ELSEVIER

Journal of Computational and Applied Mathematics 150 (2003) 157–192

JOURNAL OF COMPUTATIONAL AND APPLIED MATHEMATICS

www.elsevier.com/locate/cam

Integral equations via saddle point problems for time-harmonic Maxwell's equations

Francis Collino[a], Bruno Despres[b,*]

[a] *Formerly at INRIA, Domaine de Rocquencourt 78153 Le Chesnay, France and CERFACS, 42 avenue G. Coriolis 31057 Toulouse, France*

[b] *Commissariat à l'Energie Atomique, BP12, 91680, Bruyeres le Chatel, Cedex, and Laboratoire Jacques Louis Lions, Universite Pierre et Marie Curie, 175 rue Chevaleret, 75013, Paris, France*

Received 21 March 2001; received in revised form 11 April 2002

Abstract

We propose a new system of integral equations for the exterior time harmonic Maxwell's equation. This system is derived first from elementary manipulations of classical equations then by the minimization of a quadratic functional associated to incoming and outgoing electromagnetic waves. We analyze the inf–sup condition and various penalized problems related to this system. Then we prove that an iterative algorithm for the solution of the system of integral equations is convergent. Other numerical issues are also discussed.

Keywords: Maxwell; Integral equation; Saddle point; Scattering; Impedance boundary conditions; Iterative algorithm

1. Introduction

We propose and study a new system of integral equations for solving obstacle scattering by time-harmonic electromagnetic waves, with a particular emphasis on impedance boundary conditions: see [2,4,1] for some recent works about scattering with impedance boundary conditions. We think that the method presented in this work could have some advantages from the computational point of view when compared to more classical integral equations like electric field integral equation (EFIE) or magnetic or combined field integral equation (MFIE,CFIE), [9,20]. For instance, the particular algebraic structure of our system allows us to obtain *convergence theorems for iterative algorithms* (we give a proof for one of them in this paper), which is known to be difficult to get with classical integral equations. This difficulty is not only theoretical but also practical since it is

* Corresponding author.

PII: S0377-0427(02)00658-1

known that the lack of coercivity of classical integral equations can result in poor convergence rates for iterative methods, at least for general geometries and in general situations, [21,23,19]. In this work, we show that it is possible to solve scattering problems by means of a system of integral equations with a positive real spectrum. This kind of system has been derived in [12,13], in the 2-D acoustic, in [11] for Maxwell's equations and in [27,26] for Maxwell's equations coupled with a domain decomposition strategy in a bounded domain. However, the approach presented here is expected to be less complicated as it uses more general arguments than in those original papers. The idea of the method is to construct a positive quadratic functional whose minimum is the outgoing electromagnetic wave satisfying a given boundary condition. This functional is defined on a set of outgoing and incoming electromagnetic waves. We show that this set can be parameterized by means of electric and magnetic currents plus a kernel of an electromagnetic Herglotz wave pair, [9], these three parameters satisfying a set of linear constraints. The new system of integral equations is obtained by writing down the optimality conditions of the associated Lagrangian. It can also be obtained by some direct manipulations of classical integral operators for electromagnetism.

One interesting feature of the new system of equations, comparing to the classical ones, is that the space of solutions is L^2. A penalization procedure may be used to get more coercivity on the multiplier, even if the inf–sup condition is already true for the nonpenalized formulation. This coercive framework might appear as unusual when compared with the standard theory of integral operators for electromagnetism. However, the proof of well-posedness of the weak formulation that we give is based on some well known fundamental properties of the exterior Calderon projectors for Maxwell's equations. Let us mention that in previous work [12] most of the properties of the coercive formulation were derived from the analysis of special functions related to Helmholtz equations and not via Calderon projectors.

The outline of the paper is the following. First, we derive the new system by simple manipulations of classical integral operators for electromagnetism. Then, we introduce the space $\mathscr{W}$ of incoming and outgoing solutions of Maxwell's equations, on which we define a suitable quadratic functional depending on a boundary data. We prove that the minimum of this functional is reached at a point that coincides with the outgoing solution of Maxwell's equation satisfying a boundary condition linked to the data boundary. Using a parameterization of $\mathscr{W}$ by means of currents, we show that the system has an interpretation as a minimization problem with constraints. Introducing a Lagrange multiplier and the associated Lagrangian, the optimality conditions allow us to recover the system previously obtained. Then, we first discuss well-posedness of the system. Second, we introduce a penalization to get more tractable problems from the numerical point of view. We also prove well-posedness of the penalized system. Finally, we propose an iterative algorithm to solve the system and prove its convergence. We end up with a discussion on some numerical issues. In order to give a better understanding of the new system, we analytically determine, in Appendix C, the spectrum of our integral operator in the special case of a spherical scatterer.

2. A first derivation of the integral equation system

Let $D^- \subset R^3$ be a bounded domain with regular boundary Γ (at least of class C^2), and let $n(x)$ be the unit normal vector to Γ directed into $D^+ \subset R^3$, the exterior domain of D^-. The problem we address is the determination of the outgoing electromagnetic field scattered by Γ. Let $k > 0$ be the wavenumber and Z_0 the vacuum impedance. The field is a solution to (1), (2) where

$$\nabla \wedge E^+ - \mathrm{i}kZ_0H^+ = 0 \quad \text{in } D^+,$$
$$\nabla \wedge H^+ + \mathrm{i}kZ_0^{-1}E^+ = 0 \quad \text{in } D^+ \tag{1}$$

is the Maxwell system and

$$\lim_{|x|\to\infty} |x| \left(Z_0H^+ \wedge \frac{x}{|x|} - E^+ \right) = 0 \tag{2}$$

is the Silver–Müller radiation condition at infinity. System (1)–(2) must be supplemented by a boundary condition on Γ. Let us assume for now that the boundary condition is an absorbing boundary condition

$$n(x) \wedge (E^+_{/\Gamma}(x) \wedge n(x)) + Z_0(H^+_{/\Gamma}(x) \wedge n(x)) = G^{\mathrm{in}}. \tag{3}$$

Other boundary conditions will be considered in Section 5. It is well known (e.g. Theorem 6.6 in [9]) that the determination of the solution (E^+, H^+) amounts to the knowledge of the equivalent currents

$$J = n \wedge H^+_{/\Gamma}, \quad M = -n \wedge E^+_{/\Gamma}, \tag{4}$$

through the relations

$$E^+(x) = \mathrm{i}Z_0\tilde{T}J(x) + \tilde{K}M(x),$$
$$H^+(x) = -\tilde{K}J(x) + \mathrm{i}Z_0^{-1}\tilde{T}M(x), \tag{5}$$

where $\tilde{T}$ and $\tilde{K}$ are the classical potential operators defined by

$$\tilde{T}J(x) = \frac{1}{k}\nabla_x \wedge \left(\nabla_x \wedge \int_\Gamma G(x, y)J(y)\,\mathrm{d}\Gamma(y) \right),$$
$$\tilde{K}J(x) = -\nabla_x \wedge \int_\Gamma G(x, y)J(y)\,\mathrm{d}\Gamma(y), \tag{6}$$

or, in another form,

$$\tilde{T}J(x) = k\int_\Gamma \left(G(x, y)J(y) + \frac{1}{k^2}\nabla_x G(x, y)\nabla^t \cdot J(y) \right) \mathrm{d}\Gamma(y),$$
$$\tilde{K}J(x) = \int_\Gamma \nabla_y G(x, y) \wedge J(y)\,\mathrm{d}\Gamma(y). \tag{7}$$

Here $\nabla^t \cdot J(y)$ denotes the surface divergence of J. The kernel $G(x, y)$ is the radiating Green function for the 3-D Helmholtz equation

$$G(x, y) = \frac{\exp^{\mathrm{i}k|x-y|}}{4\pi|x - y|}. \tag{8}$$

The same definitions apply to $\tilde{T}M$ and $\tilde{K}M$. When x in D^+ approaches the boundary Γ, it is well-known that the potentials $\tilde{T}J(x)$ and $\tilde{T}M(x)$ possess continuous tangential traces on Γ while the tangential components of $\tilde{K}J(x)$ and $\tilde{K}M(x)$ have a jump across the boundary. In particular, we have

$$n(x) \wedge (E^+_{/\Gamma}(x) \wedge n(x)) = \mathrm{i}Z_0 TJ(x) + KM(x) + \tfrac{1}{2}n(x) \wedge M(x),$$
$$n(x) \wedge (H^+_{/\Gamma}(x) \wedge n(x)) = -KJ(x) + \mathrm{i}Z_0^{-1}TM - \tfrac{1}{2}n(x) \wedge J(x), \tag{9}$$

where T and K are defined by

$$TJ(x) = \lim_{y \to x} n(x) \wedge (\tilde{T}J(y)) \wedge n(x)),$$
$$KJ(x) = \left(\int_\Gamma n(x) \wedge (\nabla_y G(x, y) \wedge J(y))\, \mathrm{d}\Gamma(y) \right) \wedge n(x). \tag{10}$$

Using expressions (4) that relate fields and currents, we obtain

$$0 = \mathrm{i}Z_0(TJ)(x) + (KM)(x) - \tfrac{1}{2}n(x) \wedge M(x),$$
$$0 = (KJ)(x) - \tfrac{1}{2}n(x) \wedge J(x) - \mathrm{i}Z_0^{-1}(TM)(x), \tag{11}$$

or, in a matrix form,

$$\mathbf{Su} = 0, \tag{12}$$

where

$$\mathbf{S} = \begin{bmatrix} T & K - \frac{1}{2}n\wedge \\ K - \frac{1}{2}n\wedge & T \end{bmatrix}, \tag{13}$$

and

$$\mathbf{u} = \begin{bmatrix} J_1 \\ M_1 \end{bmatrix} = \begin{bmatrix} \sqrt{\mathrm{i}Z_0}J \\ \sqrt{\mathrm{i}Z_0}^{-1}M \end{bmatrix}. \tag{14}$$

Eq. (12) is a compatibility condition and gives profound insights into electromagnetic behavior. It expresses the fact that not all pairs of tangential fields on Γ are composed of tangential traces of radiating electromagnetic field. To be such a trace, it is necessary (and sufficient, actually) that the pair lies in the kernel of the integral operator $\mathbf{S}$.

All previous calculations are completely standard. To go further, we introduce the decomposition in real and imaginary part of the Green function $G(x, y)$,

$$G(x, y) = \frac{\cos(k|x - y|)}{4\pi|x - y|} + \mathrm{i}\,\frac{\sin(k|x - y|)}{4\pi|x - y|} = G_r(x, y) + \mathrm{i}G_i(x, y). \tag{15}$$

We obtain a similar decomposition for $\mathbf{S}$ in real and imaginary parts

$$\mathbf{S} = \mathbf{T} + \mathrm{i}\mathbf{R}, \tag{16}$$

with

$$\mathbf{T}=\begin{bmatrix} T_r & K_r-\frac{1}{2}n\wedge \\ K_r-\frac{1}{2}n\wedge & T_r \end{bmatrix}, \quad \mathbf{R}=\begin{bmatrix} T_i & K_i \\ K_i & T_i \end{bmatrix}. \tag{17}$$

At this point, we introduce the new unknown

$$\mathbf{v}=\mathrm{i}\mathbf{u}=\begin{bmatrix} \mathrm{i}J_1 \\ \mathrm{i}M_1 \end{bmatrix}=\begin{bmatrix} J_2 \\ M_2 \end{bmatrix}=\begin{bmatrix} \mathrm{i}\sqrt{\mathrm{i}Z_0}J \\ \mathrm{i}\sqrt{\mathrm{i}Z_0}^{-1}M \end{bmatrix}, \tag{18}$$

allowing us to read $\mathbf{Su}=\mathbf{Tu}+\mathrm{i}\mathbf{Ru}=0$ as

$$-\mathbf{Tu}-\mathbf{Rv}=0, \tag{19}$$

or, in an equivalent way,

$$\mathbf{Tv}-\mathbf{Ru}=0. \tag{20}$$

The operators T_r and K_r are symmetric, while $n(x)\wedge$ is antisymmetric. If $\mathbf{T}^*$ is the adjoint of $\mathbf{T}$, and

$$\Pi=\begin{bmatrix} 0 & -n(x)\wedge \\ -n(x)\wedge & 0 \end{bmatrix}, \tag{21}$$

we have

$$\mathbf{T}-\mathbf{T}^*=\Pi. \tag{22}$$

The operators T_i and K_i are symmetric and regularizing ($G_i(x,y)$ is a smooth symmetric kernel); so, $\mathbf{R}$ is as a symmetric and regularizing operator. Now, we derive the following decomposition of $\mathbf{R}$. If $\hat{d}$ is a given direction on the unit sphere S^2 and if $(J(x),M(x))$ are two given tangential fields on Γ, we define the far field operators $\mathbf{a}^\infty$ and $\mathbf{A}^\infty$ by

$$\mathbf{a}^\infty J(\hat{d})=\frac{k}{4\pi}\int_\Gamma \hat{d}\wedge(J(x)\wedge\hat{d})\mathrm{e}^{-\mathrm{i}kx\cdot\hat{d}}\,\mathrm{d}\Gamma(x), \tag{23}$$

$$\mathbf{A}^\infty:\begin{bmatrix} J(x) \\ M(x) \end{bmatrix}\to\mathbf{A}^\infty\begin{bmatrix} J \\ M \end{bmatrix}(\hat{d})=\mathbf{a}^\infty J(\hat{d})-\mathrm{i}\hat{d}\wedge\mathbf{a}^\infty M(\hat{d}). \tag{24}$$

Using the integral identity

$$\frac{\sin k|\eta|}{|\eta|}=\frac{k}{4\pi}\int_{S^2}\mathrm{e}^{-\mathrm{i}k\eta\cdot\hat{d}}\,\mathrm{d}\sigma(\hat{d}), \qquad \eta\in\mathbf{R}^3, \tag{25}$$

we prove in Appendix A the following factorization:

$$\mathbf{R}=(\mathbf{A}^\infty)^*\mathbf{A}^\infty, \tag{26}$$

i.e.,

$$\left(\mathbf{R}\begin{bmatrix} J \\ M \end{bmatrix}, \begin{bmatrix} J' \\ M' \end{bmatrix}\right)_{TL^2(\Gamma)\times TL^2(\Gamma)} = \left(\mathbf{A}^\infty\begin{bmatrix} J \\ M \end{bmatrix}, \mathbf{A}^\infty\begin{bmatrix} J' \\ M' \end{bmatrix}\right)_{TL^2(S^2)}.$$

Here $TL^2(\Gamma)$ (resp. $TL^2(S^2)$) denotes the set of tangential fields on Γ (resp. on S^2),

$$TL^2(\Gamma) = \{\varphi \in L^2(\Gamma)^3,\ \varphi.n = 0)\}.$$

A similar definition applies to $TL^2(S^2)$.

Eq. (26) shows that $\mathbf{R}$ is a symmetric positive operator. It allows us to rewrite (19) as

$$-\mathbf{T}\mathbf{u} - (\mathbf{A}^\infty)^*\mathbf{A}^\infty\mathbf{v} = 0. \tag{27}$$

Eq. (27) will be the first equation of our system of integral equations.

