D. (2002) On traces for H(curl , Ω) in Lipschitz domains. *J. Math. Anal. Appl.*, **276**, 847–867.

Chen, X. & Friedman, A. (1991) Maxwell's equations in a periodic structure. *Trans. AMS*, **323**, 465–507.

Colton, D. & Kress, R. (1998) *Inverse Acoustic and Electromagnetic Scattering Theory*. Berlin: Springer.

Dobson, D. C. (1994) A variational method for electromagnetic diffraction in biperiodic structures. *RAIRO Model. Math. Anal. Numer*, **28**, 419–439.

Elschner, J. & Hu, G. (2010) Variational approach to scattering of plane elastic waves by diffraction gratings. *Math. Meth. Appl. Sci*, **33**, 1924–1941.

Elschner, J. & Schmidt, G. (1998) Diffraction in periodic structures and optimal design of binary gratings I. Direct problems and gradient formulas. *Math. Methods Appl. Sci.*, **21**, 1297–1342.

Gohberg I. C. & Krein, M. G. (1969) *Introduction to the Theory of Linear Nonselfadjoint Operators in Hilbert Space*. Translations of Mathematical Monographs, vol. 18. Providence, RI: American Mathematical Society.

Gotlib, V. Yu. (2000) Solutions of the Helmholtz equation, concentrated near a plane periodic boundary. *J. Math. Sci.*, **102**, 4188–4194.

Haddar, H. & Lechleiter, A. (2011) Electromagnetic wave scattering from rough penetrable layers. *SIAM J. Math. Anal.*, **43**, 2418–2443.

Hiptmaier, R. (2012) Coupling of finite elements and boundary elements in electromagnetic scattering. *SIAM J. Numer. Anal.*, **41**, 919–944.

Hu, G., Qu, F. & Zhang, B. (2010) Direct and inverse problems for electromagnetic scattering by a doubly periodic structure with a partially coated dielectric. *Math. Methods Appl. Sci.*, **33**, 147–156.

Hu, G. & Rathsfeld, A. (2012) Convergence analysis of the FEM coupled with Fourier-mode expansion for the electromagnetic scattering by biperiodic structures. *WIAS preprint*, no. 1744, Berlin.

Hu, G., Yang, J. & Zhang, B. (2009) Inverse electromagnetic scattering for a biperiodic inhomogeneous layer on perfectly conducting plates. *Appl. Anal.*, **90**, 317–333.

Huber, M., Schoeberl, J., Sinwel, A. & Zaglmayr, S. (2009) Simulation of diffraction in periodic media with a coupled finite element and plane wave approach. *SIAM J. Sci. Comput.*, **31**, 1500–1517.

Kamotski, I. V. & Nazarov, S. A. (2002) The augmented scattering matrix and exponentially decaying solutions of an elliptic problem in a cylindrical domain. *J. Math. Sci.*, **111**, 3657–3666.

Kleemann, B. (2002) Elektromagnetische Analyse von Oberflächengittern von IR bis XUV mittels einer parametrisierten Randintegralmethode: Theorie, Vergleich und Anwendungen. Dissertation, Technischen Universität Ilmenau.

Li, P., Wu, H. & Zheng, W. (2011) Electromagnetic scattering by unbounded rough surfaces. *SIAM J. Math. Anal.*, **43**, 1205–1231.

Monk, P. (2003) *Finite element method for Maxwell's equations*, Oxford: Oxford University Press.

Nédélec, J. C. & Starling, F. (1991) Integral equation methods in a quasi-periodic diffraction problem for the time-harmonic Maxwell equation. *SIAM J. Math. Anal.*, **22**, 1679–1701.

Nitsche, J. (1970) Über ein Variationsprinzip zur Lösung von Dirichlet-Problemen bei Verwendung von Teilräumen. *Abh. Math. Sem. Univ. Hamburg*, **36**, 9–15.

Rathsfeld, A. (2013) Shape derivatives for the scattering by biperiodic gratings. *Appl. Numer. Math.*, **72**, 19–32.

Schmidt, G. (2003) On the diffraction by biperiodic anisotropic structures. *Applicable Anal.*, **82**, 75–92.

Schmidt, G. (2004) Electromagnetic scattering by periodic structures. *J. Math. Sci.*, **124**, 5390–5405.

