NON-REFLECTING BOUNDARY CONDITIONS
FOR WAVEGUIDES

A. BENDALI AND PH. GUILLAUME

ABSTRACT. New non-reflecting boundary conditions are introduced for the solution of the Helmholtz equation in a waveguide. These boundary conditions are perfectly transparent for all propagating modes. They do not require the determination of these propagating modes but only their propagation constants. A quasi-local form of these boundary conditions is well suited as terminating boundary condition beyond finite element meshes. Related convergence properties to the exact solution and optimal error estimates are established.

1. INTRODUCTION

We consider time-harmonic scalar waves propagating in a domain extending to infinity. The unbounded regions of this domain consist of semi-bounded waveguides. For this type of domain, any standard solution procedure is based more or less explicitly on the characterization of the guided part of the wave from some special solutions called modes of the waveguide. Each mode is expressed in terms of an eigenfunction of a related boundary-value problem set on a fixed waveguide cross-section S. When the governing equation for the propagation is the Helmholtz equation, the eigenfunction problem is relative to the transverse Laplacian in the interior of S endowed with the boundary condition involved in the complete domain. Truncating the infinite part of the waveguide beyond S leads straightforwardly to a boundary condition on S involving these modes. This expression is nothing else but an explicit writing of the Dirichlet-to-Neumann operator, which is made possible by a separation of variables. This is the so-called Steklov-Poincaré or Calderòn operator. It is well known that it is not generally easy to handle such an operator from a numerical standpoint. When dealing with Steklov-Poincaré operators, the difficulty stems from their so-called non-local character. Here, this non-local character comes from the fact that any function with a compact support in S is mapped in a function whose support is S. Another difficulty in the numerical approximation of the present boundary condition stems from the fact that "*the propagation in a waveguide can be multimodal and dispersive*" [29]. Hence, great care has to be paid to obtain effective numerical schemes to approximate the non-local boundary condition (cf. [29]). Several methods have already been proposed.

The most direct and natural approach consists in using a truncated expansion of the trace of the solution on section S in terms of previous eigenfunctions, the unknowns therein being the level of each mode, as in the mode-matching methods (cf. e.g. [31], [26]). Goldstein [14] and later Lenoir and Tounsi [23] have improved this approach. Once the continuous problem is discretized by a finite element method, the improvement gained by their method lies in the fact that only nodal values remain as degrees of freedom. However, both methods have two flaws. They require the determination of the eigenfunctions, and the non-local character of the boundary condition remains present in the numerical scheme: all degrees of freedom relative to S are coupled together.

In the context of free wave propagation in unbounded domains, Engquist and Majda [12] introduced a local boundary condition to accurately terminate the domain of computation. This condition approximates the theoretical non-local one for waves propagating in directions close to the normal to the boundary. A few years later, Higdon [20] generalized their approach by introducing a higher order boundary operator which can absorb waves propagating in a half space at almost all angles of incidence. Higdon's method was adapted later to numerical modeling of microstrips in [8]. The approximate boundary condition involves a differential operator of order N, the number of modes propagating within the waveguide. The idea of considering higher order boundary operators has already been suggested in [14] following the pioneered work of Bayliss and Turkel [2] dealing with the construction of absorbing boundary conditions for the two-dimensional wave equation in an exterior domain. Unfortunately, numerical approximation of these boundary conditions by a finite difference or finite element scheme is rather unpractical when N is greater than two. Recently an absorbing layer which perfectly matches any outgoing wave has been introduced by Berenger [5]. Particularly efficient in exterior domains, this method has been tested successfully for waveguides (cf. [27]). However, it seems that this approach fails when the waveguide is fed through prescribed incident modes, since the layer is designed only to absorb outgoing waves. Furthermore, it is not clear that such an approach can be adequately used in the frequency domain in conjunction with a finite element approximation, since it does not lead to a natural variational formulation. Finally, it was reported recently in [32] that the PML method is ineffective in absorbing evanescent waves.

The boundary condition which we propose uses a rational approximation of the Steklov-Poincaré operator, involving only the eigenvalues of the transverse Laplacian. Thus, eigenmodes are needed only if they are used to feed the waveguide or if reflection coefficients have to be calculated. This is an important feature of the present approach, since modes propagating with the same phase velocity need not be distinguished.

Although Padé approximations of the Steklov-Poincaré operator symbol have been considered for a long time [12], they have apparently not been used for waveguides yet. Previous derivations of absorbing boundary conditions use an approximation of this symbol either locally or globally, depending on the desired properties of the resulting scheme [12], [13], [20], [19]. In the present case, the approximation process is based rather on an interpolation of this symbol. In another context, a similar approach has been adopted by Bayliss and Turkel [2] to write out a hierarchy of boundary conditions which annihilates successive terms in the Wilcox expansion of any solution to the two-dimensional wave equation in an exterior domain.

The boundary condition that we propose is an extension of a Fourier-Robin condition valid for one propagating mode, which has been used for a long time in electrical engineering calculations [21]. Though non-local, we call it a quasi-local boundary condition because the resulting boundary-value system involves only local (that is, differential) operators when some adequate auxiliary unknowns join the formulation. This idea of considering auxiliary unknowns to deal only with differential operators seems to go back to Lindman (cf. [24]).

The quasi-local boundary condition is expressed in terms of partial derivatives of order no more than two. Accordingly, its effective numerical approximation can be performed through standard low order finite element schemes. It is perfectly transparent for the propagating modes, that is, it has the same effect on the propagating modes as the exact boundary condition. This property is stronger than low reflecting, because incident waves propagate through the boundary without perturbation. Although some auxiliary functions are introduced, a lumping process makes it possible to keep only the nodal values of the solution as unknowns in the discrete problem. The matrix of the resulting linear system remains sparse everywhere and can be obtained through a standard assembly process. Numerical experiments confirm that the method is capable of effectively solving various problems in waveguides. They show a very low-level reflection of the incident wave, comparable to the one found by using Berenger's perfectly matched layer in the solution of the problem by a finite-difference time-domain method [27].

A quasi-local boundary condition generating no reflection of the first N propagating modes can be written by adjusting N coefficients. Introducing an additional coefficient permits us to control the stability of the related boundary-value problem. As a result, we will be able to prove that the discrete problem is uniquely solvable and to give optimal error estimates as well. However, numerical experiments indicate that the quasi-local condition with only N coefficients, which results in a reduction of one auxiliary unknown function, is almost as accurate as the condition with $N + 1$ coefficients, a feature which will be studied elsewhere. Similarly, we found that the method still gives accurate results when the quasi-local boundary condition is designed for N propagating or evanescent modes. Numerical experiments show that taking non-propagating modes into account results in substantially reducing the amount of calculation, since only a small part of the waveguide is sufficient to get great accuracy. Hence, even in the most usual case where there is only one propagating mode, the quasi-local boundary condition can be advantageously used to improve the accuracy without a significant growth in the amount of calculation.

The outline of this paper is as follows. The mathematical background is introduced in section 2. A general boundary condition generating no reflection for propagating modes is studied in section 3, and exponential convergence to the exact solution is proved. Stability and regularity results are given in section 4. The construction of the quasi-local boundary condition is detailed in section 5. Next the convergence of the finite-element method is established. In a final section, we give an account of some numerical experiments.