So far, we have not used the boundary condition on Γ. As mentioned before, we assume that the electromagnetic field satisfies an absorbing boundary condition of the type

$$n(x)\wedge(E^+_{/\Gamma}(x)\wedge n(x)) + Z_0(H^+_{/\Gamma}(x)\wedge n(x)) = G^{\text{in}}, \tag{28}$$

where G^{in} is some incoming field (source term) on Γ. From (4) and (14), we have the two equivalent relations

$$J_1(x) - \mathrm{i}n(x)\wedge M_1(x) = -\mathrm{i}\frac{1}{\sqrt{\mathrm{i}Z_0}}G^{\text{in}}(x),$$
$$-\mathrm{i}n(x)\wedge J_1(x) + M_1(x) = -n(x)\wedge\frac{1}{\sqrt{\mathrm{i}Z_0}}G^{\text{in}}(x). \tag{29}$$

Defining

$$\mathbf{g} = \begin{bmatrix} -\mathrm{i}\frac{1}{\sqrt{\mathrm{i}Z_0}}G^{\text{in}}(x) \\ -n(x)\wedge\dfrac{1}{\sqrt{\mathrm{i}Z_0}}G^{\text{in}}(x) \end{bmatrix} \tag{30}$$

and using Definitions (14) and (18) for $\mathbf{u}$ and $\mathbf{v}$, we get

$$\mathbf{u} - \begin{bmatrix} 0 & n(x)\wedge \\ n(x)\wedge & 0 \end{bmatrix}\mathbf{v} = \mathbf{g}, \tag{31}$$

or, using (22),

$$\mathbf{u} + \Pi\mathbf{v} = \mathbf{u} + \mathbf{T}\mathbf{v} - \mathbf{T}^*\mathbf{v} = \mathbf{g}. \tag{32}$$

We use (20) to obtain

$$\mathbf{u} + \mathbf{R}\mathbf{u} - \mathbf{T}^*\mathbf{v} = \mathbf{g}. \tag{33}$$

We join (33), with the factorization (26) of $\mathbf{R}$, to Eq. (27) to obtain the final system

$$\mathbf{u} + (\mathbf{A}^\infty)^*\mathbf{A}^\infty\mathbf{u} - \mathbf{T}^*\mathbf{v} = \mathbf{g},$$
$$-\mathbf{T}\mathbf{u} - (\mathbf{A}^\infty)^*\mathbf{A}^\infty\mathbf{v} = 0. \tag{34}$$

This will constitute the main integral system we shall discuss in this work. System (34) differs from the other classical integral equations by involving only real operators and being symmetric. Furthermore, it can be seen as the optimality conditions of a saddle point problem. Let us assume for example that $\mathbf{T}$ is a continuous operator on $TL^2(\Gamma)^2$. If we define the spaces

$$\mathbf{U} = TL^2(\Gamma) \times TL^2(\Gamma), \quad Z = TL^2(S^2), \tag{35}$$

and the Lagrangian

$$\mathscr{L}(\mathbf{u}^*, \gamma^*, \mathbf{v}^*) = \tfrac{1}{2}\|\mathbf{u}^*\|_{\mathbf{U}}^2 + \tfrac{1}{2}\|\mathbf{A}^\infty\mathbf{u}^*\|_Z^2 + \tfrac{1}{2}\|\gamma^*\|_Z^2$$
$$+ \Re e < \mathbf{T}\mathbf{u}^* - \mathrm{i}(\mathbf{A}^\infty)^*\gamma^*, \quad \mathbf{v}^* >_{\mathbf{U}} - \Re e < \mathbf{g}, \ \mathbf{u}^* >_{\mathbf{U}}, \tag{36}$$

where $(\mathbf{u}^*, \mathbf{v}^*) \in \mathbf{u}$ and $\gamma \in Z$, then, the optimality conditions of the saddle point problem

$$\mathscr{L}(\mathbf{u}, \mathbf{v}, \gamma) = \min_{\mathbf{u}^*, \gamma^*} \max_{\mathbf{v}^*} \mathscr{L}(\mathbf{u}^*, \gamma^*, \mathbf{v}^*), \tag{37}$$

are nothing but

$$\mathbf{u} + (\mathbf{A}^\infty)^*\mathbf{A}^\infty\mathbf{u} - \mathbf{T}^*\mathbf{v} = \mathbf{g},$$
$$\gamma + \mathrm{i}\mathbf{A}^\infty\mathbf{v} = 0,$$
$$-\mathbf{T}\mathbf{u} - \mathrm{i}(\mathbf{A}^\infty)^*\gamma = 0. \tag{38}$$

In other words, $(\mathbf{u}, \mathbf{v})$ is a solution of (34) and

$$\gamma = \mathrm{i}\mathbf{A}^\infty\mathbf{v}. \tag{39}$$

3. A second derivation of the system of equations

The particular structure of system (34) gives some hints that it might be possible to derive it through the minimization of some quadratic functional: actually, this system was first obtain through this procedure. This section is devoted to the association of (32) with the corresponding positive quadratic functional.

3.1. Definition of incoming and outgoing electromagnetic fields

The idea is to consider a space of incoming and outgoing electromagnetic fields on D^+. We assume D^+ to be a domain of class C^2.

Definition 1. We define $\mathscr{W}$, the space of the fields (E,H) on D^+ such that

- (E,H) is in $(L^2_{\text{loc}}(D^+))^3 \times (L^2_{\text{loc}}(D^+))^3$
- (E,H) satisfies to Maxwell's equations

$$\begin{aligned} &\nabla \wedge E - \mathrm{i}kZ_0 H = 0 \quad \text{in } D^+ \\ &\nabla \wedge H + \mathrm{i}kZ_0^{-1} E = 0 \quad \text{in } D^+. \end{aligned} \tag{40}$$

- The tangential traces of (E,H) exist in $TL^2(\Gamma)$, i.e. the electromagnetic field possess tangential traces on Γ of square integrable modulus.
- (E,H) has the asymptotic behavior at infinity.

$$\lim_{R\to\infty} \frac{1}{R} \int_{D_R^+} (|E|^2(x) + Z_0|H|^2(x))\,\mathrm{d}x < \infty, \tag{41}$$

where

$$D_R^+ = \{|x| \leqslant R,\ x \in D^+\}.$$

Note that every solution (E,H) of Maxwell's equations in D^+ is regular (it has analytic Cartesian components) at points far enough from the boundary Γ.

All these properties are usual, except the behavior at infinity. It is given in the following lemma.

Lemma 1. *Let (E,H) be in $\mathscr{W}$, then there exists two fields $a_\infty^{\text{out}}(E,H)$ and $a_\infty^{\text{in}}(E,H)$ in $Z = TL^2(S^2)$ such that, if we define*

$$E^\infty(x) = \frac{e^{\mathrm{i}k|x|}}{|x|} a_\infty^{\text{out}}(E,H;\hat{x}) + \frac{e^{-\mathrm{i}k|x|}}{|x|} a_\infty^{\text{in}}(E,H;\hat{x}),$$

$$Z_0 H(x) = \frac{e^{\mathrm{i}k|x|}}{|x|} (\hat{x} \wedge a_\infty^{\text{out}}(E,H;\hat{x})) - \frac{e^{-\mathrm{i}k|x|}}{|x|} (\hat{x} \wedge a_\infty^{\text{in}}(E,H;\hat{x})) \tag{42}$$

where $\hat{x} = x/|x|$, then

$$\lim_{R\to\infty} \frac{1}{R} \int_{R \leqslant |x| \leqslant 2R} |E(x) - E^\infty(x)|^2\,\mathrm{d}x = 0, \tag{43}$$

$$\lim_{R\to\infty} \frac{1}{R} \int_{R \leqslant |x| \leqslant 2R} |H(x) - H^\infty(x)|^2\,\mathrm{d}x = 0. \tag{44}$$

The proof of this lemma is postponed to Section 3.4 (see Lemma 4). The field $a_\infty^{\text{out}}(E,H;\hat{x})$ can be seen as the far field for the outgoing part of the electromagnetic field: $a_\infty^{\text{in}}(E,H;\hat{x})$ corresponds to the far field for the incoming part of the electromagnetic field.

3.2. Definition of the quadratic functional; minimization

We define for (E,H) in $\mathscr{W}$,

$$G_\Gamma^{\rm in} \equiv G_\Gamma^{\rm in}(E,H) = n(x) \wedge (E_{/\Gamma} \wedge n(x)) + Z_0 H_{/\Gamma} \wedge n(x),$$

$$G_\Gamma^{\rm out} \equiv G_\Gamma^{\rm out}(E,H) = -n(x) \wedge (E_{/\Gamma} \wedge n(x)) + Z_0 H_{/\Gamma} \wedge n(x), \tag{45}$$

and consider a given tangential field $G^{\rm in}$ in $TL^2(\Gamma)$, which will play, as previously, the role of the right hand side of the (absorbing) boundary condition on Γ: $G^{\rm in}$ is therefore a data. We then introduce the following functional:

$$\begin{aligned} I(E,H) = {} & \tfrac{1}{4}\|G_\Gamma^{\rm in}(E,H)\|^2 + \tfrac{1}{4}\|G_\Gamma^{\rm out}(E,H)\|^2 \\ & + \|a_\infty^{\rm out}(E,H)\|_Z^2 + \|a_\infty^{\rm in}(E,H)\|_Z^2 - \Re e(G_\Gamma^{\rm in}(E,H), G^{\rm in}), \end{aligned} \tag{46}$$

where $(.,.)$ and $\|.\|$ denote, respectively, the inner product and the corresponding norm in $TL^2(\Gamma)$.

Theorem 1. *The minimum of $I(E,H)$ for (E,H) in $\mathscr{W}$ is reached by the solution of the following problem*

$$\begin{aligned} &\nabla \wedge E^+ - \mathrm{i}kZ_0 H^+ = 0 \quad \textit{in } D^+, \\ &\nabla \wedge H^+ + \mathrm{i}kZ_0^{-1} E^+ = 0 \quad \textit{in } D^+, \\ &n(x) \wedge (E^+_{/\Gamma}(x) \wedge n(x)) + Z_0(H^+_{/\Gamma}(x) \wedge n(x)) = G^{\rm in} \quad \textit{on } \Gamma, \\ &a_\infty^{\rm in}(E^+,H^+) = 0 \quad \textit{at infinity}. \end{aligned} \tag{47}$$

This result means in particular that it is possible to relax both the radiation condition at infinity and the boundary condition in the formulation of the problem, and to recover them through the minimization process. Note that the condition at infinity is treated exactly the same as the boundary condition on Γ. We will show later that minimizing the functional I amounts to solve system (33) derived in the first section.

The key point of the proof lies on the following isometry lemma, which is exactly equivalent to the unitarity of the scattering matrix in scattering theory [22].

Lemma 2 (Isometry lemma). *Let (E,H) be some electromagnetic field in $\mathscr{W}$. Then, the following equality holds*:

$$\tfrac{1}{4}\|G_\Gamma^{\rm in}(E,H)\|^2 + \|a_\infty^{\rm in}(E,H)\|_Z^2 = \tfrac{1}{4}\|G_\Gamma^{\rm out}(E,H)\|^2 + \|a_\infty^{\rm out}(E,H)\|_Z^2. \tag{48}$$

Proof. We introduce the truncated domain

$$D_r^+ = \{x \in D^+,\ |x| < r\},$$

where r is a large positive number. The boundary of D_r^+ splits into two parts: Γ (the interior boundary) and S_r^2 (the exterior spherical boundary):

$$\partial D_r^+ = \Gamma \cup S_r^2.$$

The outward normal is denoted by $v(x)$; $v(x) = -n(x)$ on Γ and $v(x) = \hat{x} = x/|x|$ on the spherical part of the boundary S_r^2.

Maxwell's equations (40) imply that

$$0 = \int_{D_r^+} ((\nabla \wedge E - \mathrm{i}kZ_0 H) \cdot \bar{H} - \overline{(\nabla \wedge H + \mathrm{i}kZ_0^{-1} E)} \cdot E)\, \mathrm{d}x.$$

Using the usual Stokes formula

$$\int_D (\nabla \wedge U \cdot V - \nabla \wedge V \cdot U)\, \mathrm{d}x = \int_{\partial D} (v \wedge (U_{/\partial D} \wedge v)) \cdot (V_{/\partial D} \wedge v)\, \mathrm{d}\Gamma, \tag{49}$$

one obtains

$$\int_{\partial D_r^+} ((v \wedge (E \wedge v)) \cdot \bar{H} \wedge v)\, \mathrm{d}\Gamma = \mathrm{i}k \int_{D_r^+} (Z_0 |H|^2 - Z_0^{-1} |E|^2)\, \mathrm{d}x.$$

Therefore we get taking the real part

$$\Re e \int_{\partial D_r^+} (v \wedge (E \wedge v)) \cdot (\bar{H} \wedge v)\, \mathrm{d}\Gamma = \Re e \int_{\Gamma \cup S_r^2} (v \wedge (E \wedge v)) \cdot (\bar{H} \wedge v)\, \mathrm{d}\Gamma = 0,$$

which is equivalent to

$$\int_{\Gamma \cup S_r^2} |-(v \wedge (E \wedge v)) + Z_0 (H \wedge v)|^2\, \mathrm{d}\Gamma = \int_{\Gamma \cup S_r^2} |+(v \wedge (E \wedge v)) + Z_0 (H \wedge v)|^2\, \mathrm{d}\Gamma.$$

We then take the mean value of this equality for r between R and $2R$ and let R goes to infinity. With the help of Lemma 1 and $\int_R^{2R} \mathrm{d}r \int_{S_r^2} \mathrm{d}\sigma = \int_{R \leqslant |x| \leqslant 2|R|} \mathrm{d}x$, we get

$$\begin{aligned}
&\lim_{R\to\infty} \frac{1}{R} \int_R^{2R} \int_{\Gamma \cup S_r^2} |-(v \wedge (E \wedge v)) + Z_0 (H \wedge v)|^2\, \mathrm{d}\Gamma\, \mathrm{d}r \\
&\quad = \int_\Gamma |-(n(x) \wedge (E_{/\Gamma} \wedge n(x))) - Z_0 (H(x) \wedge n(x)|^2\, \mathrm{d}\Gamma(x) \\
&\qquad + \lim_{R\to\infty} \frac{1}{R} \int_R^{2R} \mathrm{d}r \int_{S^2} r^2\, \mathrm{d}\sigma(\hat{d}) \left| -\frac{\mathrm{e}^{\mathrm{i}kr}}{r} a_\infty^{\mathrm{out}}(E, H; \hat{d}) - \frac{\mathrm{e}^{-\mathrm{i}kr}}{r} a_\infty^{\mathrm{in}}(E, H; \hat{d}) \right. \\
&\qquad \left. + \frac{\mathrm{e}^{\mathrm{i}kr}}{r} ((\hat{d} \wedge a_\infty^{\mathrm{out}}(E, H; \hat{d})) \wedge \hat{d}) - \frac{\mathrm{e}^{-\mathrm{i}kr}}{r} ((\hat{d} \wedge a_\infty^{\mathrm{in}}(E, H; \hat{d})) \wedge \hat{d}) \right|^2 \\
&\quad = \int_\Gamma |G_\Gamma^{\mathrm{in}}(x)|^2\, \mathrm{d}\Gamma(x) + 4 \int_{S^2} |a_\infty^{\mathrm{in}}(E, \hat{H}; d))|^2\, \mathrm{d}\sigma(\hat{d}).
\end{aligned}$$

We get in a same way,

$$\lim_{R\to\infty}\frac{1}{R}\int_R^{2R}\int_{\partial D_r^+}|+(v\wedge(E\wedge v))+Z_0(H\wedge v)|^2\,\mathrm{d}\Gamma\,\mathrm{d}r$$
$$=\int_\Gamma |G_\Gamma^{\mathrm{out}}(x)|^2\,\mathrm{d}\Gamma(x)+4\int_{S^2}|a_\infty^{\mathrm{out}}(E,\hat{H};d))|^2\,\mathrm{d}\sigma(\hat{d}). \tag{50}$$

This ends the proof of Lemma 1.

Proof (Theorem 1). Once the isometry lemma has been obtained, the minimization of $I(E,H)$ becomes obvious. Indeed, Definition (46) and Eq. (48) gives

$$I(E,H)=\tfrac{1}{2}\|G_\Gamma^{\mathrm{in}}(E,H)\|^2+2\|a_\infty^{\mathrm{in}}(E,H)\|_Z^2-\Re e(G_\Gamma^{\mathrm{in}}(E,H),G^{\mathrm{in}})$$
$$=\tfrac{1}{2}\|G_\Gamma^{\mathrm{in}}(E,H)-G^{\mathrm{in}}\|^2+2\|a_\infty^{\mathrm{in}}(E,H)\|_Z^2-\tfrac{1}{2}\|G^{\mathrm{in}}\|^2. \tag{51}$$

It is then clear that the minimum is $-\frac{1}{2}\|G^{\mathrm{in}}\|^2$ and is reached exactly for the electromagnetic field such that both the condition at infinity and the boundary condition are satisfied:

$$a_\infty^{\mathrm{in}}(E,H)=0,\quad G_\Gamma^{\mathrm{in}}(E,H)=G^{\mathrm{in}}. \tag{52}$$

Now, all the remaining difficulty is to choose an appropriate—and useful for practical computations—parameterization of space $\mathscr{W}$ of incoming and outgoing electromagnetic fields to derive the expressions of the related quantities G_Γ^{in}, G_Γ^{out}, a_∞^{in} and a_∞^{out}.

3.3. Representation of the $\mathscr{W}$-electromagnetic fields

Let (E,H) be in $\mathscr{W}$ and consider $(\tilde{E},\tilde{H})$ the extension by zero of (E,H) to D^-. It is classical to show that $(\tilde{E},\tilde{H})$ satisfies in the sense of distributions on $\mathbf{R}^3$

$$k^2\tilde{E}+\vec{\Delta}\tilde{E}=-\mathrm{i}kZ_0\left(J\delta_\Gamma+\frac{1}{k^2}\vec{\nabla}\vec{\nabla}\cdot(J\delta_\Gamma)\right)+\vec{\nabla}\wedge(M\delta_\Gamma),$$
$$k^2\tilde{H}+\vec{\Delta}\tilde{H}=-\frac{\mathrm{i}k}{Z_0}\left(M\delta_\Gamma+\frac{1}{k^2}\vec{\nabla}\vec{\nabla}\cdot(M\delta_\Gamma)\right)\Bigg)-\vec{\nabla}\wedge(J\delta_\Gamma),$$
$$\vec{\nabla}\cdot\tilde{E}=\frac{Z_0}{\mathrm{i}k}\vec{\nabla}\cdot(J\delta_\Gamma),\quad \vec{\nabla}\cdot\tilde{H}=+\frac{1}{Z_0\mathrm{i}k}\vec{\nabla}\cdot(M\delta_\Gamma), \tag{53}$$

where δ_Γ is the Dirac measure supported by Γ and

$$J=n\wedge H_{/\Gamma},\quad M=-n\wedge E_{/\Gamma}. \tag{54}$$

The general solution of system (53) can be written in the form

$$\tilde{E}(x)=\mathrm{i}Z_0\tilde{T}_rJ(x)+\tilde{K}_rM(x)+E^{\mathrm{Her}}(x),$$
$$\tilde{H}(x)=-\tilde{K}J(x)+\mathrm{i}Z_0^{-1}\tilde{T}_rM(x)+H^{\mathrm{Her}}(x), \tag{55}$$

where $\tilde{T}_r$, $\tilde{K}_r$ are defined in the same manner as $\tilde{T}$, $\tilde{K}$ given in (6) and (7) except that $G(x,y)$ is replaced with $\Re e\, G(x,y)$, and where $(E^{\mathrm{Her}}(x), H^{\mathrm{Her}}(x))$ is an entire solution to Maxwell's equations

$$k^2 E^{\mathrm{Her}} + \vec{\Delta} E^{\mathrm{Her}} = 0, \quad k^2 H^{\mathrm{Her}} + \vec{\Delta} H^{\mathrm{Her}} = 0,$$
$$\vec{\nabla}\cdot E^{\mathrm{Her}} = 0, \quad \vec{\nabla}\cdot H^{\mathrm{Her}} = 0. \tag{56}$$

Note that the choice of the kernel $\Re e\, G(x,y)$ for the particular solution of system (53) is completely arbitrary at this stage, but it is convenient for our purpose. Since (E,H) is in $\mathscr{W}$, the pair $(\tilde{E},\tilde{H})$ satisfies the growth property (41) and it is easy to show that the potentials in the right hand side of (55) also satisfy the same growth property (see (65)). We therefore deduce that $(E^{\mathrm{Her}}, H^{\mathrm{Her}})$ satisfies (41). Now, we use Theorem 6.30 of [9] that asserts that every entire solution to the Maxwell system satisfying the growth property (41) is an Herglotz pair with kernel γ: it means that there exists some tangential field γ in $L^2(S^2)$ such that

$$E^{\mathrm{Her}}(x) = \frac{\mathrm{i}k\sqrt{\mathrm{i}Z_0}}{4\pi}\int_{S^2}\gamma(\hat{d})\mathrm{e}^{\mathrm{i}k\hat{d}x}\,\mathrm{d}\sigma(\hat{d}),$$
$$H^{\mathrm{Her}}(x) = \frac{\mathrm{i}k}{4\pi\sqrt{\mathrm{i}Z_0}}\int_{S^2}(\mathrm{i}\hat{d}\wedge\gamma(\hat{d}))\mathrm{e}^{\mathrm{i}k\hat{d}x}\,\mathrm{d}\sigma(\hat{d}). \tag{57}$$

Once again, the normalization constant $\mathrm{i}k\sqrt{\mathrm{i}Z_0}/4\pi$ is here just for convenience.