Sternberg, R (1988) Mortaring by a method of J. A. Nitsche, Computational Mechanics, New Trends and Applications (S. Idelsohn, E. Onate, E. Dvorkin eds). Spain: CIMNE Barcelona.

Turunen, J., Kuittinen, M. & Wyrowski, F. (2000) Diffractive optics: electromagnetic approach. *Progress in Optics XL* (E. Wolf ed.). Amsterdam: Elsevier Science B. V., pp. 343–388.

Turunen, J. & Wyrowski, F. (eds) (2003) *Diffractive Optics for Industrial and Commercial Applications*. Berlin: Akademie.

Appendix

For the reader's convenience, we prove that the variational formulation (2.7) is uniquely solvable for small wavenumbers $k > 0$. Since Rayleigh frequencies can be excluded for small wavenumbers, by

Remark 3.1 we see that such a unique solvability also applies to our variational formulation (3.6) provided k is sufficiently small.

LEMMA A.1 There is a $k_0 > 0$ such that the variational formulation (2.7) admits a unique solution $E \in X$ for all wavenumbers k with $0 < k \leqslant k_0$.

Proof. To prove Lemma A.1, we need to replace equation (2.7) on the k-dependent α-quasi-periodic space $H(\mathrm{curl}\,,\Omega_b)$ by an equivalent variational problem acting on the (Λ_1,Λ_2)-periodic Sobolev space. Introduce the spaces $H^1_p(\Omega_b)$, $H_p(\mathrm{curl}\,,\Omega_b)$, $H^s_{t,p}(\Gamma_b)$, $H^s_{t,p}(\mathrm{Div}\,,\Gamma_b)$ and $H^s_{t,p}(\mathrm{Curl}\,,\Gamma_b)$ in the same way as $H^1(\Omega_b)$, $H(\mathrm{curl}\,,\Omega_b)$, $H^s_t(\Gamma_b)$, $H^s_t(\mathrm{Div}\,,\Gamma_b)$ and $H^s_t(\mathrm{Curl}\,,\Gamma_b)$ in Section 2, but with $\alpha = (0,0)^\top$. Furthermore, define the operator $\nabla_\alpha := \nabla + \mathrm{i}(\alpha,0)^\top$ and, analogously to the space X defined in Section 2, set $D := \{F : \Omega_b \to \mathbb{C}^3, F \in H_p(\mathrm{curl}\,,\Omega_b), \nu \times F = 0 \text{ on } \Gamma\}$. Let $\tau_n := (2\pi n_1/\Lambda_1, 2\pi n_2/\Lambda_2)^\top = \alpha_n - \alpha$ for $n = (n_1,n_2)^\top \in \mathbb{Z}^2$. Given

$$\tilde{F}(x') = \sum_{n\in\mathbb{Z}^2} \tilde{F}_n\, \mathrm{e}^{\mathrm{i}\tau_n\cdot x'} \in H^{-1/2}_{t,p}(\mathrm{Div}\,,\Gamma_b), \tag{A.1}$$

the definition of the operator $\mathscr{R}$ (cf. (2.8)) implies $\mathscr{R}(\tilde{E}) = \mathscr{T}(\tilde{F})\exp(\mathrm{i}\alpha\cdot x')$ for $\tilde{E}(x') = \mathrm{e}^{\mathrm{i}\alpha\cdot x'}\tilde{F}(x') \in H^{-1/2}_t(\mathrm{Div}\,,\Gamma_b)$, where the operator $\mathscr{T} : H^{-1/2}_{t,p}(\mathrm{Div}\,,\Gamma_b) \to H^{-1/2}_{t,p}(\mathrm{Curl}\,,\Gamma_b)$ is the Dirichlet-to-Neumann map over the space $H^{-1/2}_{t,p}(\mathrm{Div}\,,\Gamma_b)$ defined by

$$(\mathscr{T}\tilde{F})(x') = -\sum_{n\in\mathbb{Z}^2} \frac{1}{\mathrm{i}\beta_n}[k^2\tilde{F}_n - (\alpha_n\cdot\tilde{F}_n)\alpha_n]\exp(\mathrm{i}\tau_n\cdot x'), n = (n_1,n_2)^\top \in \mathbb{Z}^2. \tag{A.2}$$

Note that $\mathscr{T}$ is well defined for small wavenumbers k with $0 < k \leqslant k_0$, since $\beta_n \neq 0$ if k_0 is sufficiently small. The spaces $H^{-1/2}_{t,p}(\Gamma_b), H^{-1/2}_{t,p}(\mathrm{Div}\,,\Gamma_b)$ and $H^{-1/2}_{t,p}(\mathrm{Curl}\,,\Gamma_b)$ will be equipped with the norms analogous to the quasi-periodic ones in Section 2, but with the coefficient E_n replaced by $\tilde{F}_n$ given in (A.1) and α_n replaced by τ_n.