2. General framework

In this section, a model boundary-value problem involving the Helmholtz equation in a perturbed semi-infinite waveguide $\Omega' \subset \mathbb{R}^d$ is stated. The boundary
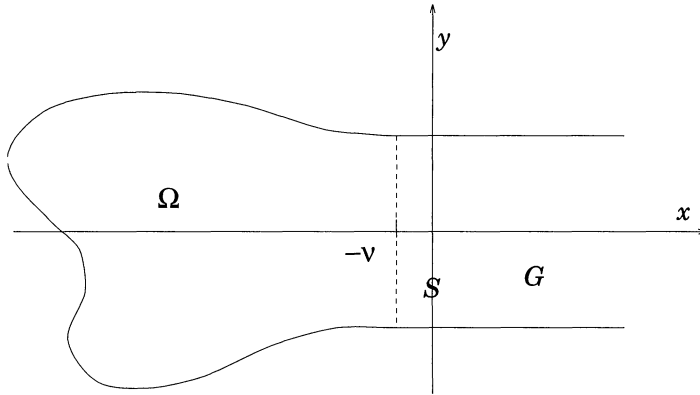
FIGURE 1. A semi-infinite waveguide

condition which we consider is either the Dirichlet or the Neumann boundary condition, since these cases provide the most representative situations for this class of problems. Junctions of several waveguides could be treated in the same way. More general (scalar) operators and boundary conditions could be treated by simply adapting the notation, provided that they rely upon a self-adjoint operator in $L^2(\Omega')$.

First, some notation is introduced. Then some known results on the exact non-reflecting condition are recalled.

2.1. **Notation.** The geometry of a semi-infinite perturbed waveguide (see figure 1) is described through the following data:
- a number $\nu > 0$,
- an open bounded domain S of $\mathbb{R}^{d-1}$, $d = 2, 3$, whose boundary ∂S is of class $\mathcal{C}^\infty$,
- the cylinder $G =]-\nu, +\infty[\times S$,
- an open bounded set Ω of $\mathbb{R}^d$ such that

$$\Omega \cap \{(x, y) \in \mathbb{R} \times \mathbb{R}^{d-1}; \ -\nu \leq x\} = [-\nu, 0[\times S.$$

The perturbed semi-finite waveguide then can be described through $\Omega' := \Omega \cup G$. Its closure $\overline{\Omega}'$ is assumed to be a $\mathcal{C}^\infty$-manifold imbedded in $\mathbb{R}^d$ whose boundary is denoted by Γ. **n** is the unit normal to Γ outwardly directed to Ω', and $\partial_\mathbf{n}$ is the related normal derivative to Γ.

A generic point in Ω' is designated by (x, y) with $x \in \mathbb{R}$ and $y \in \mathbb{R}^{d-1}$.

Standard notation and function spaces from the theory of partial differential equations are used without further comment (cf. e.g., [25], [30], [9]).

We denote by λ_n^2, $n \geq 1$, the eigenvalues of the transverse Laplacian relative to either the Dirichlet or the Neumann boundary condition on ∂S. The eigenvalues are counted with their order of multiplicity and increasingly ordered, $0 \leq \lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n \leq \ldots$, in such a way that each of them corresponds to an eigenfunction e_n:

$$(1) \qquad \begin{cases} -\Delta_S\, e_n = \lambda_n^2\, e_n & \text{in } S, \\ e_n = 0 \text{ or } \partial_\mathbf{n} e_n = 0 & \text{on } \partial S, \end{cases}$$

and $\{e_n\}_{n\geq 1}$ defines an orthonormal Hilbert basis of $L^2(S)$. We denote by $\Delta_S = \sum_{i=1}^{d-1} \partial_{y_i}^2$ the transverse Laplacian relative to a function defined on S.

Let $k > 0$ be the wavenumber involved in the problem. The constant of propagation k_n related to the nth mode is given by

$$k_n = \sqrt{k^2 - \lambda_n^2},$$

the determination of the square root being fixed by $\sqrt{-1} = i$. We assume that k does not correspond to a cut-off frequency, that is,

$$k_n \neq 0, \quad \text{for } n \geq 1.$$

The guided modes of the waveguide G then can be made explicit by

$$\varepsilon_n^{\pm}(x, y) = e_n(y)\, e^{\pm i k_n x}.$$

Waves ε_n^- represent incident or incoming waves whereas waves ε_n^+ are outgoing waves.

The number N of modes which propagate without attenuation is defined through the following relations:

$$k_n = \sqrt{k^2 - \lambda_n^2} > 0, \quad 1 \leq n \leq N,$$
$$i k_n = -\sqrt{\lambda_n^2 - k^2} < 0, \quad n > N.$$

We assume that the perturbed semi-bounded waveguide is loaded by a localized distribution of sources modeled by a given function f in $L^2(\Omega)$ as well as by an incident wave

$$u^i = \sum_{n=1}^{N} u_n^i\, \varepsilon_n^-.$$

The incident wave is completely defined by the respective levels u_n^i of the incident modes ε_n^-. Here, we have considered an incident wave to show how to deal with the problem in the case of a junction of waveguides.

2.2. Function spaces.

For $s \in \mathbb{R}$, the norm of the Sobolev space $H^s(\Omega)$ is denoted by $|\,.\,|_{s,\Omega}$. For any open region $\mathcal{O}$ of Ω', we denote by $V(\mathcal{O})$ either $\{u \in H^1(\mathcal{O}); \ u_{|\Gamma} = 0\}$ or $H^1(\mathcal{O})$, according to whether the Dirichlet or Neumann boundary condition is considered. The associate Fréchet space V_{loc} is

$$V_{\text{loc}}(\overline{\mathcal{O}}) = \{u \in L^2(\Omega'); \ \varphi\, u \in V(\mathcal{O}), \ \forall \varphi \in \mathcal{D}(\mathbb{R}^d)\}.$$

The Fourier coefficients u_n of a function $u \in L^2(S)$ are defined by

$$u_n = \int_S u(y)\, e_n(y)\, dy.$$

For any real number $s \geq 0$, the following Hilbert space X^s and its spectral norm

$$X^s = \{u \in L^2(S); \ \sum_{n\geq 1} |\lambda_n^s u_n|^2 < \infty\},$$
$$|u|_{X^s}^2 = \sum_{n\geq 1} |\lambda_n^s u_n|^2,$$

play a fundamental role. It follows from interpolation theory [25], [7] that for $0 \leq s < 1/2$ the space X^s coincides with the Sobolev space $H^s(S)$. For $1/2 \leq s \leq 1$ the identification of these spaces depends on the boundary condition being considered on Γ. For the Dirichlet boundary condition, $X^s = H_0^s(S)$, that is, the closure

of $\mathcal{D}(S)$ in $H^s(S)$ when $1/2 < s < 1$, and $X^{1/2} = H_{00}^{1/2}(S)$, a strict subspace of $H^{1/2}(S)$ whose elements vanish at the boundary of S in some special sense. The case of a Neumann boundary condition simply leads to $X^s = H^s(S)$ (cf. [25] and [16]).

The dual space of X^s, denoted by X^{-s}, can be identified with the one consisting of all sequences $\{u_n\}_{n\geq 1}$ with complex coefficients such that

$$|u|_{X^{-s}}^2 := \sum_{n\geq 1} |\lambda_n^{-s} u_n|^2 < \infty.$$

Using the characterization of the spectral spaces above, we can easily show that the usual trace operator $u(x, \cdot)$ is in addition a continuous map from $V(\Omega)$ into $X^{1/2}$ for $-\nu \leq x \leq 0$ and u in $H^1(\Omega)$. Moreover, as we will see below in a more general setting, this mapping is surjective and has a continuous right inverse.