Thus, a possible parameterization of incoming and outgoing fields might be $(\mathbf{u},\gamma)$ with

$$\mathbf{u} = \begin{bmatrix} J_1 \\ M_1 \end{bmatrix} = \begin{bmatrix} \sqrt{\mathrm{i}Z_0} & J(x) \\ \sqrt{\mathrm{i}Z_0}^{-1} & M(x) \end{bmatrix} \tag{58}$$

and with γ, the kernel of the Herglotz pair $(E^{\mathrm{Her}}, H^{\mathrm{her}})$ in (57). Reciprocally, to every $(\mathbf{u},\gamma)$ in $\mathbf{U}\times Z$ we can associate an electromagnetic field (E,H) in $\mathscr{W}$ through (55), (57), (58). But it remains to verify that the associated fields $\tilde{E}$ and $\tilde{H}$ vanish in D^-, the open complement to D^+. To prove that $(\tilde{E},\tilde{H})$ vanishes in D^-, it is enough to ensure that both tangential interior traces on Γ are zero. Using once again the jump conditions for the potentials, we get

$$0 = n(x)\wedge(\tilde{E}^-_{/\Gamma}(x)\wedge n(x))$$
$$= \mathrm{i}Z_0(T_rJ)(x) + (K_rM)(x) - \tfrac{1}{2}n(x)\wedge M(x) + \mathrm{i}\mathrm{e}^{\mathrm{Her}}(x), \tag{59}$$
$$0 = n(x)\wedge(H^+_{/\Gamma}(x)\wedge n(x))$$
$$= -(K_rJ)(x) + \mathrm{i}Z_0^{-1}(T_rM)(x) + \tfrac{1}{2}n(x)\wedge J(x) + \mathrm{i}h^{\mathrm{Her}}(x), \tag{60}$$

with

$$\mathrm{e}^{\mathrm{Her}}(x) = n(x)\wedge\left(\frac{k\sqrt{\mathrm{i}Z_0}}{4\pi}\int_{S^2}\gamma(\hat{d})\mathrm{e}^{\mathrm{i}k\hat{d}x}\,\mathrm{d}\sigma(\hat{d})\wedge n(x)\right),$$
$$h^{\mathrm{Her}}(x) = n(x)\wedge\left(\frac{k}{4\pi\sqrt{\mathrm{i}Z_0}}\int_{S^2}(\mathrm{i}\hat{d}\wedge\gamma(\hat{d}))\mathrm{e}^{\mathrm{i}k\hat{d}x}\,\mathrm{d}\sigma(\hat{d})\wedge n(x)\right). \tag{61}$$

Let us normalize (59) by $\sqrt{\mathrm{i}Z_0}^{-1}$ and (60) by $-\sqrt{\mathrm{i}Z_0}$. A look at Definitions (17) and (24), after transposition, gives us immediately

$$\mathbf{Tu} + \mathrm{i}(\mathbf{A}^{\infty})^{*}\gamma = 0. \tag{62}$$

Eq. (62) defines a closed linear sub-manifold $\mathscr{M}$ of $\mathbf{U} \times Z$ in which lies the pair $(\mathbf{u}, \gamma)$.

Lemma 3. *Let $\mathscr{M}$ be the closed linear sub-manifold of $\mathbf{U} \times Z$ defined by*

$$\mathscr{M} = \{(\mathbf{U}, \gamma) \in \mathbf{U} \times Z,\ \ \mathbf{Tu} + \mathrm{i}(\mathbf{A}^{\infty})^{*}\gamma = 0\}.$$

If $m = (\mathbf{u}, \gamma)$ is in $\mathscr{M}$, (55)–(57) *define a unique electromagnetic field (E, H)* in *$\mathscr{W}$ with*

$$\mathbf{u} = \begin{bmatrix} \sqrt{\mathrm{i}Z_0}(n(x) \wedge H_{/\Gamma}) \\ \sqrt{\mathrm{i}Z_0}^{-1}(-n(x) \wedge E_{/\Gamma}) \end{bmatrix}. \tag{63}$$

Reciprocally, every electromagnetic field (E, H) in $\mathscr{W}$ can be written in the form (55)–(57) *via* (58) *and therefore can be associated to an element of $\mathscr{M}$.*

The following paragraph is devoted to the interpretation of the Herglotz kernel γ in terms of the asymptotic behavior of the associated electromagnetic field.

3.4. Asymptotic behavior of electromagnetic pairs in $\mathscr{W}$

We study the behavior of the fields $(\tilde{E}, \tilde{H})$, as given in (55) when x goes to infinity. For the potential, the calculations are classical (see [9, p. 157]). From,

$$\begin{aligned} \nabla_x \wedge G(x,y)a(y) &= \frac{\mathrm{e}^{\mathrm{i}k|x|}}{4\pi|x|}\left(+\mathrm{i}k\hat{x} \wedge a(y)\mathrm{e}^{-\mathrm{i}k\hat{x}y} + \mathrm{O}\left(\frac{\|a\|_Z}{|x|}\right)\right), \\ \nabla_x \wedge (\nabla_x \wedge G(x,y)a(y)) &= \frac{k^2\mathrm{e}^{\mathrm{i}k|x|}}{4\pi|x|}\left(\hat{x} \wedge (a(y) \wedge \hat{x})\mathrm{e}^{-\mathrm{i}k\hat{x}y} + \mathrm{O}\left(\frac{\|a\|_Z}{|x|}\right)\right), \end{aligned} \tag{64}$$

as $|x|$ goes to infinity, uniformly for all y in Γ, we get

$$\begin{aligned} \frac{E(x) - E^{\mathrm{Her}}(x)}{\sqrt{\mathrm{i}Z_0}} &= \frac{1}{2}\mathbf{A}^{\infty}(J_1, M_1; \hat{x})\frac{\mathrm{e}^{\mathrm{i}k|x|}}{|x|} \\ &\quad + \frac{1}{2}\mathbf{A}^{\infty}(J_1, M_1; -\hat{x})\frac{\mathrm{e}^{-\mathrm{i}k|x|}}{|x|} + \mathrm{O}\left(\frac{1}{|x|^2}\right), \\ (H(x) - H^{\mathrm{Her}}(x))\sqrt{\mathrm{i}Z_0} &= \frac{1}{2}(-\mathrm{i}\hat{x} \wedge \mathbf{A}^{\infty}(J_1, M_1; \hat{x}))\frac{\mathrm{e}^{\mathrm{i}k|x|}}{|x|} \\ &\quad + \frac{1}{2}(+\mathrm{i}\hat{x} \wedge \mathbf{A}^{\infty}(J_1, M_1; -\hat{x}))\frac{\mathrm{e}^{-\mathrm{i}k|x|}}{|x|} + \mathrm{O}\left(\frac{1}{|x|^2}\right). \end{aligned} \tag{65}$$

For regular Herglotz kernel, the asymptotic behavior of the Herglotz wave can be obtained thanks to the stationary phase Theorem; if γ_0 is C^1, we have

$$F(\gamma_0;x) = \int_{S^2} \gamma_0(\hat{d})\mathrm{e}^{\mathrm{i}k\hat{d}x}\,\mathrm{d}\sigma(\hat{d}) = F^{\mathrm{asym}}(\gamma_0;x) + \mathrm{O}\left(\frac{1}{|x|^2}\right), \tag{66}$$

$$F^{\mathrm{asym}}(\gamma_0;x) = \frac{2\pi}{\mathrm{i}}\left(\gamma_0(\hat{x})\frac{\mathrm{e}^{\mathrm{i}k|x|}}{k|x|} - \gamma_0(-\hat{x})\frac{\mathrm{e}^{-\mathrm{i}k|x|}}{k|x|}\right).$$

We use a well known result about the asymptotic behavior of Herglotz waves [17] to prove that this result can be extended in a weaker sense when γ is only L^2; more precisely, we have

$$\forall\gamma_0\in Z,\quad \lim_{R\to\infty}\frac{1}{R}\int_{R\leqslant|x|\leqslant 2R}|F(\gamma_0;x) - F^{\mathrm{asym}}(\gamma_0;x)|^2\,\mathrm{d}x = 0. \tag{67}$$

We do not know if some stronger convergence occurs, nevertheless, (67) is enough to our purpose (indeed, it was enough to get the isometry lemma).

From (67), we deduce

$$\lim_{R\to\infty}\frac{1}{R}\int_{R\leqslant|x|\leqslant 2R}\frac{E^{\mathrm{Her}}(x)}{\sqrt{\mathrm{i}Z_0}} - \left(\frac{1}{2}\gamma(\hat{x})\frac{\mathrm{e}^{\mathrm{i}k|x|}}{k|x|} - \frac{1}{2}\gamma(-\hat{x})\frac{\mathrm{e}^{-\mathrm{i}k|x|}}{k|x|}\right) = 0$$

$$\lim_{R\to\infty}\frac{1}{R}\int_{R\leqslant|x|\leqslant 2R}H^{\mathrm{Her}}(x)\sqrt{\mathrm{i}Z_0} + \mathrm{i}\hat{x}\wedge\left(\frac{1}{2}\gamma(\hat{x})\frac{\mathrm{e}^{\mathrm{i}k|x|}}{k|x|} + \frac{1}{2}\gamma(-\hat{x})\frac{\mathrm{e}^{-\mathrm{i}k|x|}}{k|x|}\right) = 0.$$

Finally, gathering this result with asymptotics (65) provides

Lemma 4. *Let* $(\mathbf{u},\gamma)$ *in* $\mathscr{M}$ *be the associated element to* (E,H) *in* $\mathscr{W}$ *introduced in Lemma* 3 *and let* $a^{\mathrm{in}}(E,H)$ *and* $a^{\mathrm{out}}(E,H)$ *be defined in* $L^2(S^2)$ *by*

$$a^{\mathrm{out}}_{\infty}(E,H;\hat{x}) = \frac{\sqrt{\mathrm{i}Z_0}}{2}(\mathbf{A}^{\infty}(\mathbf{u};\hat{x}) + \gamma(\hat{x})),$$

$$a^{\mathrm{in}}_{\infty}(E,H;\hat{x}) = \frac{\sqrt{\mathrm{i}Z_0}}{2}(\mathbf{A}^{\infty}(\mathbf{u};-\hat{x}) - \gamma(-\hat{x})), \tag{68}$$

then, the asymptotics (43), (44) *hold.*

3.5. Reformulation of the minimization problem

Once the parameterization has been constructed, it only remains to rewrite the functional $I(E,H)$. At first, we have

$$\frac{1}{4}\|G^{\mathrm{in}}_{\Gamma}(E,H)\|^2 + \frac{1}{4}\|G^{\mathrm{out}}_{\Gamma}(E,H)\|^2 = \frac{1}{2}\|E\wedge n\|^2 + \frac{Z_0^2}{2}\|H\wedge n\|^2$$

$$= \frac{Z_0}{2}(\|J_1\|^2 + \|M_1\|^2) = \frac{Z_0}{2}\|\mathbf{u}\|^2_{\mathbf{u}}, \tag{69}$$

$$(G^{\mathrm{in}}_{\Gamma}(E,H),G^{\mathrm{in}}) = Z_0\int_{\Gamma}(J_1 - \mathrm{i}n\wedge M_1)\cdot\overline{\frac{G^{\mathrm{in}}}{\mathrm{i}\sqrt{\mathrm{i}Z_0}}}\,\mathrm{d}\Gamma = Z_0(\mathbf{u},\mathbf{g})_{\mathbf{U}}, \tag{70}$$

where $\mathbf{g}$ is defined in (30). Second, Lemma 4 gives

$$\|a_\infty^{\mathrm{out}}(E,H)\|_Z^2 + \|a_\infty^{\mathrm{in}}(E,H)\|_Z^2 = \frac{Z_0}{2}\|\gamma\|_Z^2 + \frac{Z_0}{2}\|\mathbf{A}^\infty(J_1,M_1)\|_Z^2, \tag{71}$$

and the functional is, finally,

$$I(E,H) = Z_0 J(\mathbf{u},\gamma),$$

$$J(\mathbf{u},\gamma) = \tfrac{1}{2}\|\mathbf{u}\|_{\mathbf{u}}^2 + \tfrac{1}{2}\|\gamma\|_Z^2 + \tfrac{1}{2}\|\mathbf{A}^\infty\mathbf{u}\|_Z^2 - \Re e(\mathbf{u},\mathbf{g})_{\mathbf{U}}. \tag{72}$$

We can now reformulate Theorem 1 as

Theorem 2. *Let G^{in} be given in $L^2(\Gamma)$. Define $\mathbf{g}$ as in* (30). *The minimum of the functional $J(\mathbf{u},\gamma)$ given in* (72), *over all the pairs $(\mathbf{u},\gamma)\in\mathscr{M}$, i.e. satisfying*

$$\mathbf{T}\mathbf{u} + \mathrm{i}(\mathbf{A}^\infty)^*\gamma = 0, \tag{73}$$

is reached at

$$\mathbf{u} = \begin{bmatrix} \sqrt{\mathrm{i}Z_0}\, n\wedge H^+_{/\Gamma} \\ -\sqrt{\mathrm{i}Z_0}^{-1} n\wedge E^+_{/\Gamma} \end{bmatrix}, \quad \gamma = \frac{a_\infty^{\mathrm{out}}(E^+,H^+)}{\sqrt{\mathrm{i}Z_0}}, \tag{74}$$

where (E^+,H^+) is the radiating (outgoing) solution of the Maxwell system (47).

3.6. Optimality conditions

We define the Hilbert space

$$\mathbf{V} = \{\mathbf{v}\in\mathbf{u},\ \text{such that}\ \mathbf{T}^*\mathbf{v}\in\mathbf{u}\}, \tag{75}$$

equipped with the following norm:

$$\|\mathbf{v}\|_{\mathbf{V}} = \|\mathbf{v}\|_{\mathbf{U}} + \|\mathbf{T}^*\mathbf{v}\|_{\mathbf{U}}. \tag{76}$$

If $\mathbf{V}'$ is the dual space of $\mathbf{V}$, constraint (73) can be viewed as an equality in $\mathbf{V}'$

$$\forall \mathbf{v}\in\mathbf{V},\quad (\mathbf{u},\mathbf{T}^*\mathbf{v})_{\mathbf{U}} + \mathrm{i}(\gamma,(\mathbf{A}^\infty)\mathbf{v})_Z = 0. \tag{77}$$

At this point it is classical to dualize the constraint, introducing the Lagrangian

$$\mathscr{L}(\mathbf{u},\gamma,\mathbf{v}) = J(\mathbf{u},\gamma) - \Re e((\mathbf{u},\mathbf{T}^*\mathbf{v})_{\mathbf{U}} - \mathrm{i}(\gamma,(\mathbf{A}^\infty)\mathbf{v})_Z), \tag{78}$$

where $\mathbf{W}$ is the space for the multiplier, i.e. the quotient Hilbert space

$$\mathbf{W} = \frac{\mathbf{V}}{\mathrm{Ker}\,\mathbf{T}^*} \tag{79}$$

equipped with the norm

$$\|\mathbf{v}\|_{\mathbf{W}} = \inf_{\mathbf{v}_0\in\mathrm{Ker}\,\mathbf{T}^*}\|\mathbf{v}-\mathbf{v}_0\|_{\mathbf{V}} = \|\mathbf{T}^*\mathbf{v}\|_{\mathbf{U}} + \inf_{\mathbf{v}_0\in\mathrm{Ker}\,\mathbf{T}^*}\|\mathbf{v}-\mathbf{v}_0\|_{\mathbf{U}}.$$

It is well known (cf. [5]) that if the Lagrangian admits a saddle point, then its first argument is the minimum argument of J in $\mathscr{M}$

$$\mathscr{L}(\mathbf{u},\gamma,\mathbf{v}) = \inf_{(\mathbf{u}^*,\gamma^*)\in \mathbf{U}\times Z} \sup_{\mathbf{v}^*\in \mathbf{W}} \mathscr{L}(\mathbf{u}^*,\gamma^*,\mathbf{v}^*) = \min_{(\mathbf{u}^*,\gamma^*)\in\mathscr{M}} J(\mathbf{u}^*,\gamma^*), \tag{80}$$

Furthermore, since J is quadratic, such a saddle point exists if and only if $D_{\mathbf{u}}\mathscr{L}(\mathbf{u},\gamma,\mathbf{v}) = 0$, $D_\gamma\mathscr{L}(\mathbf{u},\gamma,\mathbf{v}) = 0$, $D_{\mathbf{v}}\mathscr{L}(\mathbf{u},\gamma,\mathbf{v}) = 0$, i.e.

$$\begin{aligned} &\mathbf{u} + (\mathbf{A}^\infty)^*\mathbf{A}^\infty\mathbf{u} - \mathbf{T}^*\mathbf{v} = \mathbf{g},\\ &\gamma + \mathrm{i}\mathbf{A}^\infty\mathbf{v} = 0,\\ &-\mathbf{T}\mathbf{u} - \mathrm{i}(\mathbf{A}^\infty)^*\gamma = 0. \end{aligned} \tag{81}$$

Discarding γ, system (34) is then recovered. We have already proved in the first derivation that $\mathbf{v}=\mathrm{i}\mathbf{u}$ is a solution. It is possible to recover this property directly. Since $a_\infty^{\mathrm{in}}(E,H) = 0$ at the optimum, Eq. (68) implies that $\mathbf{A}^\infty\mathbf{u} = \gamma = -\mathrm{i}\mathbf{A}^\infty\mathbf{v}$ and therefore $\mathbf{v} = \mathrm{i}\mathbf{u}$ up to an element in $\operatorname{Ker}\mathbf{A}^\infty$. It is shown in Appendix B that $\operatorname{Ker}\mathbf{A}^\infty = \operatorname{Ker}\mathbf{T}^*$ and consequently we obtain $\mathbf{v} = \mathrm{i}\mathbf{u}$ in $\mathbf{W}$.