Set $F^{\mathrm{in}}(x) := \exp(-\mathrm{i}\alpha\cdot x')E^{\mathrm{in}}(x)$, $F(x) = \exp(-\mathrm{i}\alpha\cdot x')E(x)$, as well as $\psi(x) = \exp(-\mathrm{i}\alpha\cdot x')\varphi(x) \in H_p(\mathrm{curl}\,,\Omega_b)$ for $E,\varphi \in H(\mathrm{curl}\,,\Omega_b)$. We now consider the variational formulation

$$\begin{aligned} a_p(F,\psi) &:= \int_{\Omega_b} [\nabla_\alpha \times F \cdot \nabla_\alpha \times \overline{\psi} - k^2 F\cdot\overline{\psi}]\,\mathrm{d}x - \int_{\Gamma_b} \mathscr{T}(e_3\times F)\cdot(e_3\times\overline{\psi})\,\mathrm{d}s \\ &= \int_{\Gamma_b} [(\nabla_\alpha\times F^{\mathrm{in}})_T - \mathscr{T}(e_3\times F^{\mathrm{in}})]\cdot(e_3\times\overline{\psi})\,\mathrm{d}s \end{aligned} \tag{A.3}$$

for all $\psi \in D$, which is the counterpart of problem (2.7) in the periodic space $H_p(\mathrm{curl}\,,\Omega_b)$.

The problem (A.3) can be rewritten as the operator equation $B(F) = f$ in the dual space D' of D, where for $\psi \in D$ the dualities $\langle B(F),\psi\rangle$ and $\langle f,\psi\rangle$ between D' and D are defined by the sesquilinear form $a_p(F,\psi)$ and the right hand of (A.3), respectively. By Lemma 4.1, we have the Hodge-decomposition $D = D_0 \oplus D_1$, with the two subspaces $D_1 := \{\nabla_\alpha q : q \in H^1_p(\Omega_b), q = 0 \text{ on } \Gamma\}$ and $D_0 := \{F_0 \in D : \int_{\Omega_b} \nabla_\alpha q\cdot\overline{F}_0\,\mathrm{d}x = 0 \forall\ \nabla_\alpha q \in D_1\}$. This allows the decompositions $F = F_0 + \nabla_\alpha q$ and $\psi = G_0 + \nabla_\alpha g$ with $F_0, G_0 \in D_0$ and $\nabla_\alpha q, \nabla_\alpha g \in D_1$. Now, the sesquilinear form a_p in (A.3) can be written as

$$a_p(F,\psi) = a_p(F_0,G_0) + a_p(\nabla_\alpha q, G_0) + a_p(\nabla_\alpha q, \nabla_\alpha g) + a_p(F_0, \nabla_\alpha g),$$

and operator B takes the form

$$B = \begin{pmatrix} B_1 & B_3 \\ B_2 & B_4 \end{pmatrix}, \quad \begin{array}{lll} B_1 : D_0 \to D_0', & \langle B_1(F_0), G_0\rangle = a_p(F_0,G_0), & \forall G_0 \in D_0, \\ B_2 : D_0 \to D_1', & \langle B_2(F_0), \nabla_\alpha g\rangle = a_p(F_0,\nabla_\alpha g), & \forall \nabla_\alpha g \in D_1, \\ B_3 : D_1 \to D_0', & \langle B_3(\nabla_\alpha q), G_0\rangle = a_p(\nabla_\alpha q, G_0), & \forall G_0 \in D_0, \\ B_4 : D_1 \to D_1', & \langle B_4(\nabla_\alpha q), \nabla_\alpha g\rangle = a_p(\nabla_\alpha q, \nabla_\alpha g), & \forall \nabla_\alpha g \in D_1. \end{array}$$