2.3. **Exact bounded domain formulation.** Every solution $u \in V_{\text{loc}}(\overline{G})$ of the equation

$$-\Delta u - k^2 u = 0$$

satisfying the boundary condition considered on Γ has a unique decomposition

$$\begin{aligned}
u(x,y) &= u^+(x,y) + u^-(x,y), \\
u^+(x,y) &= \sum_{n\geq 1} u_n^+ \varepsilon_n^+(x,y), \\
u^-(x,y) &= \sum_{n\geq 1} u_n^- \varepsilon_n^-(x,y),
\end{aligned}$$

with

$$u_n = \int_S u(0,y)e_n(y)\,dy = u_n^+ + u_n^-.$$

General results on regularity up to the boundary of elliptic boundary-value problems (cf. e.g., [1], [25], etc.) show that this solution is $\mathcal{C}^\infty$ in $]-\nu, \infty[\times \overline{S}$. The outgoing part u^+ satisfies the boundary condition on S

$$\partial_{\mathbf{n}} u^+(0, \cdot) = i \sum_{n\geq 1} k_n u_n^+ e_n.$$

This suggests that we should consider the pseudodifferential operator of order one

$$\begin{aligned}
K : X^{1/2} &\longrightarrow X^{-1/2}, \\
\sum_{n\geq 1} u_n e_n &\longmapsto i \sum_{n\geq 1} k_n u_n e_n.
\end{aligned}$$

Furthermore, since $k_n \neq 0$ for $n \geq 1$, and $k_n \simeq i\lambda_n$, K is indeed an isomorphism from X^s onto X^{s-1} for all real numbers s.

Then the problem to be solved can be stated as follows:

$$(2) \qquad \begin{cases} -\Delta u - k^2 u = f & \text{in } \Omega, \\ u = 0 \text{ or } \partial_{\mathbf{n}} u = 0 & \text{on } \Gamma, \\ \partial_{\mathbf{n}} u - Ku = \partial_{\mathbf{n}} u^i - Ku^i & \text{on } S_0, \end{cases}$$

where f and u^i are respectively the given localized sources and incident wave.

It has been proved in [14] that this problem is well-posed except for an at most countable set Λ of values of k^2 with no finite accumulation point. This set Λ, which

may be empty, is the point spectrum of the operator acting in $L^2(\Omega')$ given by $-\Delta$ associated with the considered boundary condition on Γ [22]. We will always assume in this paper that $k^2 \notin \Lambda$ without further comment. Hence problem (2) is well-posed. Below, we will more precisely state the stability relative to the data.

3. THE APPROXIMATE BOUNDARY CONDITIONS

The boundary condition in problem (2) needs the computation of the eigenfunctions e_n to be made explicit. We introduce a new class of boundary conditions which need only the knowledge of some of the eigenvalues λ_n. These boundary conditions constitute an approximation of the exact one in that they leave each propagating mode unchanged. We give precise details about this approximation by estimating the error which affects the solution to the initial problem. Finally, we see that the approximate boundary conditions can take into account incident waves just as the exact one does.

Let R be a function defined for $t \geq 0$ satisfying

(3)
$$\begin{cases} t_n := R(\lambda_n^2) > 0 \quad \forall n \geq 1, \\ \lim\limits_{t \to +\infty} \frac{R(t)}{t} = a > 0. \end{cases}$$

The function R plays the role of a symbol. Exactly as for the above operator K, we can associate to the latter a pseudodifferential operator $T = R(-\Delta_S)$ of order 2 through

$$T : X^1 \longrightarrow X^{-1},$$
$$\sum_{n \geq 1} u_n e_n \longmapsto i \sum_{n \geq 1} t_n u_n e_n.$$

Again as for the operator K, T is indeed an isomorphism from X^s onto X^{s-2} for all s, since $t_n \neq 0$ for $n \geq 1$ and $t_n \simeq a\lambda_n^2$.

For $-\nu \leq L$, we denote by S_L the cross-section of the waveguide $\{(L, y); \ y \in S\}$, and for $0 \leq L$, we denote by Ω_L the open set $\Omega_L = \Omega \cup ([0, L[\times S)$. When it is sufficiently clear from the context, we do not distinguish between S_0 and S.

The problem with the approximate boundary condition can now be stated as follows: find u in $H^2(\Omega_L)$ such that

(4)
$$\begin{cases} -\Delta u - k^2 u = f & \text{in } \Omega_L, \\ u = 0 \text{ or } \partial_{\mathbf{n}} u = 0 & \text{on } \Gamma, \\ \partial_{\mathbf{n}} u - Tu = \partial_{\mathbf{n}} u^i - Tu^i & \text{on } S_L. \end{cases}$$

Whenever there is no risk of confusion, we do not distinguish between a function or a distribution defined in S_L and its identification, when this is possible, to an element of X^s.

Remark. If the symbol R is chosen such that $R(\lambda_n^2) = k_n$ for $1 \leq n \leq N$, then the boundary condition of (4) is transparent for all propagating modes, since $\partial_{\mathbf{n}} u - Tu = \partial_{\mathbf{n}} u - Ku$ for any solution u in the form $u = \sum_{n=1}^{N} u_n^+ \varepsilon_n^+ + u^i$.

Let $\mathcal{V}_L$ be the space of $v \in V(\Omega_L)$ such that $v_{|S_L} \in X^1$ equipped with the norm

$$|v|_{\mathcal{V}_L} = \left(|v|_{H^1(\Omega_L)}^2 + |v|_{X^1}^2 \right)^{1/2}.$$

Problem (4) has the following variational formulation: find $u \in \mathcal{V}_L$ such that

$$a_L(u, v) = l_L(v), \qquad \forall v \in \mathcal{V}_L,$$

with

$$a_L(u,v) = \int_{\Omega_L} \left(\nabla u . \nabla v - k^2 u \, v \right) \, dx \, dy - \langle Tu, v \rangle_{-1,1},$$

$$l_L(v) = \int_{\Omega_L} f \, v \, dx \, dy + \int_{S_L} (\partial_{\mathbf{n}} u^i - Tu^i) \, v \, dy,$$

where the brackets $\langle \cdot, \cdot \rangle_{-s,s}$ denote the duality product between X^{-s} and X^s.

The following theorem gives the general framework which permits us to analyze the stability of the problem relative to the boundary conditions which are introduced, and to show that the error resulting from these approximate conditions decays exponentially as $L \to +\infty$.