Thus, we have given another derivation of the mixed integral system. The interesting feature of this second derivation is that the saddle point problem is explained and is related to the isometry lemma. The additional unknown, whose introduction might seem strange in the first Section 1 can now be interpreted as the Lagrange multiplier of our constrained minimization problem.

4. Variational formulation and well-posedness. The penalized systems

Variational formulation is useful for minimization problems. It provides a good framework for the study of uniqueness and existence, and also for discretization and convergence of the discrete solution. A possible variational formulation of our problem is the following.

$$\begin{aligned} &\mathbf{u}\in V_{\mathbf{u}}, \quad \mathbf{v}\in V_{\mathbf{v}},\\ &(\mathbf{u},\tilde{\mathbf{u}}) + (\mathbf{A}_\infty\mathbf{u},\mathbf{A}_\infty\tilde{\mathbf{u}}) - (\mathbf{T}^*\mathbf{v},\tilde{\mathbf{u}}) = (\mathbf{g},\tilde{\mathbf{u}}), \quad \forall\tilde{\mathbf{u}}\in V_{\mathbf{u}},\\ &(\mathbf{u},\mathbf{T}^*\tilde{\mathbf{v}}) + (\mathbf{A}_\infty\mathbf{v},\mathbf{A}_\infty\tilde{\mathbf{v}}) = 0, \quad \forall\tilde{\mathbf{v}}\in V_{\mathbf{v}}, \end{aligned} \tag{82}$$

where it remains to define the functional spaces $V_{\mathbf{u}}$ and $V_{\mathbf{v}}$. Due to the L^2 coerciveness of the formulation it is clear that an interesting choice for $V_{\mathbf{u}}$ is

$$V_{\mathbf{u}} = \mathbf{U} = TL^2(\Gamma)\times TL^2(\Gamma).$$

But, it may seem at first sight impossible to define an associated space $V_{\mathbf{v}}$ such that the inf–sup condition of Babuska–Brezzi holds, [5],

$$\max\frac{(\mathbf{u},\mathbf{T}^*\mathbf{v})}{\|\mathbf{u}\|_{L^2}} \geqslant k\|\mathbf{v}\|_{W_{\mathbf{v}}},\ k>0 \quad \text{with } W_{\mathbf{v}} = \frac{V_{\mathbf{v}}}{\operatorname{Ker}\mathbf{T}^*}, \tag{83}$$

just because standard functional spaces in which $\mathbf{T}$ is continuous are known to be based on

$$H^{-1/2}(\mathrm{div},\Gamma) \quad \text{and} \quad H^{-1/2}(\mathrm{curl},\Gamma).$$

Nevertheless it is at least possible to provide an abstract framework in which the inf–sup condition holds. Let us take $V_{\mathbf{u}} = \mathbf{U}$, $V_{\mathbf{v}} = \mathbf{V}$, defined in (75) and $\mathbf{W}$ defined in (79).

System (34) is well-posed as soon as the inf–sup condition

$$\sup_{\mathbf{u}\in\mathbf{U}} \frac{(\mathbf{u}, \mathbf{T}^*\mathbf{v})_{\mathbf{U}}}{\|\mathbf{u}\|_{\mathbf{U}}} \geqslant C\|\mathbf{v}\|_{\mathbf{W}} \tag{84}$$

holds for some positive constant $C > 0$. This inequality can be derived as follows. Picking $\mathbf{u} = \mathbf{T}^*\mathbf{v}$ in (84), we find

$$\sup_{\mathbf{u}\in\mathbf{U}} \frac{(\mathbf{u}, \mathbf{T}^*\mathbf{v})_{\mathbf{U}}}{\|\mathbf{u}\|_{\mathbf{U}}} \geqslant \|\mathbf{T}^*\mathbf{v}\|_{\mathbf{U}}. \tag{85}$$

Let $\mathbf{v} \in \mathbf{V}$. In Appendix B it is proved that $-\Pi\mathbf{T}$ with Π defined in (21), is a projector (it is a Calderon Projector):

$$\mathbf{T}\Pi\mathbf{T} = -\mathbf{T}. \tag{86}$$

Let $\mathbf{v}_1 = -\Pi^*\mathbf{T}^*\mathbf{v}$. We have that $\mathbf{T}^*(\mathbf{v} - \mathbf{v}_1) = \mathbf{T}^*\mathbf{v} + \mathbf{T}^*\Pi^*\mathbf{T}^*\mathbf{v} = 0$. So $\mathbf{v}_0 = \mathbf{v} - \mathbf{v}_1 \in Ker\,\mathbf{T}^*$ hence

$$\inf\nolimits_{\mathbf{v}_0\in\mathrm{Ker}\,\mathbf{T}^*} \|\mathbf{v} - \mathbf{v}_0\|_{\mathbf{U}} \leqslant \|\mathbf{v}_1\|_{\mathbf{U}} = \|\Pi^*\mathbf{T}^*\mathbf{v}\|_{\mathbf{U}} = \|\mathbf{T}^*\mathbf{v}\|_{\mathbf{U}}, \tag{87}$$

where we have used the isometric property $\Pi^*\Pi = I$. So we have

$$\sup_{\mathbf{u}\in\mathbf{U}} \frac{(\mathbf{u}, \mathbf{T}^*\mathbf{v})_{\mathbf{U}}}{\|\mathbf{u}\|_{\mathbf{U}}} \geqslant \frac{1}{2}\|\mathbf{T}^*\mathbf{v}\|_{\mathbf{U}} + \frac{1}{2}\inf_{\mathbf{v}_0\in\mathrm{Ker}\,\mathbf{T}^*} \|\mathbf{v} - \mathbf{v}_0\|_{\mathbf{U}} = \frac{1}{2}\|\mathbf{v}\|_{\mathbf{W}}. \tag{88}$$

Thus it gives

Lemma 5. *The* inf–sup *condition* (84) *in space* $\mathbf{U} \times \mathbf{V}$ *is true with* $C = \frac{1}{2}$.

Since both the continuity $\mathbf{T}^* : \mathbf{V} \to \mathbf{U}$ and the bound $\|\mathbf{R}\mathbf{v}\|_{\mathbf{U}} \leqslant C'\|\mathbf{v}\|_{\mathbf{W}}$ for some $C' > 0$ hold (see inequality (B.13) in Appendix B), we have, following [5]

Theorem 3. *The variational system* (82) *is well posed, that is, for every* $\mathbf{g} \in \mathbf{u}' = \mathbf{u}$ *there exists a unique* $(\mathbf{u}, \mathbf{v}) \in \mathbf{u} \times \mathbf{W}$ *such that*

$$\begin{aligned} &(\mathbf{u}, \tilde{\mathbf{u}}) + (\mathbf{A}_\infty\mathbf{u}, \mathbf{A}_\infty\tilde{\mathbf{u}}) - (\mathbf{T}^*\mathbf{v}, \tilde{\mathbf{u}}) = (\mathbf{g}, \tilde{\mathbf{u}}), \quad \forall\tilde{\mathbf{u}} \in \mathbf{U},\\ &(\mathbf{u}, \mathbf{T}^*\tilde{\mathbf{v}}) + (\mathbf{A}_\infty\mathbf{v}, \mathbf{A}_\infty\tilde{\mathbf{v}}) = 0, \quad \forall\tilde{\mathbf{v}} \in \mathbf{V}. \end{aligned} \tag{89}$$

However when discretization is considered a difficulty arises with the use of spaces $(\mathbf{U}, \mathbf{V}) = (V_{\mathbf{u}}, V_{\mathbf{v}})$. The reason is that we want to avoid the construction of some discrete space compatible with the L^2 based space $(\mathbf{U}, \mathbf{V})$. We would like to take a classical integral code based on the duality $H^{-1/2}(\mathrm{div}, \Gamma)$ and $H^{-1/2}(\mathrm{curl}, \Gamma)$ and use the iterative algorithms described later in order to solve our new discrete integral system. Then the question of the convergence of the discrete solution to the exact one arises. All our efforts to prove the convergence using this strategy failed. The reason seems to be that the classical discretization of integral operators is based on $H^{-1/2}(\mathrm{div},\Gamma)$ and $H^{-1/2}(\mathrm{curl},\Gamma)$, and not on $(\mathbf{U},\mathbf{V})$, [6]. Moreover, numerical results in 2-D for the Helmholtz equation, [3], show that this problem may be a real one; there are cases where the discrete solution obtained through the strategy described above does not converge to the exact solution, even in some very simple and regular cases. Of course this conclusion has to be re-evaluated if the discretization of the integral operators are compatible with $(\mathbf{U},\mathbf{V})$. It is our purpose now to modify the system and to present what we will call the penalized problem, with a much stronger coercivity.

Let β be some positive penalization parameter (for instance $\beta=1$). Recalling that $\mathbf{v}=\mathrm{i}\mathbf{u}$, we modify system (26) to obtain the penalized system

$$\begin{aligned}&(1+\beta)\mathbf{u}+(\mathbf{A}^\infty)^*\mathbf{A}^\infty\mathbf{u}-\mathbf{T}^*\mathbf{v}+\mathrm{i}\beta\mathbf{v}=\mathbf{g},\\&+\mathbf{T}\mathbf{u}-\mathrm{i}\beta\mathbf{u}+\beta\mathbf{v}+(\mathbf{A}^\infty)^*\mathbf{A}^\infty\mathbf{v}=0,\end{aligned}\tag{90}$$

which is a system of the form

$$\mathscr{A}_\beta\begin{bmatrix}\mathbf{u}\\ \mathbf{v}\end{bmatrix}=\begin{bmatrix}\mathbf{g}\\ 0\end{bmatrix}.\tag{91}$$

The associated variational system is for a given pair $(\mathbf{g},\mathbf{g}_\mathbf{v})\in(\mathbf{U}',\mathbf{V}')=(\mathbf{U},\mathbf{V}')$, find $(\mathbf{u},\mathbf{v})\in\mathbf{U}\times\mathbf{V}$ such that $\forall(\tilde{\mathbf{u}},\tilde{\mathbf{v}})\in\mathbf{U}\times\mathbf{V}$

$$\begin{aligned}&(1+\beta)(\mathbf{u},\tilde{\mathbf{u}})+(\mathbf{A}_\infty\mathbf{u},\mathbf{A}_\infty\tilde{\mathbf{u}})-(\mathbf{T}^*\mathbf{v},\tilde{\mathbf{u}})+\mathrm{i}\beta(\mathbf{v},\tilde{\mathbf{u}})=(\mathbf{g},\tilde{\mathbf{u}}),\\&(\mathbf{u},\mathbf{T}^*\tilde{\mathbf{v}})-\mathrm{i}\beta(\mathbf{u},\tilde{\mathbf{v}})+(\mathbf{A}_\infty\mathbf{v},\mathbf{A}_\infty\tilde{\mathbf{v}})+\beta(\mathbf{v},\tilde{\mathbf{v}})=\langle g_\mathbf{v},\tilde{\mathbf{v}}\rangle_{\mathbf{V},\mathbf{V}'}.\end{aligned}\tag{92}$$

Simple calculations show that

$$\begin{aligned}\mathfrak{Re}\left(\mathscr{A}_\beta\begin{bmatrix}\mathbf{u}\\ \mathbf{v}\end{bmatrix},\begin{bmatrix}\mathbf{u}\\ \mathbf{v}\end{bmatrix}\right)_{\mathbf{U}\times\mathbf{U}}&=\|\mathbf{u}\|_\mathbf{U}^2+\beta\|\mathbf{u}+\mathrm{i}\mathbf{v}\|_\mathbf{U}^2+\|\mathbf{A}^\infty\mathbf{v}\|_Z^2+\|\mathbf{A}^\infty\mathbf{u}\|_Z^2\\&\geqslant C_\beta(\|\mathbf{u}\|_\mathbf{U}^2+\|\mathbf{v}\|_\mathbf{U}^2),\end{aligned}\tag{93}$$

with

$$C_\beta=\beta+\frac{1}{2}-\sqrt{\beta^2+\frac{1}{4}}\geqslant\min\left(\frac{\beta}{2},\frac{1}{3}\right)$$

and the system is now coercive in the $\mathbf{u}$ variable and in the $\mathbf{v}$ variable, even if the norm $\|\mathbf{v}\|_\mathbf{U}^2$ in (93) is not the norm in $\mathbf{V}$ required to have a true coercivity property.

Theorem 4. *The variational system* (92) *has a unique solution* $(\mathbf{u},\mathbf{v})$ *in* $\mathbf{U}\times\mathbf{V}$. *For* $\mathbf{g}$ *in* $\mathbf{U}$, *Eq*. (90) *has a unique solution in* $\mathbf{U}\times\mathbf{V}$.

Proof. Uniqueness of the solution is obvious. Existence is obtained by using the arguments used to prove existence for saddle points problems; we transform the operator $\mathscr{A}_\beta$ into $\mathscr{A}_{\beta,\varepsilon}$ by adding the term $\varepsilon\mathbf{T}\mathbf{T}^*\mathbf{v}$ to the second equation in (90). The modified system is

$$(1+\beta)\mathbf{u}^\varepsilon + (\mathbf{A}^\infty)^*\mathbf{A}^\infty\mathbf{u}^\varepsilon - \mathbf{T}^*\mathbf{v}^\varepsilon + \mathrm{i}\beta\mathbf{v}^\varepsilon = \mathbf{g}$$

$$+\mathbf{T}\mathbf{u}^\varepsilon - \mathrm{i}\beta\mathbf{u}^\varepsilon + \beta\mathbf{v}^\varepsilon + (\mathbf{A}^\infty)^*\mathbf{A}^\infty\mathbf{v}^\varepsilon + \varepsilon\mathbf{T}\mathbf{T}^*\mathbf{v}^\varepsilon = \mathbf{g}_v,$$

$$\Leftrightarrow \mathscr{A}_{\beta,\varepsilon}\begin{bmatrix}\mathbf{u}^\varepsilon\\ \mathbf{v}^\varepsilon\end{bmatrix} = \begin{bmatrix}\mathbf{g}\\ \mathbf{g}_v\end{bmatrix}. \tag{94}$$

Note that the operator $\mathscr{A}_{\beta,\varepsilon}$ is coercive due to

$$\begin{aligned}\Re e(\mathscr{A}_{\beta,\varepsilon}[\mathbf{u},\mathbf{v}]^{\mathrm{t}},[\mathbf{u},\mathbf{v}]^{\mathrm{t}}) &\geqslant C_\beta(\|\mathbf{u}\|_{\mathbf{U}}^2 + \|\mathbf{v}\|_{\mathbf{U}}^2) + \varepsilon\|\mathbf{T}^*\mathbf{v}\|_{\mathbf{U}}^2)\\ &\geqslant C_{\beta,\varepsilon}(\|\mathbf{u}\|_{\mathbf{U}}^2 + \|\mathbf{v}\|_{\mathbf{V}}^2).\end{aligned} \tag{95}$$

Here $(.,.)^{\mathrm{t}}$ denotes vector transpose. Since continuity is obvious, the Lax-Milgram Theorem gives existence and uniqueness of the solution $(\mathbf{u}^\varepsilon,\mathbf{v}^\varepsilon)$. Now, the inequality of coercivity and the continuity of $\mathscr{A}_{\beta,\varepsilon}$ provide the estimate

$$\|\mathbf{u}^\varepsilon\|_{\mathbf{U}}^2 + \|\mathbf{v}^\varepsilon\|_{\mathbf{U}}^2 \leqslant \frac{1}{C_\beta}\,\|\mathscr{A}_{\beta,\varepsilon}\|\;\|(\mathbf{g},\mathbf{g}_v)\|_{\mathbf{U}\times\mathbf{V}'}\|(\mathbf{u}^\varepsilon,\mathbf{v}^\varepsilon)\|_{\mathbf{U}\times\mathbf{V}}. \tag{96}$$

Using a triangular inequality in first equation of system (94), gives

$$\|\mathbf{T}^*\mathbf{v}^\varepsilon\|_{\mathbf{U}} \leqslant \|\mathbf{g}\|_{\mathbf{U}} + (1+\beta+\|(\mathbf{A}^\infty)^*\mathbf{A}^\infty\|)\|\mathbf{u}^\varepsilon\|_{\mathbf{U}} + \beta\|\mathbf{v}^\varepsilon\|_{\mathbf{U}}. \tag{97}$$

It is not difficult to check that $\|\mathscr{A}_{\beta,\varepsilon}\|$ is bounded by $\|\mathscr{A}_\beta\| + \varepsilon$. We deduce from (96)–(97) that there is a constant C (which is function of β and is independent of ε) such that

$$\|\mathbf{u}^\varepsilon\|_{\mathbf{U}} \leqslant C\|(\mathbf{g},\mathbf{g}_v)\|_{\mathbf{U}\times\mathbf{V}'} \quad \|\mathbf{v}^\varepsilon\|_{\mathbf{V}} \leqslant C\|(\mathbf{g},\mathbf{g}_v)\|_{\mathbf{U}\times\mathbf{V}'}.$$

Consequently, it is possible to extract a sub-sequence converging weakly in $\mathbf{U}\times\mathbf{V}$. Writing down the variational formulation of the ε-problem, and passing to the limit, we obtain a solution of problem (92).