We first prove that the form a_p is coercive over D_0 for a small wavenumber k. Using the explicit representation of $\mathscr{T}$, we obtain, for $\tilde{F}$ given in (A.1),

$$\mathrm{Re} \int_{\Gamma_b} \mathscr{T}(\tilde{F}) \cdot \overline{\tilde{F}}\, \mathrm{d}s = \Lambda_1 \Lambda_2 \sum_{n:|\alpha_n|>k} \frac{1}{\sqrt{|\alpha_n|^2 - k^2}} [k^2 |\tilde{F}_n|^2 - |\alpha_n \cdot \tilde{F}_n|^2],$$

$$-\mathrm{Re} \int_{\Gamma_b} \mathscr{T}(\tilde{F}) \cdot \overline{\tilde{F}}\, \mathrm{d}s \geqslant -\Lambda_1 \Lambda_2 \sum_{n:|\alpha_n|>k} \frac{1}{\sqrt{|\alpha_n|^2 - k^2}} k^2 |\tilde{F}_n|^2 \geqslant -C_1 \Lambda_1 \Lambda_2 \sum_{n\in\mathbb{Z}^2} \frac{1}{\sqrt{|\tau_n|^2+1}} k^2 |\tilde{F}_n|^2$$
$$= -C_1 \Lambda_1 \Lambda_2 k^2 \|\tilde{F}\|^2_{H_{t,p}^{-1/2}(\Gamma_b)} \geqslant -C_1 \Lambda_1 \Lambda_2 k^2 \|\tilde{F}\|^2_{H_{t,p}^{-1/2}(\mathrm{Div}\,,\Gamma_b)}.$$

Applying the previous estimate to the trace $e_3 \times F_0$ for $F_0 \in D_0$ and using the continuity of the trace mapping from $H_p(\mathrm{curl}\,, \Omega_b)$ to $H_{t,p}^{-1/2}(\mathrm{Div}\,, \Gamma_b)$, we arrive at

$$-\mathrm{Re} \left\{ \int_{\Gamma_b} \mathscr{T}(e_3 \times F_0) \cdot (e_3 \times \overline{F}_0)\, \mathrm{d}s \right\} \geqslant -k^2 C_1 \Lambda_1 \Lambda_2 \|e_3 \times F_0\|^2_{H_{t,p}^{-1/2}(\mathrm{Div}\,,\Gamma_b)} \geqslant -k^2 C_2 \|F_0\|^2_{H(\mathrm{curl}\,,\Omega_b)^3}.$$

Therefore,

$$\mathrm{Re}\, a_p(F_0,F_0) \geqslant \|\nabla_\alpha \times F_0\|^2_{L^2(\Omega_b)^3} - k^2 \|F_0\|^2_{L^2(\Omega_b)^3} - k^2 C_2 \|F_0\|^2_{H(\mathrm{curl}\,,\Omega_b)^3}. \tag{A.4}$$

Recalling that the function $E_0 := \exp(\mathrm{i}\alpha \cdot x')F_0$ belongs to the space $X_0 \subset X$ which is divergence free, we have the Friedrichs-type estimate $\|E_0\|^2_{L^2(\Omega_b)^3} \leqslant C_3 \|\nabla \times E_0\|^2_{L^2(\Omega_b)^3}$ (cf. Monk, 2003, Cor. 4.8) for some constant $C_3 > 0$ independent of $E_0 \in X_0$, which is equivalent to

$$\|F_0\|^2_{L^2(\Omega_b)^3} \leqslant C_3 \|\nabla_\alpha \times F_0\|^2_{L^2(\Omega_b)^3}. \tag{A.5}$$

Combining (A.4) and (A.5) leads to the coercivity of the form a_p over D_1 for small wavenumbers $k < k_0$. This implies that the operator B_1^{-1} exists with the bounded norm $\|B_1^{-1}\|_{D_0 \to D_0'} \leqslant C$ for some constant $C > 0$ independent of k with $0 < k \leqslant k_0$.