Theorem 3.1. *Problem* (4) *has one and only one solution. Let* u_L *be the solution to problem* (4) *and* u_E *be that of problem* (2). *Assume that* $R(\lambda_n^2) = k_n$ *for* $1 \leq n \leq N$; *then there exist two positive constants* c *and* ρ, *independent of the data* f *and* u^i *and of* L, *such that*

$$|u_E - u_L|_{1,\Omega} \leq ce^{-\rho L} \left(|f|_{0,\Omega} + |u^i|_{X^{1/2}} \right).$$

Proof. Since $t_n > 0$ for $n \geq 1$, and $t_n \simeq a\lambda_n^2$, we readily get the existence of a positive constant c such that

$$-i \langle Tu, \overline{u} \rangle_{-1,1} = \sum_{n \geq 1} t_n |u_n|^2 \geq c \sum_{n \geq 1} \lambda_n^2 |u_n|^2 = c|u|_{X^1}^2$$

and, as a consequence,

$$\left| a_L(u, \overline{u}) + (1 + k^2) \int_{\Omega_L} |u|^2 \, dx \, dy \right| \geq c|u|_{\mathcal{V}_L}^2$$

for each u in $\mathcal{V}_L$. Well-posedness of problem (4) is thus a consequence of the Fredholm alternative. Let u be such a solution with $f = 0$ and $u^i = 0$. Since the imaginary part $\text{Im}(a_L(u, \overline{u})) = -\langle Tu, \overline{u} \rangle_{-1,1} = 0$, the traces $u_{|S_L}$ and $\partial_{\mathbf{n}} u$ are zero on S_L. Extending u by 0 beyond S_L, we readily obtain that the extension is still a solution to the Helmholtz equation in Ω'. Since the function obtained is analytic therein, it must vanish.

To establish the error estimate, we expand the solution in G as

$$(5) \qquad u_L - u^i = \sum_{n \geq 1} (u_{Ln}^+ e^{ik_n x} + \left(u_{Ln}^- - u_n^i \right) e^{-ik_n x}) e_n(y)$$

with $u_n^i = 0$ for $n > N$. Our aim is to compare the restriction $(u_L)_{|\Omega}$ with the solution u to problem (2). The details being similar to those given in [18], we limit ourselves here only to the main steps in the proof.

Taking the derivative relatively to x in (5) and using the boundary condition of problem (4), we get the following expressions for the normal derivative on S_L:

$$\partial_{\mathbf{n}}(u_L - u^i) = \sum_{n \geq 1} ik_n(u_{Ln}^+ e^{ik_n L} - \left(u_{Ln}^- - u_n^i \right) e^{-ik_n L}) e_n(y),$$

$$= \sum_{n \geq 1} it_n(u_{Ln}^+ e^{ik_n L} + \left(u_{Ln}^- - u_n^i \right) e^{-ik_n L}) e_n(y).$$

Hence,

$$(k_n - t_n)u_{Ln}^+ e^{ik_n L} = (k_n + t_n) \left(u_{Ln}^- - u_n^i \right) e^{-ik_n L}, \quad \forall n \geq 1.$$

Defining

$$d_n = \frac{k_n - t_n}{k_n + t_n} \quad \text{and} \quad r_n = k_n \frac{1 - d_n \mathrm{e}^{2ik_n L}}{1 + d_n \mathrm{e}^{2ik_n L}},$$

we obtain the following relation on the cross-section S_0 (i.e. for $L = 0$):

$$\partial_x(u_L - u^i) = \sum_{n \geq 1} i r_n (u_{Ln}^+ + (u_{Ln}^- - u_n^i)) e_n(y).$$

Hence, the restriction $(u_L)_{|\Omega}$ of u_L to the open set Ω, also denoted by u_L for simplicity, satisfies

$$\begin{cases} -\Delta u_L - k^2 u_L = f & \text{in } \Omega, \\ u_L = 0 \text{ or } \partial_\mathbf{n} u_L = 0 & \text{on } \Gamma, \\ \partial_\mathbf{n} u_L - T_L u_L = \partial_\mathbf{n} u^i - T_L u^i & \text{on } S_0, \end{cases}$$

where T_L is the continuous operator

$$T_L : X^{1/2} \longrightarrow X^{-1/2},$$
$$\sum_{n \geq 1} u_n e_n \longmapsto i \sum_{n \geq 1} r_n u_n e_n.$$

The key point in the proof is the following. Since $R(\lambda_n^2) = k_n$ for $1 \leq n \leq N$, we have $|d_n| = 0$ for $1 \leq n \leq N$ and $|d_n| = 1$ for $n > N$. Hence, $r_n = k_n$ for $1 \leq n \leq N$ and

$$(K - T_L)u = i \sum_{n > N} k_n \left(\frac{2 d_n \mathrm{e}^{2ik_n L}}{1 + d_n \mathrm{e}^{2ik_n L}} u_n e_n \right),$$

which yields

$$|K - T_L|_{\mathcal{L}(X^{1/2},\, X^{-1/2})} \leq c \frac{2 \mathrm{e}^{2ik_{N+1} L}}{1 - \mathrm{e}^{2ik_{N+1} L}},$$

c being a constant independent of N and L.

As usual in partial differential equation theory, c will stand for any constant.

Since problem (2) is well posed and invertible elements of a Banach algebra constitute an open set, the desired estimate holds with $\rho = -2ik_{N+1} > 0$. $\qquad\square$

4. STABILITY AND REGULARITY ESTIMATES

In this section, we state the stability and regularity results which will be used in the next one.

It will be convenient to suppose here that the solution to problem (4) is sufficiently close to the solution to problem (2) for $L = 0$. This holds for a sufficiently large ν. Numerical computations show that a distance of one wavelength is enough.

4.1. Stability.
We will use the following technical lemma in several instances. Its proof is straightforward from standard techniques of interpolation theory and square integrable functions valued in Hilbert spaces (cf. [25] and [16]), and hence is omitted.

Lemma 4.1. *For $m = 1$ or $m = 2$, the usual trace operator given for sufficiently smooth functions v by $v_{|S}$ is bounded from $H^m(\Omega) \cap V(\Omega)$ into $X^{m-1/2}$ and has a continuous right inverse explicitly given for $v = \sum_{n \geq 1} v_n e_n \in X^{m-1/2}$ by*

$$v = \varphi \sum_{n \geq 1} v_n \varepsilon_n^-,$$

where φ is a fixed element of $C^{\infty}(\overline{\Omega})$ such that $\varphi = 1$ near S and supp $\varphi \subset\ ]-\nu, 0]\times S$. *Moreover, this explicit mapping is also continuous from $X^{m-1/2}$ into $H^m(\Omega)$ for all positive integers m. Finally, the trace given for sufficiently smooth functions v by $(\partial_{\mathbf{n}} v)_{|S} = (\partial_x v)_{|S}$ defines a bounded mapping from $H^2(\Omega) \cap V(\Omega)$ into $X^{1/2}$.*

Remark. For $m > 2$, the trace $v_{|S}$ of $v \in H^m(\Omega) \cap V(\Omega)$ is not necessarily in $X^{m-1/2}$. Some additional boundary conditions must be fulfilled by v, e.g. $\Delta_S v = 0$ on ∂S.

To establish the stability of the finite element scheme, we need the following regularity result for the solution to the problem with the exact boundary condition, but with data f whose support extends up to S and data g in $X^{1/2}$.

Lemma 4.2. *Let $f \in L^2(\Omega)$ and $g \in X^{1/2}$ be given. The problem*

$$(6) \qquad \begin{cases} -\Delta u - k^2 u = f & in\ \Omega, \\ u = 0\ or\ \partial_{\mathbf{n}} u = 0 & on\ \Gamma, \\ \partial_{\mathbf{n}} u - Ku = g & on\ S, \end{cases}$$

has one and only one solution belonging to $H^2(\Omega)$. Moreover, there exists a constant c independent of f and g such that

$$|u|_{2,\Omega} \le c\left(|f|_{0,\Omega} + |g|_{X^{1/2}}\right).$$

Proof. Using notation similar to that of lemma 4.1, we set

$$v = -\varphi \sum_{n \ge 1} \frac{g_n}{2ik_n} \varepsilon_n^-.$$

Since $k_n \simeq i\lambda_n$, this lemma yields $v \in H^2(\Omega)$ with

$$(7) \qquad |v|_{2,\Omega} \le c\,|g|_{X^{1/2}}.$$

Observing that $\partial_{\mathbf{n}} v - Kv = g$ and replacing u by $u - v$, we prove the lemma for $g = 0$.