Remark. Another type of penalization consists in modifying system (26) according to

$$\begin{aligned}&(1-\beta)\mathbf{u} + (\mathbf{A}^\infty)^*\mathbf{A}^\infty\mathbf{u} - (\mathbf{T}^*\mathbf{v} + \mathrm{i}\beta\mathbf{v}) = \mathbf{g},\\ &(\mathbf{T}\mathbf{u} - \mathrm{i}\beta\mathbf{u}) + \beta\mathbf{v} + (\mathbf{A}^\infty)^*\mathbf{A}^\infty\mathbf{v} = 0,\end{aligned} \tag{98}$$

where β is now some positive number less than 1 (let $\beta = \frac{1}{2}$). The interest of (98) is that it corresponds to a saddle point for the Lagrangian

$$\begin{aligned}\mathscr{L}_\beta(\mathbf{u},\gamma,\mathbf{v}) = {}&(1-\beta)\tfrac{1}{2}\|\mathbf{u}\|_{\mathbf{U}}^2 - \beta\,\tfrac{1}{2}\,\|\mathbf{v}\|_{\mathbf{U}}^2 - \Re e(\mathbf{g},\mathbf{u})_{\mathbf{U}}\\ &+\tfrac{1}{2}\,\|\gamma\|_Z^2 + \tfrac{1}{2}\,\|\mathbf{A}^\infty\mathbf{u}\|_Z^2 - \Re e((\mathbf{u},\mathbf{T}^*\mathbf{v} + \mathrm{i}\beta\mathbf{v})_{\mathbf{U}} - \mathrm{i}(\gamma,(\mathbf{A}^\infty)\mathbf{v})_Z).\end{aligned} \tag{99}$$

This problem is a penalized saddle point problem.

5. System for general impedance boundary conditions

We turn now our attention to the case of a general boundary condition. We assume that the electro-magnetic field satisfies some impedance boundary condition of the type

$$n(x) \wedge (E^{+}_{/\Gamma}(x) \wedge n(x)) + Z_0 \mathscr{Z}_r (H^{+}_{/\Gamma}(x) \wedge n(x)) = F, \tag{100}$$

where $\mathscr{Z}_r$ is some impedance operator that we assume symmetric with a positive real part, i.e.

$$(\Re e\, \mathscr{Z}_r J, J) \geqslant 0, \quad \forall J \in D(Z_r). \tag{101}$$

We associate to $\mathscr{Z}_r$ its reflection coefficient operator

$$\mathscr{R} = (Id - \mathscr{Z}_r)(Id + \mathscr{Z}_r)^{-1}, \tag{102}$$

which, thanks to (101), satisfies

$$\|\mathscr{R}\|_{\mathscr{L}(TL^2(\Gamma))} \leqslant 1. \tag{103}$$

We first rewrite the boundary condition in terms of $\mathscr{R}$. We have

$$G^{\text{in}} = -\mathscr{R} G^{\text{out}} + (Id + \mathscr{R})F, \tag{104}$$

where G^{in} is defined in (28) while

$$G^{\text{out}} = n(x) \wedge (E^{+}_{/\Gamma}(x) \wedge n(x)) - Z_0 (H^{+}_{/\Gamma}(x) \wedge n(x)). \tag{105}$$

Setting

$$F_0 = \frac{1}{\sqrt{\mathrm{i}Z_0}}(Id + \mathscr{R})F, \quad \mathbf{f} = \begin{bmatrix} -\mathrm{i}F_0 \\ -n(x) \wedge F_0 \end{bmatrix}, \tag{106}$$

and using definition (30) for $\mathbf{g}$ with (104) for G^{in} in system (34), we have

$$\begin{aligned} &\mathbf{u} + (\mathbf{A}^{\infty})^* \mathbf{A}^{\infty} \mathbf{u} - \mathbf{T}^* \mathbf{v} = \mathbf{f} - \mathbf{N}_{\mathscr{R}} \mathbf{u}, \\ &-\mathbf{T}\mathbf{u} - (\mathbf{A}^{\infty})^* \mathbf{A}^{\infty} \mathbf{v} = 0, \end{aligned} \tag{107}$$

where

$$\mathbf{N}_{\mathscr{R}} \mathbf{u} = \mathbf{N}_{\mathscr{R}} \begin{bmatrix} J_1(x) \\ M_1(x) \end{bmatrix} = \begin{bmatrix} -\mathrm{i}\mathscr{R} \frac{G^{\text{out}}(x)}{\sqrt{\mathrm{i}Z_0}} \\ -n(x) \wedge \mathscr{R} \frac{G^{\text{out}}(x)}{\sqrt{\mathrm{i}Z_0}} \end{bmatrix}, \tag{108}$$

or

$$\mathbf{N}_{\mathscr{R}} \mathbf{u} = \begin{bmatrix} -\mathrm{i}\mathscr{R}(n \wedge M_1 - \mathrm{i}J_1) \\ -n \wedge \mathscr{R}(n \wedge M_1 - \mathrm{i}J_1) \end{bmatrix}. \tag{109}$$

Let β some positive parameter (for instance $\beta = 1$), remembering that $\mathbf{v} = \mathrm{i}\mathbf{u}$, we finally modify the system to obtain the final system

$$(1+\beta)\mathbf{u} + (\mathbf{A}^{\infty})^*\mathbf{A}^{\infty}\mathbf{u} - \mathbf{T}^*\mathbf{v} + \mathbf{N}_{\mathscr{R}}\mathbf{u} + \mathrm{i}\beta\mathbf{v} = \mathbf{f}$$
$$\mathbf{T}\mathbf{u} - \mathrm{i}\beta\mathbf{u} + \beta\mathbf{v} + (\mathbf{A}^{\infty})^*\mathbf{A}^{\infty}\mathbf{v} = 0. \tag{110}$$

The associated weak formulation is

$$\mathbf{u} \in \mathbf{U},\ \ \mathbf{v} \in \mathbf{V},\ \text{ and } \forall \tilde{\mathbf{u}} \in \mathbf{U}, \forall\, \tilde{\mathbf{v}} \in \mathbf{V},$$
$$(1+\beta)(\mathbf{u},\tilde{\mathbf{u}}) + (\mathbf{A}_{\infty}\mathbf{u},\mathbf{A}_{\infty}\tilde{\mathbf{u}}) - (\mathbf{T}^*\mathbf{v},\tilde{\mathbf{u}}) + \mathrm{i}\beta(\mathbf{v},\tilde{\mathbf{u}}) - (\mathbf{N}_{\mathscr{R}}\mathbf{u},\tilde{\mathbf{u}}) = (\mathbf{f},\tilde{\mathbf{u}}),$$
$$(\mathbf{u},\mathbf{T}^*\tilde{\mathbf{v}}) - \mathrm{i}\beta(\mathbf{u},\tilde{\mathbf{v}}) + (\mathbf{A}_{\infty}\mathbf{v},\mathbf{A}_{\infty}\tilde{\mathbf{v}}) + \beta(\mathbf{v},\tilde{\mathbf{v}}) = 0. \tag{111}$$

The interest of this new formulation for general boundary conditions lies on the following lemma

Lemma 6. *Let $\mathscr{R} \in \mathscr{L}(TL^2(\Gamma))$ be a general reflexion operator bounded in the space of tangent square integrable functions. Let $\mathbf{N}_{\mathscr{R}} \in \mathscr{L}(\mathbf{U})$ be the surface operator defined by* (108). *Then*

$$\forall \mathbf{u} \in \mathbf{U}, \quad |(\mathbf{N}_{\mathscr{R}}\mathbf{u},\mathbf{u})_{\mathbf{U}}| \leqslant \|\mathscr{R}\|_{\mathscr{L}(TL^2(\Gamma))}\|\mathbf{u}\|_{\mathbf{U}}^2. \tag{112}$$

Proof. For simplicity we assume that $\mathscr{R}$ is nonzero. We have

$$(\mathbf{N}_{\mathscr{R}}\mathbf{u},\mathbf{u})_{\mathbf{U}} = (-\mathrm{i}\mathscr{R}(n \wedge M_1 - \mathrm{i}J_1), J_1)_{L^2} + (-n \wedge \mathscr{R}(n \wedge M_1 - \mathrm{i}J_1), M_1)_{L^2}$$
$$= (\|\mathscr{R}\|^{1/2} n \wedge M_1 + \mathrm{i}J_1, \|\mathscr{R}\|^{-1/2}\mathscr{R}(n \wedge M_1 - \mathrm{i}J_1))_{L^2},$$

and, by Cauchy–Schwartz inequality

$$|(\mathbf{N}_{\mathscr{R}}\mathbf{u},\mathbf{u})_{\mathbf{U}}| \leqslant \tfrac{1}{2}\,\|\mathscr{R}\|\,\|n \wedge M_1 + \mathrm{i}J_1\|_{L^2}^2 + \tfrac{1}{2}\,\|\mathscr{R}\|^{-1}\|\mathscr{R}(n \wedge M_1 - iJ_1)\|_{L^2}^2.$$

Then we use the boundedness of $\mathscr{R}$ and we expand the squares to get

$$|(\mathbf{N}_{\mathscr{R}}\mathbf{u},\mathbf{u})_{\mathbf{U}}| \leqslant \frac{\|\mathscr{R}\|}{2}(\|n \wedge M_1 + iJ_1\|_{L^2}^2 + \|n \wedge M_1 - \mathrm{i}J_1\|_{L^2}^2)$$
$$= \|\mathscr{R}\|\,\|\mathbf{u}\|_{\mathbf{U}}^2.$$

This estimate proves that the additional term $\mathbf{N}_{\mathscr{R}}\mathbf{u}$ is small compared to the other terms of the system. As a result

Theorem 5. *Let us assume that the reflexion operator is strictly bounded by one in L^2: $\|\mathscr{R}\|_{\mathscr{L}(TL^2(\Gamma))} < 1$. Then for β=0, the variational formulation* (111) *is well posed in $\mathbf{U}\times\mathbf{W}$. That is for every $\mathbf{f} \in \mathbf{U}$ there exists a unique $(\mathbf{u},\mathbf{v}) \in \mathbf{U}\times\mathbf{W}$ weak solution of* (110). *For $\beta > 0$ the variational formulation of* (110) *is well posed in $\mathbf{U}\times\mathbf{V}$. That is for every $\mathbf{f} \in \mathbf{U}$ there exists a unique $(\mathbf{u},\mathbf{v}) \in \mathbf{U}\times\mathbf{V}$ weak solution of* (110).

Under the weaker assumption $\|\mathscr{R}\|_{\mathscr{L}(TL^2(\Gamma))}=1$, *then an a priori estimate implies uniqueness of the solution in* $\mathbf{U}\times\mathbf{W}$ *for* $\beta=0$ *and in* $\mathbf{U}\times\mathbf{V}$ *for* $\beta>0$.

Proof. Let us consider $\mathscr{A}_{\beta,\mathscr{R}}$, the operator associated to (110) defined in the same manner as in (94). A simple calculation shows that

$$\Re e(\mathscr{A}_{\beta,\mathscr{R}}[\mathbf{u},\mathbf{v}]^t,[\mathbf{u},\mathbf{v}]^t)\geqslant\|\mathbf{u}\|_{\mathbf{U}}^2-\Re e(\mathbf{N}_{\mathscr{R}}\mathbf{u},\mathbf{u})_{\mathbf{U}}+\beta\|\mathbf{u}+\mathrm{i}\mathbf{v}\|_{\mathbf{U}}^2+\|\mathbf{A}_\infty\mathbf{u}\|_Z^2+\|\mathbf{A}_\infty\mathbf{v}\|_Z^2$$

and so, by virtue of (112)

$$\Re e(\mathscr{A}_{\beta,\mathscr{R}}[\mathbf{u},\mathbf{v}]^t,[\mathbf{u},\mathbf{v}]^t)\geqslant(1-\|\mathscr{R}\|)\|\mathbf{u}\|_{\mathbf{U}}^2+\beta\|\mathbf{u}+\mathrm{i}\mathbf{v}\|_{\mathbf{U}}^2+\|\mathbf{A}_\infty\mathbf{u}\|_Z^2+\|\mathbf{A}_\infty\mathbf{v}\|_Z^2.$$

Uniqueness: If $\mathbf{f}=0$, we have $\beta(\mathbf{u}+\mathrm{i}\mathbf{v})=0$ and $\mathbf{A}_\infty\mathbf{u}=\mathbf{A}_\infty\mathbf{v}=0$. Using the identity $\operatorname{Ker}\mathbf{A}_\infty=\operatorname{Ker}\mathbf{T}^*$ (see Appendix B), we get $\mathbf{T}^*\mathbf{u}=0$ and $\mathbf{T}^*\mathbf{v}=0$ and so $\mathbf{v}=0$ in $\mathbf{W}$. Furthermore the second equation of (110) gives $\mathbf{T}\mathbf{u}=0$ and so $\Pi\mathbf{u}=\mathbf{T}\mathbf{u}-\mathbf{T}^*\mathbf{u}=0$ then $\mathbf{u}=0$. If $\beta\neq0$, we also have $\mathbf{v}=-\mathrm{i}\mathbf{u}=0$ in $\mathbf{u}$.

Existence: We assume $\|\mathscr{R}\|<1$. We proceed as in the proof of Theorem 4. We transform $\mathscr{A}_{\beta,\mathscr{R}}$ into $\mathscr{A}_{\beta,\mathscr{R},\varepsilon}$ by adding the perturbation $\varepsilon\mathbf{T}\mathbf{T}^*\mathbf{v}$ in the left hand side of the second equation of (110). If $\beta\neq0$, the proof follows exactly the same steps of the proof of Theorem 4. We assume $\beta=0$. Let $(\mathbf{u}^\varepsilon,\mathbf{v}^\varepsilon)$ be the solution of the perturbated problem. The coercivity property and the triangular inequality provide

$$(1-\|\mathscr{R}\|)\|\mathbf{u}^\varepsilon\|^2\leqslant\|\mathbf{f}\|\,\|\mathbf{u}^\varepsilon\|\Rightarrow\|\mathbf{u}^\varepsilon\|\leqslant\frac{\|\mathbf{f}\|}{1-\|\mathscr{R}\|}$$

$$\|\mathbf{T}^*\mathbf{v}^\varepsilon\|\leqslant(1+\|\mathscr{R}\|+\|\mathbf{A}_\infty^*\mathbf{A}_\infty\|)\,\|\mathbf{u}^\varepsilon\|+\|\mathbf{f}\|.$$

Using the inf–sup condition (84) with $C=\frac{1}{2}$ (see Lemma 5), we get

$$\frac{1}{2}\|\mathbf{v}^\varepsilon\|_{\mathbf{W}}\leqslant\sup_{\mathbf{u}\in\mathbf{U}}\frac{(\mathbf{u},\mathbf{T}^*\mathbf{v}^\varepsilon)_{\mathbf{U}}}{\|\mathbf{u}\|_{\mathbf{U}}}=\|\mathbf{T}^*\mathbf{v}^\varepsilon\|$$

$$\leqslant(1+\|\mathscr{R}\|+\|\mathbf{A}_\infty^*\mathbf{A}_\infty\|)\frac{\|\mathbf{f}\|}{1-\|\mathscr{R}\|}+\|\mathbf{f}\|.$$

Finally, $(\mathbf{u}^\varepsilon,\mathbf{v}^\varepsilon)$ is bounded in $\mathbf{U}\times\mathbf{W}$; we conclude as in the proof of Theorem 4.

6. An iterative algorithm and its convergence

To solve (110), the iterative algorithm we propose in this section is a relaxed Jacobi method. Of course, this Jacobi algorithm is not necessarily the best one to solve *in practice* our integral equation system. But what we intend to do is to show that the new structure of the system can be exploited to get convergence results for iterative algorithms, which are difficult to obtain with the other classical formulations.