Next we claim that the form $-a_p$ is also coercive over D_1. In fact, the function $Q(x')$, given by $(Q(x'),0)^\top := e_3 \times \nabla_\alpha g|_{\Gamma_b}$, can be expanded into

$$Q(x') = \sum_{n\in\mathbb{Z}^2} (-\alpha_n^{(2)}, \alpha_n^{(1)})^\top Q_n \exp(\mathrm{i}\tau_n \cdot x'), \quad Q_n \in \mathbb{C}. \tag{A.6}$$

Thus, using the representation of $\mathscr{T}$ given in (A.2), we find

$$-\mathrm{Re}\, a_p(\nabla_\alpha q, \nabla_\alpha q) = k^2 \|\nabla_\alpha q\|^2_{L^2(\Omega_b)^3} + 4\pi^2 \sum_{n:|\alpha_n|>k} |\beta_n|^{-1} k^2 \|\alpha_n\|^2 |Q_n|^2 \geqslant C_0 k^2 \|\nabla_{\alpha y} q\|^2_{H(\mathrm{curl}\,,\Omega_b)^3}.$$

As a consequence, we have $\|B_4^{-1}\|_{D_1 \to D_1'} \leqslant k^{-2} C_0^{-1}$, where the constant C_0 does not depend on k.

The operator B can be written as the matrix operator

$$B = \begin{pmatrix} B_1 & 0 \\ B_2 & B_4 \end{pmatrix} + \begin{pmatrix} 0 & B_3 \\ 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} B_1 & 0 \\ B_2 & B_4 \end{pmatrix}^{-1} = \begin{pmatrix} B_1^{-1} & 0 \\ -B_4^{-1} B_2 B_1^{-1} & B_4^{-1} \end{pmatrix} =: \mathscr{M}.$$

Thus the operator equation $B(F) = f$ is equivalent to

$$\left[\begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix} + \begin{pmatrix} 0 & B_1^{-1} B_3 \\ 0 & -B_4^{-1} B_2 B_1^{-1} B_3 \end{pmatrix} \right] \begin{pmatrix} F_0 \\ \nabla_\alpha g \end{pmatrix} = \mathscr{M} f, \tag{A.7}$$

where I denotes the identity operator. To prove the invertibility of B, it suffices to show

$$\|B_3\|_{D_1 \to D_0'} + \|B_2\|_{D_0 \to D_1'} \leqslant C_4 k^2, \tag{A.8}$$

with a $C_4 > 0$ independent of $k \in (0, k_0]$. Consider the sesquilinear form corresponding to B_2:

$$a_p(F_0, \nabla_\alpha g) = -\int_{\Gamma_b} \mathscr{T}(e_3 \times F_0) \cdot (e_3 \times \overline{\nabla_\alpha g})\, \mathrm{d}s.$$

Expand the first two components of $e_3 \times F_0$, $e_3 \times \overline{\nabla_\alpha g}$ into the series in (A.1) and (A.6), respectively. Then, by (A.2) we obtain

$$\begin{aligned} |a_p(F_0, \nabla_{\alpha y} g)| &= k^2 \left| \sum_{n \in \mathbb{Z}^2} \frac{1}{\mathrm{i}\beta_n} \tilde{F}_n \cdot (-\alpha_n^{(2)}, \alpha_n^{(1)})^\top \overline{Q}_n \right| \\ &\leqslant C_5 k^2 \left(\sum_{n \in \mathbb{Z}^2} (1 + |\tau_n|^2)^{1/2} |Q_n|^2 \right)^{1/2} \left(\sum_{n \in \mathbb{Z}^2} (1 + |\tau_n|^2)^{-1/2} |\tilde{F}_n|^2 \right)^{1/2} \\ &\leqslant C_6 k^2 \|e_3 \times \nabla_\alpha g\|_{H_{t,p}^{-1/2}(\mathrm{Div}\,,\Gamma_b)} \|e_3 \times F_0\|_{H_{t,p}^{-1/2}(\mathrm{Div}\,,\Gamma_b)}. \end{aligned}$$

This combined with the continuity of the trace mapping from $H_p(\mathrm{curl}\,,\Omega_b)$ to $H_{t,p}^{-1/2}(\mathrm{Div}\,,\Gamma_b)$ leads to the estimate in (A.8) for B_2. For the proof concerning B_3, we can proceed analogously.

We now conclude that the sesquilinear form corresponding to the operator on the left-hand side of (A.7) is positive definite for small wavenumbers. Indeed, the operator is a small perturbation of the identity for all $k < k_0$ if k_0 is sufficiently small. Hence, problem (3.6) always admits a unique solution E of the form $E = \exp(\mathrm{i}\alpha \cdot x')F$ with $F = F_0 + \nabla_\alpha q$, $F_0 \in D_0$, $\nabla_\alpha q \in D_1$. $\square$