The Fredholm alternative insures that problem (6) is uniquely solvable in $V(\Omega)$.

Extending f by zero and u by $\sum_{n \ge 1} \left(\int_S u(0, y)\, e_n(y)\, dy\right) \varepsilon_n^+$ beyond S, we find that the extension satisfies a homogeneous elliptic boundary-value problem in Ω'. The end of the proof follows from standard estimates up to the boundary for regular elliptic boundary-value problems (cf. e.g., [1], [25], etc.).   □

Finally, we establish a similar regularity result for the solution to the problem related to the approximate boundary condition.

Proposition 4.3. *As in lemma 4.2, let $f \in L^2(\Omega)$ and $g \in X^{1/2}$ be given. The solution $u \in \mathcal{V}_0$ of the problem*

$$(8) \qquad \begin{cases} -\Delta u - k^2 u = f & in\ \Omega, \\ u = 0\ or\ \partial_{\mathbf{n}} u = 0 & on\ \Gamma, \\ \partial_{\mathbf{n}} u - Tu = g & on\ S, \end{cases}$$

is in $H^2(\Omega)$ with $u_{|S} \in X^{5/2}$ and satisfies the estimate

$$|u|_{2,\Omega} + |u|_{X^{5/2}} \le c\left(|f|_{0,\Omega} + |g|_{X^{1/2}}\right)$$

with a constant c independent of f and g.

Proof. Existence and uniqueness of a solution in $\mathcal{V}_0$ have already been established in theorem 3.1. Let u_1 be the solution to (6). From lemma 4.2, u_1 belongs to $H^2(\Omega)$ and thus its trace on S is in $X^{3/2}$.

The function $v = u_1 - u$ in turn is a solution to problem (8) with $f = 0$ and $g = (K - T)u_1$, denoted by h for clarity. Note that h is now in $X^{-1/2}$ and satisfies $|h|_{X^{-1/2}} \leq c|u_1|_{X^{3/2}} \leq c|u_1|_{2,\Omega}$. Since T is an isomorphism from X^s onto X^{s-2} for each $s \in \mathbb{R}$ and $Tv = \partial_{\mathbf{n}} v - h$, it follows that

$$|v|_{X^{3/2}} \leq c(|\partial_{\mathbf{n}} v|_{X^{-1/2}} + |h|_{X^{-1/2}}).$$

From $\Delta v + k^2 v = 0$ and the definition of traces by transposition methods, we get the bound $|\partial_{\mathbf{n}} v|_{X^{-1/2}} \leq c|v|_{1,\Omega}$. Since problem (8) is well-posed, we get

$$(9) \qquad |v|_{1,\Omega} \leq c|h|_{X^{-1}} \leq c|h|_{X^{-1/2}} \leq c|u_1|_{2,\Omega}.$$

Gathering these inequalities and using lemma 4.1, we arrive at

$$(10) \qquad |v|_{X^{3/2}} \leq c|u_1|_{2,\Omega}.$$

The function $v \in V(\Omega)$ is a solution to

$$\begin{cases} \Delta v + k^2 v = 0 & \text{in } \Omega, \\ v = 0 \text{ or } \partial_{\mathbf{n}} v = 0 & \text{on } \Gamma, \\ v_{|S} \in X^{3/2}. \end{cases}$$

Using lemma 4.1, we can reduce the problem as above to the one with a homogeneous Dirichlet condition on S and a right-hand side of the partial differential equation in $L^2(\Omega)$. The geometry of the cross-section of the waveguide permits us to extend the solution to an odd function relative to the variable x. Standard elliptic estimates up to the boundary then give the bound

$$|u|_{2,\Omega} \leq c\left(|f|_{0,\Omega} + |g|_{X^{1/2}}\right).$$

Furthermore, since T is a pseudodifferential operator of order 2, there is an extra regularity for the trace $u_{|S}$ which results from the inequality

$$|\partial_{\mathbf{n}} u - g|_{X^{1/2}}^2 = |Tu|_{X^{1/2}}^2 = \sum_{n \geq 1} \lambda_n t_n^2 |u_n|^2 \geq \gamma \sum_{n \geq 1} \lambda_n^5 |u_n|^2$$

with a constant $\gamma > 0$ independent of f and g. The end of the proof is then a direct consequence of this last estimate. □

4.2. Regularity of the solution.

We will see now that the curvilinear edge resulting from the truncation of the waveguide does not generate any flaw in the regularity properties of the solution to the problem related to the approximate boundary condition as long as the support of the localized sources f does not extend up to the cross-section S. It is clear that this assumption is not restrictive for the class of problems which are considered.

Proposition 4.4. *Assume that $R(\lambda_n^2) = k_n$ for $1 \leq n \leq N$. Let $u \in \mathcal{V}_0$ be a solution to problem (8). For any integer $m \geq 0$, if $f \in H^m(\Omega)$ is such that $\operatorname{supp} f \subset \left\{(x,y) \in \mathbb{R}^d;\ x \leq -\nu\right\}$ and $g \in X^{m-1/2}$, then $u \in H^{m+2}(\Omega)$ and the estimate*

$$(11) \qquad |u|_{m+2,\Omega} \leq c\left(|f|_{m,\Omega} + |g|_{X^{m-1/2}}\right)$$

holds with a constant c independent of f and of g. Moreover, regardless of the regularity of f, $u_{|S}$ is in $X^{m+3/2}$ as soon as g is in $X^{m-1/2}$, and we have the bound

$$|u|_{X^{m+3/2}} \leq c \left(|f|_{0,\Omega} + |g|_{X^{m-1/2}}\right),$$

again with a constant c independent of f and g.

Proof. Let $\varphi \in \mathcal{C}^\infty(\overline{\Omega})$ be such that $\varphi(x,y) = 1$ on $]-\nu/2, 0[\times S$ and $\operatorname{supp}\varphi \subset [-\nu, 0] \times S$. Set

$$v = -\varphi \sum_{n \geq 1} \frac{g_n}{i(k_n + t_n)} \varepsilon_n^-$$

as in lemma 4.1. The function v satisfies the boundary condition $\partial_{\mathbf{n}} v - Tv = g$. Since $k_n + t_n \simeq a \lambda_n^2$, we immediately get

(12) $$|v|_{X^{m+3/2}} \leq c \, |g|_{X^{m-1/2}}$$

with a constant c independent of g.

It follows from lemma 4.1 that $v \in H^{m+2}(\Omega)$ with the bound

(13) $$|v|_{m+2,\Omega} \leq c \, |g|_{X^{m-1/2}},$$

again with a constant c independent of g.