Let $1 > r > 0$ be some relaxation parameter. The algorithm reads

- *computation of F_0 as given in* (106) *in function of the data F then computation of the second term f in* (106)
- *initialization*: $\mathbf{u}^0 = \mathbf{v}^0 = 0$
- *loop over p*
  - *solve*

$$\begin{aligned}&(1+\beta)\tilde{\mathbf{u}}^p + (\mathbf{A}^\infty)^*\mathbf{A}^\infty\tilde{\mathbf{u}}^p - \mathbf{T}^*\tilde{\mathbf{v}}^p = \mathbf{f} - \mathbf{N}_{\mathscr{R}}\mathbf{u}^{p-1} - \mathrm{i}\beta\mathbf{v}^{p-1},\\ &-\mathbf{T}\tilde{\mathbf{u}}^p - \beta\tilde{\mathbf{v}}^p - (\mathbf{A}^\infty)^*\mathbf{A}^\infty\tilde{\mathbf{v}}^p = -\mathrm{i}\beta\mathbf{u}^{p-1}.\end{aligned} \tag{113}$$

  - *relax*

$$\begin{aligned}&\mathbf{u}^p = (1-r)\mathbf{u}^{p-1} + r\tilde{\mathbf{u}}^p,\\ &\mathbf{v}^p = (1-r)\mathbf{v}^{p-1} + r\tilde{\mathbf{v}}^p.\end{aligned} \tag{114}$$

We assume F to be in $TL^2(\Gamma)$, so that $\mathbf{f}$ is in $\mathbf{U}$. Here we assume that solving the discrete counterpart of (113) is easy due to the strong coercivity of the operator on the left hand side. For example discrete conjugate gradient might be used for this calculation. In Appendix C, the condition number of the system is studied in the special case of a sphere. It is shown to be moderate even for large spheres. Concerning the convergence, we have the following result where we assume that the solution $(\mathbf{u},\mathbf{v})$ exists and is in $\mathbf{U}\times\mathbf{V}$ even for $\|\mathscr{R}\| = 1$.

Theorem 6. *If $\|\mathscr{R}\| < 1$ then $(\mathbf{u}^p,\mathbf{v}^p)$ converges strongly to $(\mathbf{u},\mathbf{v})$ in $\mathbf{U}\times\mathbf{V}$. Using the weaker assumption $\|\mathscr{R}\| = 1$ if a solution of* (111) *exists then $(\mathbf{u}^p,\mathbf{v}^p)$ converges weakly to $(\mathbf{u},\mathbf{v})$ in $\mathbf{U}\times\mathbf{V}$.*

For simplicity, we will give a proof only valid in the case of a constant $\mathscr{R}$. Let us consider the algorithm for the error. It is initialized by $(\mathbf{u}^0,\mathbf{v}^0) = -(\mathbf{u},\mathbf{v})$ while the associated induction consists in systems (113)–(114) with $\mathbf{f}$ replaced by 0. We begin by multiplying the first (resp. second) equation of (113) by $\tilde{\mathbf{u}}^p$ (resp. by $\tilde{\mathbf{v}}^p$) then we subtract the real part of the two results. Both terms involving $-\mathbf{T}^*\tilde{\mathbf{u}}^p$ and $\mathbf{T}\tilde{\mathbf{u}}^p$ cancel each other and we get

$$\begin{aligned}&(1+\beta)\|\tilde{\mathbf{u}}^p\|_{\mathbf{U}}^2 + \beta\|\tilde{\mathbf{v}}^p\|_{\mathbf{U}}^2 + \|\mathbf{A}^\infty\tilde{\mathbf{u}}^p\|_Z^2 + \|\mathbf{A}^\infty\tilde{\mathbf{v}}^p\|_Z^2\\ &\quad = -\,\Re e(\mathrm{i}\beta\mathbf{v}^{p-1},\tilde{\mathbf{u}}^p)_{\mathbf{U}} + \Re e(\mathrm{i}\beta\mathbf{u}^{p-1},\tilde{\mathbf{v}}^p)_{\mathbf{U}} - \Re e(\mathbf{N}_{\mathscr{R}}\mathbf{u}^{p-1},\tilde{\mathbf{u}}^p)_{\mathbf{U}},\end{aligned} \tag{115}$$

where we recall that $Z = TL^2(S^2)$. The remaining part of the proof is just technical. First, the term multiplied by β reads

$$\begin{aligned}&\|\tilde{\mathbf{u}}^p\|_{\mathbf{U}}^2 + \|\tilde{\mathbf{v}}^p\|_{\mathbf{U}}^2 + \Re e(\mathrm{i}\mathbf{v}^{p-1},\tilde{\mathbf{u}}^p)_{\mathbf{U}} + \Re e(-\mathrm{i}\mathbf{u}^{p-1},\tilde{\mathbf{v}}^p)_{\mathbf{U}}\\ &\quad = \tfrac{1}{2}\|\tilde{\mathbf{u}}^p\|_{\mathbf{U}}^2 + \tfrac{1}{2}\|\tilde{\mathbf{v}}^p\|_{\mathbf{U}}^2 - \tfrac{1}{2}\|\mathbf{u}^{p-1}\|_{\mathbf{U}}^2 - \tfrac{1}{2}\|\mathbf{v}^{p-1}\|_{\mathbf{U}}^2\\ &\qquad + \tfrac{1}{2}\|\mathrm{i}\mathbf{u}^{p-1} - \tilde{\mathbf{v}}^p\|_{\mathbf{U}}^2 + \tfrac{1}{2}\|\mathrm{i}\mathbf{v}^{p-1} + \tilde{\mathbf{u}}^p\|_{\mathbf{U}}^2.\end{aligned} \tag{116}$$

Then, denoting $\tilde{\mathbf{u}}^p$ as $(\tilde{J}_1^p, \tilde{M}_1^p)$, $\mathbf{u}^{p-1}$ as (J_1^{p-1}, M_1^{p-1}), and using

$$F^p = \mathrm{i}\tilde{J}_1^p + n \wedge \tilde{M}_1^p + \mathscr{R}(\mathrm{i}J_1^{p-1} - n \wedge M_1^{p-1}),$$

we derive from definition (109)

$$\begin{aligned}
\|\tilde{\mathbf{u}}^p\|_{\mathbf{U}}^2 + (\mathbf{N}_{\mathscr{R}}\mathbf{u}^{p-1}, \tilde{\mathbf{u}}^p)_{\mathbf{U}} &= \frac{1}{2}\|\mathrm{i}\tilde{J}_1^p - n \wedge \tilde{M}_1^p\|^2 + \frac{1}{2}\|\mathrm{i}\tilde{J}_1^p + n \wedge \tilde{M}_1^p\|^2 \\
&\quad + \Re e(\mathscr{R}(\mathrm{i}J_1^{p-1} - n \wedge M_1^{p-1}), (\mathrm{i}\tilde{J}_1^p + n \wedge \tilde{M}_1^p) \\
&= \frac{1}{2}\|\mathrm{i}\tilde{J}_1^p - n \wedge \tilde{M}_1^p\|^2 + \frac{1}{2}\|F^p\|^2 - \frac{1}{2}\|\mathscr{R}(\mathrm{i}J_1^{p-1} - n \wedge M_1^{p-1})\|^2 \\
&\leqslant \frac{1}{2}\|\mathrm{i}\tilde{J}_1^p - n \wedge \tilde{M}_1^p\|^2 + \frac{1}{2}\|F^p\|^2 \\
&\quad - \frac{\|\mathscr{R}\|^2}{2}\|(\mathrm{i}J_1^{p-1} - n \wedge M_1^{p-1})\|^2. \qquad (117)
\end{aligned}$$

(we have used here the assumption that $\mathscr{R}$ is constant). Let us define the following norm on $\mathbf{X} = \mathbf{U} \times \mathbf{U}$:

$$\|v\|_{\mathbf{X}}^2 = \|(\mathbf{u}, \mathbf{v})\|_{\mathbf{X}}^2 = \frac{\beta}{2}\|\mathbf{u}\|_{\mathbf{U}}^2 + \frac{\beta}{2}\|\mathbf{v}\|_{\mathbf{U}}^2 + \frac{1}{2}\|\mathrm{i}J_1 - n \wedge M_1\|^2, \qquad (118)$$

with $\mathbf{u} = (J_1, M_1)$. With the help of estimates (116) and (117) we have

$$\|\tilde{v}^p\|_{\mathbf{X}}^2 = \|(\tilde{\mathbf{u}}^p, \tilde{\mathbf{v}}^p)\|_{\mathbf{X}}^2 \leqslant \|(\mathbf{u}^{p-1}, \mathbf{v}^{p-1})\|_{\mathbf{X}}^2 - \mathscr{V}^p = \|v^{p-1}\|_{\mathbf{X}}^2 - \mathscr{V}^p$$

with

$$\begin{aligned}
\mathscr{V}^p &= |\mathbf{A}^\infty\tilde{\mathbf{u}}^p|_Z^2 + |\mathbf{A}^\infty\tilde{\mathbf{v}}^p|_Z^2 + \frac{1}{2}\|F^p\|_{\mathbf{U}}^2 + \frac{\beta}{2}\|\mathrm{i}\mathbf{u}^{p-1} - \tilde{\mathbf{v}}^p\|_{\mathbf{U}}^2 \\
&\quad + \frac{\beta}{2}\|\mathrm{i}\mathbf{v}^{p-1} + \tilde{\mathbf{u}}^p\|_{\mathbf{U}}^2 + \frac{1}{2}(1 - \|\mathscr{R}\|^2)\|\mathrm{i}J_1^{p-1} - n \wedge M_1^{p-1}\|^2. \qquad (119)
\end{aligned}$$

Finally, using once more time the identity

$$2\Re e(\tilde{v}^p, v^{p-1})_{\mathbf{X}} = \|\tilde{v}^p\|_V^2 + \|v^{p-1}\|_{\mathbf{X}}^2 - \|\tilde{v}^p - v^{p-1}\|_{\mathbf{X}}^2, \qquad (120)$$

we get for $v^p = (1-r)v^{p-1} + r\tilde{v}^p$

$$\begin{aligned}
\|v^p\|_{\mathbf{X}}^2 &= (1-r)^2\|v^{p-1}\|_{\mathbf{X}}^2 + r^2\|\tilde{v}^p\|_{\mathbf{X}}^2 + (1-r)r(\|v^{p-1}\|_{\mathbf{X}}^2 + \|\tilde{v}^p\|_{\mathbf{X}}^2 - \|\tilde{v}^p - v^{p-1}\|_{\mathbf{X}}^2) \\
&\leqslant \|v^{p-1}\|_V^2 - r\mathscr{V}^p - r(1-r)\|\tilde{v}^p - v^{p-1}\|_{\mathbf{X}}^2, \qquad (121)
\end{aligned}$$

performing a summation of all those inequalities over p provides

$$\|v^q\|_{\mathbf{X}}^2 + \sum_{p=0}^{q} r(1-r)\|\tilde{v}^p - v^{p-1}\|_{\mathbf{X}}^2 + \sum_{p=0}^{q}(1-r)\mathscr{V}^q \leqslant \|\mathbf{u}\|_{\mathbf{U}}^2 + \|\mathbf{v}\|_{\mathbf{U}}^2. \qquad (122)$$

This estimate implies that $\|v^p\|_{\mathbf{X}}^2$ is bounded and that the two series are convergent and so

$$\|\mathbf{A}^\infty \tilde{\mathbf{u}}^p\|_Z^2 + \|\mathbf{A}^\infty \tilde{\mathbf{v}}^p\|_Z^2 \to 0, \quad p \to \infty. \tag{123}$$

Furthermore, if $\|\mathscr{R}\| < 1$, we have that $a^p = \mathrm{i}J_1^p - n \wedge M_1^p$ approaches 0 as p tends to infinity. It is now not difficult to check that the only convergence of the a^p's implies that $\mathbf{u}^p$ and $\mathbf{v}^p$ goes to zero as p goes to infinity.

If $\|\mathscr{R}\| = 1$, the convergence does not hold any longer in the strong sense. However, it is clear that since the sequence is bounded in $\mathbf{V}$, we can extract a subsequence that converges weakly. Since the only solution of $\mathscr{A}_{\beta,\mathscr{R}}(\mathbf{u},\mathbf{v})' = 0$ is 0, the limit point can only be 0. Thus, the only accumulation point is 0 and the whole sequence converges weakly.

Remark. We have readily obtained the strong convergence when the reflection coefficient is strictly less than 1, but there is no reason for the error to be a geometric decreasing function of the iterations. Indeed, all what we have obtained is that a series is convergent and consequently its generic term must go to 0. However if $\beta = 0$ and $\|\mathscr{R}\| < 1$ then (118) and (119) imply that $\|\mathrm{i}\tilde{J}_1^p - n \wedge \tilde{M}_1^p\|^2 \leqslant \|\mathscr{R}\|^2 \|\mathrm{i}J_1^{p-1} - n \wedge M_1^{p-1}\|^2$, which in turn implies a geometric convergence to 0.

7. Discussion of numerical issues

We would like to discuss the use of this method in practical computations, and split the discussion between the method by itself and the method coupled with other problems or other algorithms.

7.1. The method by itself

At first sight, the major drawback of this new integral system when compared to a classical one is that the number of unknowns has been multiplied by a factor 4. Nevertheless, this can be tempered by the fact that the matrix of the system can be easily split into four independent blocks (real and imaginary part are uncoupled and so are the unknowns $((J_k + M_k),\ k = 1, 2)$ and $((J_k - M_k),\ k = 1, 2)$). Thus, multiplication by the matrix of the new system is simply four time more expensive than for a classical system.

Another property is that the new system might appear to be well-suited for impedance conditions rather than for perfect conductors. Actually, the reflection coefficient is 1 for a perfect conductor and strong coercivity properties does not hold any longer in this case. Nevertheless, the positivity of the system remains valid even in this case. Furthermore, the numerical experiments of [3], concerning the new system in the 2-D case, shows that the method can either be used for the computation of the scattering of electromagnetic waves by perfect conductors.

However, our new system presents some advantages. The most important is obviously that we have now a system with a structure that allows us to use many standard algorithms for solving them (with convergence theorems). We proposed one in the previous section but many others could be contemplated. A second property is that the system appears to be generic for all boundary conditions: all we have to do when a new boundary condition is considered is to implement a solver for the reflection coefficient operator (i.e. a way to obtain f^{o} in $\mathscr{R} f^{\mathrm{i}} = f^{\mathrm{o}}$ or equivalently

$f^i + Z_r f^i = f^o - Z_r f^o$, with $\|f^i\| \leqslant \|f^o\|$). For instance, it is possible to handle problems with complicated boundary conditions, involving surface-to-surface differential operators as those coming from scattering by backed obstacles, [24] (see [3] for an example).

7.2. Coupled algorithms and coupled problems

Due to the strong coercivity properties of our integral systems, they are well suited for the coupling with interior problems treated for example with domain decomposition algorithms [14,15,26,8].

A very promising method for the numerical calculation of time-harmonic obstacle scattering solutions is the multipole method, see [16,25,10] among others. Multipoles methods is a way to speed up the matrix vector product for linear systems coming from integral equations. For those who are familiar with these methods, it is well known that the usual integral equations require at least 2 scalar (i.e. with scalar far and near fields) multipole computations, and even more when impedance problems are considered. For our system, 4 multipole computations are enough to compute $\mathbf{Tu}, \mathbf{T}^*\mathbf{v}$ and 4 other multipole computations (with a very simple translation function) provide $(\mathbf{A}^\infty)^*(\mathbf{A}^\infty)\mathbf{u}$ and $(\mathbf{A}^\infty)^*(\mathbf{A}^\infty)\mathbf{v}$. Furthermore the near interaction matrix required by the method for $\mathbf{T}$ is now real and so is halved in storage with respect to the classical ones. It is reasonable to think that coupling multipole methods and our integral system should give a good compromise in terms of an accurate, fast, robust and reliable algorithm.

Acknowledgements

The authors greatly acknowledge Houssem Haddar for his helpful comments.