Now, consider the function $w := u - v$. It is the solution to

(14)
$$\begin{cases} -\Delta w - k^2 w = f_1 & \text{in } \Omega, \\ w = 0 \text{ or } \partial_{\mathbf{n}} w = 0 & \text{on } \Gamma, \\ \partial_{\mathbf{n}} w - Tw = 0 & \text{on } S, \end{cases}$$

where $f_1 = f + \Delta v + k^2 v \in H^m(\Omega)$. Moreover, according to the definition of φ and v, we have $\operatorname{supp} f_1 \subset \left\{ (x,y) \in \mathbb{R}^d;\ x \leq -\nu/2 \right\}$. Hence we can decompose w for $-\nu/2 < x < 0$ as

(15) $$w = w^+ + w^- = \sum_{n \geq 1} w_n^+ \varepsilon_n^+ + w_n^- \varepsilon_n^-.$$

Thus the boundary condition satisfied by w on S can be rewritten as

$$\partial_{\mathbf{n}} w - Kw = \partial_{\mathbf{n}} w^- - Kw^-.$$

Goldstein [14] has established that the bound

(16) $$|w|_{m+2,\Omega} \leq c \left(|f_1|_{m,\Omega} + |w^-|_{X^{m+3/2}}\right)$$

holds with a constant c independent of f_1 and $w_{|S}^-$. It remains to give an estimate of $|w^-|_{X^{m+3/2}}$.

Taking the normal derivative in (15) and using the boundary condition in (14), we obtain the relation

$$\partial_{\mathbf{n}} w = \sum_{n \geq 1} i k_n (w_n^+ - w_n^-) e_n = \sum_{n \geq 1} i t_n (w_n^+ + w_n^-) e_n,$$

which in turn yields

(17) $$w_n^- = d_n w_n^+ \quad \text{for } n \geq 1,$$

with $d_n := (k_n - t_n)/(k_n + t_n)$. Note that $d_n = 0$ for $1 \leq n \leq N$, since then $t_n = R(\lambda_n^2) = k_n$, and that $|d_n| = 1$ for $n > N$ because then both t_n and ik_n are real.

The function w is of class $\mathcal{C}^\infty$ on $]-\nu/2,\, 0[\, \times \overline{S}$, and for each fixed x in $]-\nu/2,\, 0[$ we can write

$$w(x,y) = \sum_{n\geq 1} w_n(x) e_n(y).$$

Using (15) and (17), we obtain for $n > N$

(18) $\qquad\qquad w_n(x) \;=\; w_n^+ \, e^{ik_n x} + w_n^- \, e^{-ik_n x}$

(19) $\qquad\qquad\qquad =\; d_n^{-1} \, e^{ik_n x}(1 + d_n \, e^{-2ik_n x}) w_n^-.$

Hence, we can get the following bound:

$$|e^{ik_n x} w_n^-| \leq \frac{1}{1 - e^{-2ik_{N+1}x}} |w_n(x)| \qquad \text{for } n \geq 1.$$

The sequence $ik_n x$ is positive and increasing for $n > N$ with $k_n \simeq i\lambda_n$. For each $s > 0$, we have $\lim_{n\to\infty} \lambda_n^{2s} e^{-2ik_n x} = 0$. Thus, using (19) and the fact that trace operator is bounded from $H^1(\Omega)$ into $L^2(S_x)$, we can write

$$
\begin{aligned}
|w^-|_{S_0}|^2_{X^s} \;&=\; \sum_{n\geq 1} \frac{\lambda_n^{2s}}{e^{2ik_n x}} |e^{ik_n x} w_n^-|^2 \\
&\leq\; c(x,s,N) \sum_{n\geq 1} |w_n(x)|^2 = c(x,s,N)|w_{|S_x}|^2_{X^0} \\
&\leq\; c(x,s,N)|w|^2_{1,\Omega}.
\end{aligned}
$$

Applying theorem 3.1 to problem (14), we get the estimate $|w|_{1,\Omega} + |w_{|S}|_{X^1} \leq c|f_1|_{0,\Omega}$, which combined with (12) and in view of the expression of f_1 yields

(20) $\quad |w^-|_S|^2_{X^s} \leq c|f_1|_{0,\Omega} \leq c(|f|_{0,\Omega} + |v|_{2,\Omega}) \leq c(|f|_{0,\Omega} + |g|_{X^{-1/2}}).$

Substituting (20) in (16) and taking $s = m + 3/2$, we obtain

$$|u|_{m+2,\Omega} \leq c(|f|_{m,\Omega} + |g|_{X^{m-1/2}}).$$

The second part of the proposition has already been obtained through the previous steps (12), (15), (17) and (20). $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Remark. It is well-known (cf. e.g., [15]) that the data relative to boundary conditions of different type must fulfill certain compatibility conditions for the related solution to an elliptic boundary-value problem in order to be of optimal regularity. Actually, these compatibility conditions are here implicit from the fact that the data g is taken in $X^{m-1/2}$.

5. THE QUASI-LOCAL BOUNDARY CONDITION

Now, we come to the main goal of this paper. We show how appropriate choices of the function R permit us to overcome the difficulty of solving problem (2) with a non-local boundary condition on S. Actually, the resulting boundary condition remains non-local, but the addition of some unknown functions on S leads to a problem whose formulation involves only local (that is, differential) operators. When only one unknown function is involved, we will see that a lumping technique permits us to eliminate the additional unknown at the assembly process.

It is worth noting that the system to be solved depends on the wave number k through a multiplicative term k^2 and through the coefficients of a small system described below. As a result, the computation of higher order derivatives of the solution relative to k as described in [18] can be easily implemented.

5.1. The effective finite-element method. For convenience, we now change the numbering of the eigenvalues λ_n, counting each element of the set $\{\lambda_n; \ n \geq 1\}$ only once, i.e., $0 \leq \lambda_1 < \lambda_2 < \cdots < \lambda_n < \cdots$. As a result, the construction of the approximate boundary condition has the advantage of not requiring separation of modes which propagate with the same phase velocity, nor even determining them.

The rational function R

$$R(t) = at + b + \sum_{j=1}^{N-1} \frac{c_j}{j + t}$$

depends on $N + 1$ real parameters. It follows from classical Hermite interpolation theory that the conditions

(21)
$$\begin{cases} R(\lambda_n^2) = k_n & \text{for } 1 \leq n \leq N, \\ R'(\lambda_N^2) = 0 \end{cases}$$

uniquely determine R. The choice of the poles is arbitrary. The above analysis has shown that the first N equations are necessary for the consistency of the resulting boundary-value problem with the problem to be solved. The equation $R'(\lambda_N^2) = 0$ is added to carry out the requirement that $R(t)$ be positive for $t > \lambda_N^2$. For $N = 2$ and $N = 3$ some simple manipulations show that $a > 0$ and $R(t) > 0$ for all $t > \lambda_N^2$. We have not been able to prove that these last properties are still true for the next values of N. However, we have numerically determined the function R for $N = 4, 5$ in many cases, the values of λ_n being randomly chosen. We have never observed that either $a > 0$ or $R(t) > 0$ for $t > \lambda_N^2$ fails to be true. So we conjecture that these inequalities are valid for each N and take them as assumptions in what follows.

For $v \in X^1$, we can write Tv implicitly with only local operators as

$$Tv = i \sum_{n \geq 1} R(\lambda_n^2) v_n e_n = iR(-\Delta_S)v = i(-a\Delta_S v + bv + \sum_{j=1}^{N-1} c_j g_j),$$

$$-\Delta_S g_j + j g_j = v \quad \text{in } S$$

$$g_j = 0 \text{ or } \partial_{\mathbf{n}} g_j = 0 \quad \text{on } \partial S,$$

through $N - 1$ auxiliary functions g_j, $1 \leq j \leq N - 1$. Hence for solving problem (2), we are led to solve the problem

(22)
$$\begin{cases} -\Delta u - k^2 u = 0 & \text{in } \Omega, \\ u = 0 \text{ or } \partial_{\mathbf{n}} u = 0 & \text{on } \Gamma, \\ \partial_{\mathbf{n}} u - i(-a\Delta_S u + bu + \sum_{j=1}^{N-1} c_j g_j) \\ \qquad = \partial_{\mathbf{n}} u^i - i(-a\Delta_S u^i + bu^i) & \text{on } S, \\ -\Delta_S g_j + j g_j - u = -u^i & \text{in } S, \\ g_j = 0 \text{ or } \partial_{\mathbf{n}} g_j = 0 & \text{on } \partial S. \end{cases}$$

The following result is a straightforward consequence of proposition 4.4.

Proposition 5.1. *Problem* (22) *is uniquely solvable, and* u *is the solution to problem* (2) *if and only if* $(u, g_1, \cdots, g_{N-1})$ *is a solution to problem* (22). *Moreover, if* $f \in L^2(\Omega)$ *with* $\operatorname{supp} f \subset \{(x, y) \in \mathbb{R}^d; \ x \leq -\nu\}$, *then the functions* g_j *are of class* $\mathcal{C}^\infty$ *on* $\overline{S}$.

For solving the discrete problem, we use the following variational formulation. Let $\mathcal{V} = \mathcal{V}_0 \times \{X^1\}^{N-1}$ and denote by $y := (u, g)$ a generic element of $\mathcal{V}$ with

$g = (g_1, \ldots, g_{N-1})$. The space $\mathcal{V}$ is equipped with the norm

$$|y|_\mathcal{V} = \left(|u|^2_{\mathcal{V}_0} + \sum_{j=1}^{N-1} |g_j|^2_{X^1} \right)^{1/2}.$$

Then, $y \in \mathcal{V}$ is a solution to problem (22) if and only if

$$a(y, z) = l(z), \qquad \forall z = (v, s) \in \mathcal{V},$$

where

$$
\begin{aligned}
a(y, z) &= \int_\Omega \left(\nabla u . \nabla v - k^2 u\, v \right) dx dy - i \int_S a \nabla_S u . \nabla_S v + buv + \sum_{j=1}^{N-1} c_j g_j v \; dy \\
&+ \int_S \sum_{j=1}^{N-1} \nabla_S g_j . \nabla_S s_J + j g_j s_J - u s_J \; dy,
\end{aligned}
$$

$$
l(z) = \int_\Omega f\, v \; dx dy + \int_S \partial_\mathbf{n} u^i\, v - i(a \nabla_S u^i . \nabla_S v + bu^i\, v) - \sum_{j=1}^{N-1} u^i s_J \; dy.
$$

Let $\mathcal{T}_h$ be a mesh of Ω in the general meaning of usual approximations by a finite element method [10]. The trace of this mesh on S gives a mesh of the $(d-1)$-domain. Similarly, when designing a conforming nodal finite-element method of order m to approximate $H^1(\Omega)$, we define a similar method giving an approximation of the space $H^1(S)$. Taking into account the Dirichlet condition when it is involved, we can define $\mathcal{V}_h \subset \mathcal{V}$, a finite-element space of order m, and consider the following approximation of problem (22): find $y_h = (u_h, g_h) \in \mathcal{V}_h$ such that

$$(23) \qquad\qquad a(y_h, z_h) = l(z_h) \qquad \forall z_h = (v_h, s_h) \in \mathcal{V}_h.$$

As usual for simplicity, we do not take into account the consistency error coming from the approximation of the geometry. However, this aspect does not lead to any specific difficulty and could be treated by the usual techniques (see e.g., [4] for a similar situation) and Bernardi's general results [6].

The other result of this paper establishes that the discrete problem (23) is uniquely solvable if h is taken sufficiently small, and gives an error estimate for the solution which is effectively calculated.

Theorem 5.2. *Let y be the solution to problem (22). There exists $h^* > 0$ such that for $0 < h \le h^*$, problem (23) has a unique solution $y_h \in \mathcal{V}_h$ satisfying*

$$|y - y_h|_\mathcal{V} \le c\, h^m \left(|f|_{m-1,\Omega} + |u^i|_{X^{m+3/2}} \right)$$

with a constant c independent of f and u^i.

Proof. General results on the numerical analysis of conforming approximations of variational problems by a finite-element method (cf., e.g., [10], [28], and for instance [4] for the case of the approximation of a Fredholm alternative), and the above propositions 4.3 and 4.4 reduce the proof to checking that any solution $y = (u, g) \in \mathcal{V}$ to the coercive problem

$$
\begin{cases}
-\Delta u + u = f & \text{in } \Omega, \\
u = 0 \text{ or } \partial_\mathbf{n} u = 0 & \text{on } \Gamma, \\
\partial_\mathbf{n} u + ia\Delta_S u = f_0 & \text{on } S, \\
-\Delta g_J = f_J & \text{on } S,
\end{cases}
$$

satisfies the estimates

$$|u|_{2,\Omega} + |u|_{X^2} \;\leq\; c\,(|f|_{0,\Omega} + |f_0|_{X^{1/2}}),$$
$$\sum_{j=1}^{N-1} |g_j|_{X^2} \;\leq\; c\sum_{j=1}^{N-1} |f_j|_{X^0},$$

with a constant c independent of the data $(f, \{f_j\}_{j=0}^{N-1})$. Proposition 4.3 gives the first estimate by taking $Tu := -ia\Delta_S u$ and substituting -1 for k^2. Observe that T in proposition 4.3 was independent of k. Estimates up to the boundary for the solution to regular elliptic boundary-value problems complete the proof.   □

Abandoning now the assumption $L = 0$, which is simply a convenient notation, we can summarize the results of this paper as follows.

Theorem 5.3. *Let u_E be the solution to problem (2), u_L the solution to problem (4), and $u_{h,L}$ the finite-element solution for an approximation of order $m \geq 1$. There exist four positive constants h^*, c, ρ and c_L, independent of f in $H^{m-1}(\Omega)$, of u^i and of $0 < h < h^*$, such that*

$$|u_E - u_h|_{1,\Omega} \leq (c\,e^{-\rho L} + c_L h^m)(|f|_{m-1,\Omega} + |u^i|_{X^{m+3/2}})$$

Remark. The way that the constant c_L depends on L seems to be a difficult question. However, in view of the usual properties of coercive problems and of the fact that the usual approximations of the Helmholtz equation give sufficient accuracy with a fixed number of degrees of freedom by wavelength, it is reasonable to assume that c_L grows at most like L. The previous error estimate suggests taking the mesh length h of order $\exp(-\rho L/m)$. Numerical results presented in table 1 (see Section 5.3) confirm this observation.

5.2. **Implementation of the finite-element method.** For simplicity, we suppose in this section that only two modes are propagating, i.e., in the previous notation $N = 2$. We show that the additional unknown can be eliminated through a lumping technique which preserves the sparsity of the matrix and can be implemented at the element level. The same reduction can be applied when considering more propagating modes, but then with a less obvious advantage.