Appendix A. Decomposition of the regular part of operator S

We start from

$$\frac{\sin(k|x-y|)}{4\pi|x-y|} = \frac{k}{4\pi^2}\int_{S^2} e^{ik(x-y)\cdot\hat{d}}\, d\sigma(\hat{d}), \tag{A.1}$$

where S^2 is the unit sphere. If J and J' are two tangential fields defined on Γ, we begin with $(T_i J, J')$. We have

$$(T_i J, J') = k\int_\Gamma\int_\Gamma\left(\frac{\sin(k|x-y|)}{4\pi|x-y|} J(y)\cdot J'(x)\right.$$
$$\left. - \frac{1}{k^2}\nabla_y^t\cdot J(y)\nabla_x^t\cdot J'(x)\right) d\Gamma(y)\, d\Gamma(\hat{d}). \tag{A.2}$$

Using (A.1) and interchanging the integrals, we get

$$(T_i J, J') = \frac{k^2}{4\pi^2}\int_{S^2}(\vec{A}J(\hat{d})\cdot\overline{\vec{A}J'}(\hat{d}) - AJ(\hat{d})\overline{AJ'}(\hat{d}))\, d\sigma(\hat{d}), \tag{A.3}$$

with

$$\vec{A}J(\hat{d}) = \int_\Gamma J(x)\mathrm{e}^{-\mathrm{i}kx\cdot\hat{d}}\,\mathrm{d}\Gamma(x), \quad AJ(\hat{d}) = \frac{1}{\mathrm{i}k}\int_\Gamma \nabla_x^t \cdot J(x)\mathrm{e}^{-\mathrm{i}kx\cdot\hat{d}}\,\mathrm{d}\Gamma(x). \tag{A.4}$$

Using an integration by parts provides

$$AJ(\hat{d}) = \frac{1}{\mathrm{i}k}\int_\Gamma J(x)\cdot\nabla_x^t \mathrm{e}^{\mathrm{i}kx\cdot\hat{d}}\,\mathrm{d}\Gamma(x) = \int_\Gamma J(x)\mathrm{e}^{\mathrm{i}kx\cdot\hat{d}}\,\mathrm{d}\Gamma(x)\cdot\hat{d}. \tag{A.5}$$

Thus (A.2) reads

$$(T_iJ,J') = \frac{k^2}{4\pi^2}\int_{S^2} \vec{A}J(\hat{d})\cdot\overline{\vec{A}J'}(\hat{d}) - ((\vec{A}J(\hat{d})\cdot\hat{d})\hat{d})\cdot((\overline{\vec{A}J'(\hat{d})}\cdot\hat{d})\hat{d})\,\mathrm{d}\sigma(\hat{d}). \tag{A.6}$$

Now, we consider the splitting

$$\vec{A}J(\hat{d}) = \hat{d}\wedge(\vec{A}J(\hat{d})\wedge\hat{d}) + (\vec{A}J(\hat{d})\cdot\hat{d})\hat{d}, \tag{A.7}$$

and the similar expression for $(\vec{A}J')(\hat{d})$. Remarking that the two vectors in the above decomposition are orthogonal, we have

$$(T_iJ,J') = \frac{k^2}{4\pi^2}\int_{S^2}(\hat{d}\wedge(\vec{A}J(\hat{d})\wedge\hat{d}))\cdot(\hat{d}\wedge(\overline{\vec{A}J'}(\hat{d})\wedge\hat{d}))\,\mathrm{d}\sigma(\hat{d}), \tag{A.8}$$

or (see (23))

$$(T_iJ,J') = \int_{S^2} a^\infty J(\hat{d})\cdot\overline{a^\infty J'(\hat{d})}\,\mathrm{d}\sigma(\hat{d}). \tag{A.9}$$

We can either get a similar expression for (T_iM,M'), but we modify it into

$$(T_iM,M') = \int_{S^2}(-\mathrm{i}\hat{d}\wedge a^\infty M(\hat{d}))\cdot(\overline{-\mathrm{i}\hat{d}\wedge a^\infty M'(\hat{d})})\,\mathrm{d}\sigma(\hat{d}). \tag{A.10}$$

Now we turn to (K_iM,J'), from the definition of K_i we have

$$(K_iM,J') = \int_\Gamma\int_\Gamma\left(\nabla_y^t\frac{\sin(k|x-y|)}{4\pi|x-y|}\wedge M(y)\right)\cdot\overline{J'(x)}\,\mathrm{d}\Gamma(y)\,\mathrm{d}\Gamma(\hat{d}), \tag{A.11}$$

or

$$\begin{aligned}(K_iM,J') &= \frac{k}{(4\pi)^2}\int_{S^2}\left(\int_\Gamma\nabla_y^t\mathrm{e}^{-\mathrm{i}ky\hat{d}}\wedge M(y)\,\mathrm{d}\Gamma(y)\cdot\int_\Gamma\overline{J'(x)\mathrm{e}^{-\mathrm{i}kx\hat{d}}}\,\mathrm{d}\Gamma(\hat{d})\right)\mathrm{d}\sigma(\hat{d})\\ &= \frac{k^2}{(4\pi)^2}\int_{S^2}\left(-\mathrm{i}\hat{d}\wedge\int_\Gamma M(y)\mathrm{e}^{-\mathrm{i}ky\hat{d}}\,\mathrm{d}\Gamma(y)\cdot\int_\Gamma\overline{J'(x)\mathrm{e}^{-\mathrm{i}kx\hat{d}}}\,\mathrm{d}\Gamma(\hat{d})\right)\mathrm{d}\sigma(\hat{d}).\end{aligned}$$

Using once again the decomposition

$$J' = \hat{d} \wedge (J'(\hat{d}) \wedge \hat{d}) + (J'(\hat{d}) \cdot \hat{d})\hat{d}, \tag{A.12}$$

we get

$$(K_i M, J') = \int_{S^2} (-\mathrm{i}\hat{d} \wedge a^\infty M(\hat{d})) \cdot (\overline{a^\infty J'(\hat{d})})\, \mathrm{d}\sigma(\hat{d}). \tag{A.13}$$

Similar calculations show that

$$(K_i J, M') = \int_{S^2} (a^\infty J(\hat{d})) \cdot (\overline{-\mathrm{i}\hat{d} \wedge a^\infty M'(\hat{d})})\, \mathrm{d}\sigma(\hat{d}), \tag{A.14}$$

and finally

$$\begin{aligned}&(T_i J + K_i M, J') + (K_i J + T_i M, M')\\ &\quad = \int_{S^2} (\mathbf{a}^\infty J(\hat{d}) - \mathrm{i}\hat{d} \wedge \mathbf{a}^\infty M(\hat{d})) \cdot \overline{(\mathbf{a}^\infty J'(\hat{d}) - \mathrm{i}\hat{d} \wedge \mathbf{a}^\infty M'(\hat{d})}\, \mathrm{d}\sigma(\hat{d}).\end{aligned} \tag{A.15}$$

Appendix B. Calderon projectors

Our aim in this section is to make the link between the operators we have defined in Section 2 with the Calderon projectors, [6, p. 87], [7, p. 93].

Let $(\hat{J}_1, \hat{M}_1)$ a pair of tangential fields given on Γ, (not necessarily corresponding to the tangential traces of an exterior electro-magnetic field). We can associate to $(\hat{J}_1, \hat{M}_1)$ the two fields in Ω^+:

$$\begin{aligned}&\frac{E^+(x)}{\sqrt{\mathrm{i}Z_0}} = \tilde{T}\hat{J}_1(x) + \tilde{K}\hat{M}_1(x),\\ &-\sqrt{\mathrm{i}Z_0}H^+(x) = \tilde{K}\hat{J}_1(x) + \tilde{T}\hat{M}_1(x),\end{aligned} \tag{B.1}$$

where $\tilde{T}J, \tilde{K}M$ are given in (7). If x approaches a point of Γ, the jump conditions provide

$$n(x) \wedge \left(\frac{1}{\sqrt{\mathrm{i}Z_0}} E^+_{/\Gamma}(x) \wedge n(x)\right) = T\hat{J}_1(x) + K\hat{M}_1(x) + \tfrac{1}{2} n(x) \wedge \hat{M}_1(x), \tag{B.2}$$

$$n(x) \wedge (-\sqrt{\mathrm{i}Z_0} H^+_{/\Gamma}(x) \wedge n(x)) = T\hat{M}_1(x) + K\hat{J}_1(x) + \tfrac{1}{2} n(x) \wedge \hat{J}_1(x). \tag{B.3}$$

Now, we can proceed exactly as in Section 2: from the exterior traces, we construct the fields (J_1, M_1) and $\mathbf{u}$ by (14) and we have

$$\mathbf{Su} = \mathbf{S}\begin{bmatrix} J_1 \\ M_1 \end{bmatrix} = 0. \tag{B.4}$$

But, it is easy to see that (B.3)–(B.2) reads (see Definitions (21) and (17))

$$-\Pi\begin{bmatrix} J_1 \\ M_1 \end{bmatrix} = (\mathbf{S}-\Pi)\begin{bmatrix} \hat{J}_1 \\ \hat{M}_1 \end{bmatrix}. \tag{B.5}$$

Multiplying by $\mathbf{S}\Pi$ and using both $-\Pi^2 = \mathbf{Id}$ and (B.4) we get

$$0 = \mathbf{S}\begin{bmatrix} J_1 \\ M_1 \end{bmatrix} = (\mathbf{S}\Pi\mathbf{S}+\mathbf{S})\begin{bmatrix} \hat{J}_1 \\ \hat{M}_1 \end{bmatrix}. \tag{B.6}$$

In other words,

$$-\Pi\mathbf{S} = (-\Pi\mathbf{S})^2, \tag{B.7}$$

appears as a projector: it is one of the Calderon projector. With our decomposition in real and imaginary part $\mathbf{S} = \mathbf{T} + \mathrm{i}\mathbf{R}$, we get

$$\begin{aligned} &\mathbf{T}\Pi\mathbf{T} - \mathbf{R}\Pi\mathbf{R} = -\mathbf{T}, \\ &\mathbf{T}\Pi\mathbf{R} + \mathbf{R}\Pi\mathbf{T} = -\mathbf{R}. \end{aligned} \tag{B.8}$$

So far, all these calculations corresponds to Green function (8). We can proceed exactly in the same manner with

$$G(x,y) = \frac{\exp^{-\mathrm{i}k|x-y|}}{4\pi|x-y|} \tag{B.9}$$

the only modification being the radiation condition at infinity. The $1/(4\pi|x-y|)$ singularity of the kernel remaining unchanged, equality (B.7) holds also for $\tilde{\mathbf{S}} = \mathbf{T} - \mathrm{i}\mathbf{R}$. It expands into

$$\begin{aligned} &\mathbf{T}\Pi\mathbf{T} + \mathbf{R}\Pi\mathbf{R} = -\mathbf{T}, \\ &\mathbf{T}\Pi\mathbf{R} + \mathbf{R}\Pi\mathbf{T} = -\mathbf{R}. \end{aligned} \tag{B.10}$$

By comparison we get

$$\mathbf{T}\Pi\mathbf{T} = -\mathbf{T} \quad \text{and} \quad \mathbf{R}\Pi\mathbf{R} = 0 \tag{B.11}$$

and $-\Pi\mathbf{T}$ is found to be also a projector. It remains to prove that

$$\mathbf{R}\Pi\mathbf{T}^* = \mathbf{R}. \tag{B.12}$$

This is a consequence of the fact that free fields are in the kernel of "exterior" integral operators [9]. To see that, let us consider a free field (i.e. a field with continuous traces on Γ) in the form $\mathbf{R}\mathbf{v}$ for some arbitrary smooth $\mathbf{v}$. Let $(J,M) = \Pi\mathbf{R}\mathbf{v}$. By integration by parts we get the equivalent of (5)

$$\begin{aligned} &0 = \mathrm{i}Z_0(\tilde{T}J)(\hat{x}) + (\tilde{K}M)(\hat{x}), \\ &0 = -(\tilde{K}J)(\hat{x}) + \mathrm{i}Z_0^{-1}(\tilde{T}M)(\hat{x}), \end{aligned}$$

for $\hat{x}$ in D^+. If x approaches a point of Γ, using (9), (17) and (21)–(22), we get that

$$(\mathbf{T}+\Pi+\mathrm{i}\mathbf{R})\Pi\mathbf{R}\mathbf{v}=0.$$

It means that $\mathbf{T}\Pi\mathbf{R}-\mathbf{R}=0$ which implies by transposition the sought equality (B.12). A consequence is the continuity bound

$$\|\mathbf{R}\mathbf{v}\|_{\mathbf{U}} \leqslant \|\mathbf{R}\|_{\mathscr{L}(\mathbf{U})}\|\mathbf{T}^*\mathbf{v}\|_{\mathbf{U}}. \tag{B.13}$$

Another consequence is that the kernel of $\mathbf{T}^*$ is imbedded in the kernel of $\mathbf{A}^\infty$. Indeed, it can be proved that $\mathbf{A}^\infty$ has a dense range in Z (see [9] for a proof about operators of the same kind) and so $(\mathbf{A}^\infty)^*$ is injective. Since $\mathbf{R}$ is $(\mathbf{A}^\infty)^*\mathbf{A}^\infty$, equation (B.12) implies

$$\mathbf{A}^\infty\Pi\mathbf{T}^* = \mathbf{A}^\infty, \tag{B.14}$$

and $\mathbf{T}^*\mathbf{u}=0$ implies $\mathbf{A}^\infty\mathbf{u}=0$. The converse also holds. Let $\mathbf{u}$ be such that $\mathbf{A}^\infty\mathbf{u}=0$. To $\mathbf{u}=(J_1,M_1)$, we associate the outgoing electromagnetic field in $\Omega^+(E^+,H^+)$ defined in (B.1). It is classical to show that this electromagnetic field is such that if

$$\begin{aligned}(\mathrm{i}Z_0)^{-1/2}E^\infty(x) &= \frac{\exp \mathrm{i}k|x|}{|x|}\,\mathbf{A}^\infty(\mathbf{u};\hat{x}),\ \ (\mathrm{i}Z_0)^{+1/2}H^\infty(x)\\ &= \frac{\exp \mathrm{i}k|x|}{|x|},\ \ \hat{x}\wedge\mathbf{A}^\infty(\mathbf{u};\hat{x}),\end{aligned}$$

we have

$$\lim_{|x|\to\infty}\frac{1}{R}\int_{|x|=R} Z_0|E^+(x)-E^\infty(x)|^2+|H^+(x)-H^\infty(x)|^2\,\mathrm{d}\sigma=0.$$

Thus, $\mathbf{A}^\infty\mathbf{u}=0$ means that the associated outgoing electromagnetic field is associated to a null far field. From both Rellich Lemma and the unique continuation principle, [9], we conclude that (E^+,H^+) vanishes in Ω^+, and, in particular the tangential trace on Γ vanish. Using the jump conditions, following the same way as in Section 2, we readily obtain that $\mathbf{T}^*\mathbf{u}+\mathrm{i}\mathbf{R}\mathbf{u}=0$, what implies $\mathbf{T}^*\mathbf{u}=0$ since $\mathbf{R}\mathbf{u}=(\mathbf{A}^\infty)^*\mathbf{A}^\infty\mathbf{u}=0$.

Appendix C. An example: the case of the sphere

We specialize our study to the special case of a spherical scatterer where all the calculations can be done analytically. We define

$$\begin{aligned}\tilde{Y}_n^m(\hat{r}) &= P_n^{|m|}(\cos\theta)\mathrm{e}^{\mathrm{i}m\varphi}, \quad n\geqslant 0,\ \ |m|\leqslant n,\\ d_{n,m} &= \frac{1}{4\pi}\,\frac{(n-|m|)!(2n+1)}{(n+|m|)!n(n+1)},\end{aligned} \tag{C.1}$$

$$u_{nm}^{(+)}(\hat{r}) = d_{n,m}^{1/2} \nabla^t \tilde{Y}_n^m(\hat{r}),$$
$$u_{nm}^{(-)}(\hat{r}) = d_{n,m}^{1/2} \hat{r} \wedge \nabla^t \tilde{Y}_n^m(\hat{r}), \tag{C.2}$$

where (θ, φ) are the usual spherical coordinates and ∇^t is $(1/a)\hat{\theta}(\partial/\partial\theta) + (1/a\sin\theta)\hat{\varphi}(\partial/\partial\varphi)$. Any tangential field on a sphere of radius a may be decomposed into

$$J(\hat{r}) = \sum_{n=1}^{\infty} \sum_{m=-n}^{+n} \sum_{\varepsilon=\pm} J_{mn}^{(\varepsilon)} u_{nm}^{(\varepsilon)}(\hat{r}). \tag{C.3}$$

The $u_{nm}^{(\varepsilon)}$'s form a complete set of orthonormal functions in the space of the square integrable complex fields on the sphere. We have

$$\|J(\hat{r})\|_{TL^2(S_a^2)}^2 = \sum_{n=1}^{\infty} \sum_{m=-n}^{+n} \sum_{\varepsilon=\pm} |J_{mn}^{(\varepsilon)}|^2. \tag{C.4}$$

In [18], Hsiao and Kleinman used this decomposition to study two potentials denoted here by T_{HS} and K_{HS} and which are related to ours by the relationships

$$T = T_r + \mathrm{i}T_i = \mathrm{i}\hat{r} \wedge T_{\mathrm{HS}},$$
$$K = K_r + \mathrm{i}K_i = -\hat{r} \wedge K_{\mathrm{HS}}. \tag{C.5}$$

It is shown in this paper that if j_n and y_n are the spherical Bessel functions and

$$J_n = kaj_n(ka), \quad J_n' = (kaj_n(ka))',$$
$$Y_n = kay_n(ka), \quad Y_n' = (kay_n(ka))', \tag{C.6}$$

then, (see (78), (79), (85), (86) of [19])

$$(T_{HK} u_{nm}^{\varepsilon})(\hat{r}) = T_n^{\varepsilon} u_{nm}^{-\varepsilon}(\hat{r}),$$
$$(K_{HK} u_{nm}^{\varepsilon})(\hat{r}) = K_n^{\varepsilon} u_{nm}^{\varepsilon}, \tag{C.7}$$

with

$$T_n^+ = -J_n'(J_n' + \mathrm{i}Y_n'), \quad T_n^- = J_n(J_n + \mathrm{i}Y_n),$$
$$K_n^+ = -\tfrac{1}{2} - \mathrm{i}J_n(J_n' + \mathrm{i}Y_n'), \quad K_n^- = -\tfrac{1}{2} + \mathrm{i}J_n'(J_n + \mathrm{i}Y_n). \tag{C.8}$$

(note: let us remark that formulae (C.8) differs from those given in Hsiao et al.'s paper by a change of sign. The reason is that there is an error of sign in formulas (68) and (69) of this paper: $-\mathrm{i}k$ and $-\mathrm{i}k^2$ must be changed into $\mathrm{i}k$ and $\mathrm{i}k^2$. But, except this point all the remaining calculations are valid hence the only change of sign in the result).

From this and after some algebraic manipulations we easily deduce for T

$$(T_r u_{nm}^+(\hat{r}) = -J_n' Y_n' u_{nm}^+(\hat{r}), \quad (T_i u_{nm}^+)(\hat{r}) = (J_n')^2 \; u_{nm}^+(\hat{r}),$$
$$(T_r u_{nm}^-)(\hat{r}) = -J_n Y_n u_{nm}^-(\hat{r}) \quad (T_i u_{nm}^-)(\hat{r}) = J_n^2 u_{nm}^-(\hat{r}), \tag{C.9}$$

while for K

$$\left(K_r + \varepsilon\frac{\hat{r}}{2}\wedge\right) u^{\varepsilon}_{nm}(\hat{r}) = -J'_n Y_n u^{-\varepsilon}_{nm}(\hat{r}), \quad (K_i u^{\varepsilon}_{nm})(\hat{r}) = J'_n J_n u^{-\varepsilon}_{nm}(\hat{r}),$$
$$\left(K_r - \varepsilon\frac{\hat{r}}{2}\wedge\right) u^{\varepsilon}_{nm}(\hat{r}) = -J_n Y'_n\ u^{-\varepsilon}_{nm}(\hat{r}). \tag{C.10}$$

Our aim is to apply this harmonic analysis for the Jacobi algorithm described in Section 6. At each step of the induction, the solution of the linear system (113) is required. It reads

$$\begin{bmatrix} (1+\beta)+T_i & K_i \\ K_i & (1+\beta)+T_i \end{bmatrix}\begin{bmatrix} J_1 \\ M_1 \end{bmatrix} - \begin{bmatrix} T_r & K_r+\frac{\hat{r}}{2} \\ K_r+\frac{\hat{r}}{2} & T_r \end{bmatrix}\begin{bmatrix} J_2 \\ M_2 \end{bmatrix} = \begin{bmatrix} G_1 \\ H_1 \end{bmatrix}, \tag{C.11}$$

and

$$-\begin{bmatrix} T_r & K_r-\frac{\hat{r}}{2} \\ K_r-\frac{\hat{r}}{2} & T_r \end{bmatrix}\begin{bmatrix} J_1 \\ M_1 \end{bmatrix} - \begin{bmatrix} \beta+T_i & K_i \\ K_i & \beta+T_i \end{bmatrix}\begin{bmatrix} J_2 \\ M_2 \end{bmatrix} = \begin{bmatrix} G_2 \\ H_2 \end{bmatrix}. \tag{C.12}$$

This system can be split into two blocks

$$\mathbf{M}^{\pm}\begin{bmatrix} J_1 \pm M_1 \\ J_2 \pm M_2 \end{bmatrix} = \begin{bmatrix} G_1 \pm H_1 \\ G_2 \pm H_2 \end{bmatrix}, \tag{C.13}$$

with

$$\mathbf{M}^{\pm} = \begin{bmatrix} (1+\beta)+T_i \pm K_i & -T_r \mp (K_r+\frac{\hat{r}}{2}) \\ -T_r \mp (K_r-\frac{\hat{r}}{2}) & -\beta-(T_i \pm K_i) \end{bmatrix}. \tag{C.14}$$

Using the decomposition with the basis functions, we have a block structure for $\mathbf{M}^{\pm}$, each block being 4×4 (caution, the sign $\pm$ does not have the same meaning as the one in (C.2))

$$\mathbf{M}^{\pm} = \bigoplus_{n,m,|m|\leqslant n} (\mathbf{M})^{\pm}_{mn} \tag{C.15}$$

with

$$\mathbf{M}^{\pm}_{mn} = \begin{bmatrix} 1+\beta+(J'_n)^2, & \pm J_n J'_n, & J'_n Y'_n, & \pm J_n Y'_n \\ \pm J_n J'_n, & 1+\beta+J_n^2, & \pm J'_n Y_n & J_n Y_n \\ J'_n Y'_n, & \pm J'_n Y_n & -\beta-(J'_n)^2 & \mp J_n J'_n \\ \pm J_n Y'_n, & J_n Y_n & \mp J_n J'_n & -\beta-J_n^2 \end{bmatrix}. \tag{C.16}$$

Remark that $\mathbf{M}^{\pm}_{mn}$ is independent of m as we can expect since the problem is invariant under any rotation. Furthermore, it is in-lighting to interpret each system through a decomposition into

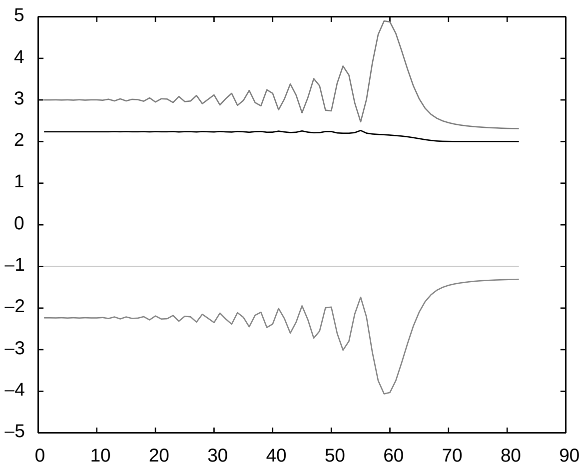

Fig. 1. The 4 eigenvalues of the matrix $\mathbf{M}^{\pm}_{mn}$ as function of n for $ka = 20\pi$.

four 2×2 blocks

$$\mathbf{M}^{\pm}_{mn} = \begin{bmatrix} (1+\beta)Id + (A^{\pm}_{nm})^*(A^{\pm}_{nm}) & (B^{\pm}_{nm})^*(A^{\pm}_{nm}) \\ (A^{\pm}_{nm})^*(B^{\pm}_{nm}) & -\beta Id - (A^{\pm}_{nm})^*(A^{\pm}_{nm}) \end{bmatrix}, \tag{C.17}$$

with

$$(A^{\pm}_{nm}) = \begin{bmatrix} J'_n \\ \pm J_n \end{bmatrix}, \quad (B^{\pm}_{nm}) = \begin{bmatrix} Y'_n \\ \pm Y_n \end{bmatrix}. \tag{C.18}$$

From the Wronskian for Bessel functions: $J_n Y'_n - J'_n Y_n = 1$, it is readily seen that

$$T^{\pm}_{nm} - (T^{\pm}_{nm})^* = (A^{\pm}_{nm})^*(B^{\pm}_{nm}) - (B^{\pm}_{nm})^*(A^{\pm}_{nm}) = \Pi^{\pm}_{nm},$$

$$\Pi^{\pm}_{nm} = \begin{bmatrix} 0 & \mp 1 \\ \pm 1 & 0 \end{bmatrix}, \quad (\Pi^{\pm}_{nm})^2 = Id. \tag{C.19}$$

For each small independent system, the algebraic structure of the original saddle point problem is recovered.

In Figs. 1 and 2 are depicted the 4 real eigenvalues of $(\mathbf{M})^{+}_{nm}$ (or those of $(\mathbf{M})^{-}_{nm}$ as they coincide) as function of n. Coefficient β was chosen equal to 1. Fig. 1 corresponds to a sphere of moderate size with respect to the wavelength ($ka = 20\pi$).

In Fig. 2, ka equals 200π and the sphere is large with a radius of 100 wavelengths. Note that the spectrum is real, as expected, and the range of the eigenvalues does not change so much between the two experiments.

Extensive calculations with Matlab has shown that the maximum modulus of the eigenvalues in the area $n \leqslant ka + 10\log(ka + \pi)$ is a low increasing function of the frequency. This area corresponds

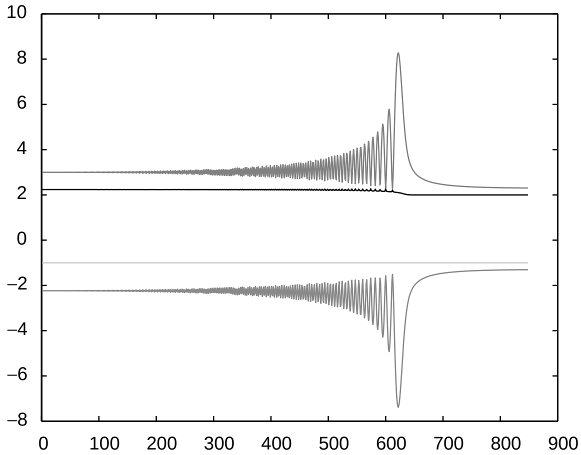

Fig. 2. The 4 eigenvalues of the matrix $\mathbf{M}_{mn}^{\pm}$ as function of n for $ka=200\pi$.

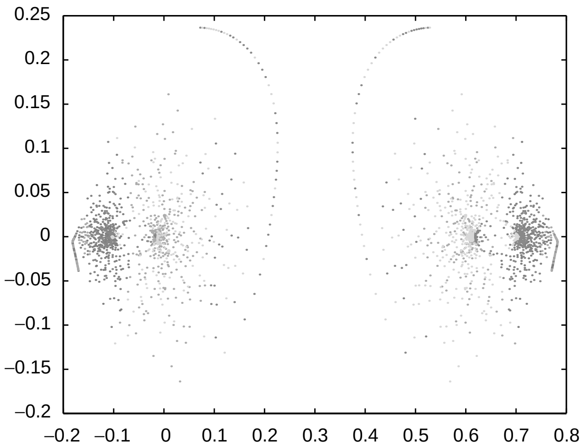

Fig. 3. Spectrum of the iteration matrix for the Jacobi algorithm applied to a sphere ($ka=200\pi$). Coefficient β is 1, relaxation parameter is $r=0.7$ and reflexion coefficient is $R=0$ (model problem).

to the excited modes for incident plane waves. For larger n, the asymptotic behavior of Bessel functions enables us to obtain $-n$, $-\beta$, $1+\beta$ and $+n$ as asymptotic eigenvalues. The fact that the L^2 spectrum is not bounded is of course to be related to the use of the functional framework $\mathbf{U}\times\mathbf{U}$ instead of $\mathbf{U}\times\mathbf{W}$ in our analysis, see Section 4.

Fig. 3 shows the spectrum of the matrix associated to the error of the Jacobi Algorithm for the large sphere. This spectrum is composed of the set of all the eigenvalues of the iteration matrices (Id is the 2×2 identity matrix)

$$\mathbf{E}_{mn}^{\pm}=(1-r)\begin{bmatrix} Id, & 0 \\ 0, & Id \end{bmatrix}-\mathrm{i}r\beta(\mathbf{M}_{mn}^{\pm})^{-1}\begin{bmatrix} 0, & Id \\ Id, & 0 \end{bmatrix}. \tag{C.20}$$

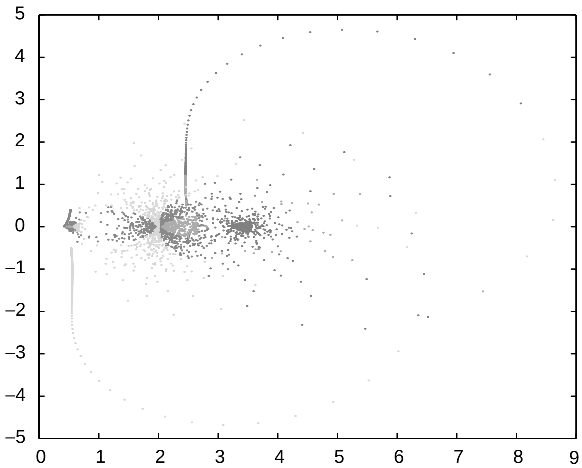

Fig. 4. Spectrum of the coercive matrix for the model problem. Case of the sphere with $ka = 200\pi$. Coefficient β is 1.

In that example, the reflexion coefficient is equal to zero (model problem) and the relaxation parameter is 0.7. All the eigenvalues are found located inside a circle of radius 0.782 and geometrical convergence occurs.

Other algorithms than Jacobi can be used to solve the problem. They can be constructed from our linear system in its coercive version. Fig. 4 shows the spectrum for the coercive matrix (90), let

$$\mathbf{A}_{\beta^{\pm}_{mn}} = \begin{bmatrix} (1+\beta)Id + (A^{\pm}_{nm})^*(A^{\pm}_{nm}) & (B^{\pm}_{nm})^*(A^{\pm}_{nm}) + \mathrm{i}\beta \\ -(A^{\pm}_{nm})^*(B^{\pm}_{nm}) - \mathrm{i}\beta & +\beta Id + (A^{\pm}_{nm})^*(A^{\pm}_{nm}) \end{bmatrix}, \tag{C.21}$$

Parameters β, R and ka are unchanged with respect to the previous examples: $\beta=1$, $R=0$, $ka=200\pi$. The coercivity property can be checked on the figure since all eigenvalues are located to the right of a line $\Re z = c$ with $c \approx 0.41$.

References

[1] T. Abboud, T. Sayah, Potentiels retardés pour les Equations de Maxwell avec condition d'impédance généralisée, Technical Report R.I. 387, Ecole Polytechnique, France, 2000.

[2] X. Antoine, H. Barucq, A. Bendali, Bayliss–Turkel-like radiation conditions on surfaces of arbitrary shapes, J. Math. Anal. Appl. 229 (2000) 184–211.

[3] N. Bartoli, F. Collino, Integral equations via saddle point problems for the 2-D electromagnetic problems, M2AN. Mathematical Modelling and Numerical Analysis 34 (5) (2000) 1023–1050.

[4] A. Bendali, Boundary element solution of scattering problems relative to a generalized impedance boundary condition, Proceedings of the Boca Raton Cerfacs Conference, London, New York, 1999.

[5] F. Brezzi, M. Fortin, Mixed and Hybrid Finite Element Methods, No. 15, Springer Series in Computational Mathematics, Springer, Berlin, 1991.

[6] M. Cessenat, Mathematical Methods in Electromagnetism, No. 41, Series on Advances in Mathematics for Applied Sciences, World-Scientific, Singapore, 1996.

[7] G. Chen, J. Zhou, Boundary Element Methods, Academic Press, New York, 1992.

[8] F. Collino, S. Ghanemi, P. Joly, Domain decomposition method for the Helmholtz equation: a general presentation, Comput. Methods Appl. Mech. Eng. 184 (2–4) (2000) 171–211.
[9] D. Colton, R. Kress, Inverse acoustic and electromagnetic scattering theory, in: Applied Mathematical Sciences, Vol. 93, Springer, Berlin, 1992.
[10] E. Darve, The fast multipole method: numerical implementation, J. Comput. Phys. 160 (1) (2000) 195–240.
[11] B. Després, Fonctionnelle quadratique et équations intégrales pour les équations de Maxwell en domaine extérieur, Comptes Rendus de l'Académie des Sciences, Paris, Série I 323 (1996) 547–552.
[12] B. Després, Fonctionnelle quadratique et Équations Integrales pour les problèmes d'onde harmonique en domaine extérieur, M2AN. Mathematical Modelling and Numerical Analysis 31 (1997) 679–732.
[13] B. Després, Quadratic Functional and Integral Equations for Harmonic Wave Equations, in Mathematical and Numerical Aspects of Wave Propagation (Golden, CO, 1998), SIAM, Philadelphia, 1998, pp. 56–64.
[14] B. Després, P. Joly, J.E. Roberts, A domain decomposition method for the harmonic Maxwell's equations, in: Proceedings of the IMACS International Symposium on Iterative Methods in Linear Algebra, North-Holland, Amsterdam, 1990.
[15] B. Després, B. Stupfel, A domain decomposition for the solution of large electromagnetic scattering problem, J. Electromagn. Waves Appl. 13 (11) (1999) 1553–1568.
[16] M. Epton, B. Dembart, Multipole translation theory for the three-dimensional Laplace and Helmholtz equations, SIAM J. Sci. Comput. 16 (1995) 865–897.
[17] P. Hartman, C. Wilcox, On solutions of the Helmholtz equations in exterior domains, Math. Z. 75 (1961) 228–255.
[18] G.C. Hsiao, R.F. Kleinman, Mathematical foundations for error estimations in numerical solutions of integral equations in electromagnetism, IEEE Trans. Antennas Propag. 45 (3) (1997) 316–328.
[19] R.E. Kleinman, P.M. Van den Berg, Iterative Methods for Solving Integral Equations, Elsevier, Amsterdam, 1988.
[20] J.C. Nédelec, Cours de l'école d'été d'analyse numérique, Technical Report, CEA-EDF-IRIA, 1977.
[21] M.D. Pocock, S.P. Walker, The complex bi-conjugate gradient solver applied to large electromagnetic scattering problems, computational costs, and costs scaling, IEEE Trans. Antennas Propag. 45 (1) (1997) 140–146.
[22] M. Reed, M. Simon, Scattering Theory, Academic Press, New York, 1979.
[23] T.K. Sarkar, Application of conjugate gradient method to electromagnetics and signal analysis, in: From Reaction Concept to Conjugate Gradient: Have we Made any Progress?, Elsevier, Amsterdam, 1988.
[24] T.B.A. Senior, J.L. Volakis, Approximate Boundary Conditions in Electromagnetism, No. 41, in: IEEE Electromagnetism Waves Series, The Institution of Electric Engineers, London, 1995.
[25] J. Song, C.C. Lu, W.C. Chew, Multilevel Fast Multipole Algorithm for Electromagnetic Scattering by Large Complex Objects, IEEE Trans. Antennas Propag. 45 (1997) 1488–1493.
[26] B. Stupfel, A hybrid finite element method and integral equation domain decomposition method for the solution of the 3-D scattering problem, J. Comput. Phys. 172 (2) (2001) 451–471.
[27] B. Stupfel, B. Despres, A domain decomposition method for the solution of large electromagnetic scattering problems, J. Electromagn. Waves Appl. 13 (1999) 1553–1568.