Without any change in the solution, we slightly modify the variational formulation of the discrete problem in such a way that the matrix of the final linear system to be solved will be symmetric although not Hermitian. The finite element unknown $y_h = (u_h, g_h) \in \mathcal{V}_h$ then reads

$$(24) \qquad a(y_h, z_h) = l(z_h), \qquad \forall z_h = (v_h, s_h) \in \mathcal{V}_h,$$

with the following notation:

$$a(y, z) \;=\; b(u, v) + d(v, g) + d(u, s) + c(g, s),$$
$$b(u, v) \;=\; \int_\Omega \left(\nabla u.\nabla v - k^2 u\,v\right)\, dx\,dy - i\int_S (a\nabla_S u.\nabla_S v + buv)\, dy,$$
$$d(v, g) \;=\; -ic\int_S g\,v\; dy,$$
$$c(g, s) \;=\; ic\int_S \left(\nabla_S g.\nabla_S s + gs\right)\, dy,$$
$$l(z) \;=\; \int_\Omega f\,v\; dx\,dy + \int_S \left(\partial_\mathbf{n} u^i\,v - i(a\nabla_S u^i.\nabla_S v + bu^i\,v) - ic\,u^i\,s\right)\, dy.$$

Let $\{e_i\}_{i \in I}$ be the nodal basis of the finite element-space $\mathcal{V}_0^h$ [10]. Using the obvious notation, we decompose the set of indices I as $I = I_\Omega \cup I_S$. For $i \in I_S$, we denote by $s_i = (e_i)_{|S}$ the trace of e_i on S. Standard properties of a nodal finite element method show that $\{s_i\}_{i \in I_S}$ is a nodal finite element approximation X_h^1 of X^1 with the same order of accuracy as $\mathcal{V}_0^h$ is of $\mathcal{V}_0$. As a result, the functions

$$\begin{aligned} E_i &= (e_i, 0), & i \in I, \\ S_i &= (0, s_i), & i \in I_S, \end{aligned}$$

constitute a basis of $\mathcal{V}_h = \mathcal{V}_0^h \times X_h^1$.

Any element of $\mathcal{V}_h$ can be written

$$(u_h, g_h) = \sum_{j \in I} u_j E_j + \sum_{j \in I_S} g_j S_j.$$

Since we are only interested in the nodal values u_j, we consider another basis of $\mathcal{V}_h$ which avoids the computation of the coefficients g_j.

By the standard lumping process, we express the bilinear form d in an approximate way by a diagonal matrix as follows: $d(e_j, s_i) = \delta_{ij} d_i$, where δ_{ij} is the Kronecker symbol. Let $I_S' = \{i \in I_S; \ d_i \neq 0\}$, $I_\Omega' = I \setminus I_S'$, and $c_{ij} = c(s_j, s_i)$. Define $S_i' = S_i$ for $i \in I_S$, $E_i' = E_i$ for $i \in I_\Omega'$ and

(25)
$$E_i' = \sum_{k \in I_S'} \frac{c_{ik}}{d_k} E_k - S_i, \qquad \text{for } i \in I_S'.$$

Observe that $c_{ik}/d_k = 0$ whenever nodes of respective index i and k do not belong to the same element.

Rewriting (u_h, g_h) in this basis as

$$(u_h, g_h) = \sum_{j \in I} u_j' E_j' + \sum_{j \in I_S} g_j' S_j',$$

and using the relations

$$a(S_j', E_i') = 0, \qquad \text{for } (i, j) \in I \times I_S,$$

which are a consequence of the lumping technique, we get the reduced linear system

(26)
$$\sum_{j \in I} a(E_j', E_i') \, u_j' = a((u_h, g_h), E_i') = l(E_i'), \qquad \text{for } i \in I,$$

where the unknowns g_j' have been eliminated.

The coefficients of the matrix of system (26) are given by

$$\begin{aligned} a_{ij} &= a((e_j, 0), (e_i, 0)) & \text{for } (i, j) \in I_\Omega' \times I_\Omega', \\ a_{ij} &= a((e_j, 0), (\sum_{k \in I_S'} c_{ik}/d_k \, e_k, -s_i)) & \text{for } (i, j) \in I_\Omega' \times I_S', \\ a_{ij} &= a((\sum_{k \in I_S'} c_{jk}/d_k \, e_k, -s_j), (\sum_{k \in I_S'} c_{ik}/d_k \, e_k, -s_i)) & \text{for } (i, j) \in I_S' \times I_S'. \end{aligned}$$

As already mentioned above, only a few terms in $\sum_{k \in I_S'} c_{kj}/d_k \, e_k$ are nonzero. The remaining terms are easy to identify and to form from the connectivity of the mesh through the assembly process at the element level. Hence, the matrix not only remains sparse but can also be assembled in the usual way. This explains why we have called our boundary condition quasi-local.

Once the unknowns u'_j have been computed, the coefficients u_j are recovered by $u_j = u'_j$, for $j \in I'_\Omega$, and $u_j = \sum_{k \in I'_S} (c_{kj}/d_j)u'_k$, for $j \in I'_S$.

Remark. Usually, the lumping process is given by an approximate quadrature formula on each element. This process can be seen as a nonconforming way to deal with the exact discrete equations [11] and, at least for the coercive case, can be tackled through a consistency estimate of the resulting approximation. In [11], it is also established that in the coercive case the order of convergence is unaltered as long as the quadrature formula exactly integrates polynomials of degree $\leq 2m - 2$ for locally polynomial shape functions of degree $\leq m$. This restricts m to be ≤ 2. Finally, as is shown in [3] in a more difficult situation, the estimates related to the consistency error for the coercive case remain valid for a Fredholm alternative obtained through compact perturbation for any sufficiently small meshsize h.

5.3. Numerical results.

5.3.1. *A first example.* The first example deals with a two-dimensional semi-infinite waveguide

$$\Omega' = \mathbb{R}^+ \times\ ]0, \pi[$$

with a homogeneous Neumann boundary condition along the boundary.

The eigenvalues are given explicitly here by

$$\lambda_n = (n - 1)^2 \quad \text{for } n \geq 1.$$

Since the case where only one mode is propagating is well known (cf. e.g., [21]), we consider here a wavenumber $k = 1.3$. In this situation, there are two modes which propagate. The waveguide is loaded by a point source located at $(0, 3\pi/4)$ modeling a coaxial loading (Fig. 2).

The computational domain is denoted by $\Omega_L :=]0, L[\times]0, \pi[$. The reference solution u_E is the restriction to $\Omega := \Omega_\pi$ of the solution computed on $\Omega_{4\pi}$. The solution computed on Ω_L for $L \geq \pi$ is denoted by u_L. The contour curves of $|u_\pi|$, $|u_{3\pi/2}|$ and $|u_{2\pi}|$ are shown on Fig. 2. The left column corresponds to $h = \pi/32$ and the right one to $h = \pi/48$.

The relative error in decibels, $20 \log_{10}(|u_L - u_E|_{\infty,\Omega}/|u_E|_{\infty,\Omega})$, is given in table 1 for different values of L and of the mesh length h. The results confirm the exponential decay of the error with respect to L as it is theoretically predicted in theorem 3.1. Moreover, the value of the reflected energy due to the terminating boundary condition is of the same order as the one resulting from using the Berenger's perfectly matched layer method [32].

TABLE 1. Exponential decay of the error $|u_L - u_E|/|u_E|$ (dB) with respect to L

L	$h = \pi/16$	$h = \pi/32$	$h = \pi/48$
0	-21.7	-23.1	-23.6
$\pi/4$	-42.7	-43.9	-44.4
$\pi/2$	-48.7	-63.1	-66.3
$3\pi/4$	-53.8	-75.8	-85.7
π	-62.2	-67.3	-72.6

$(h = \pi/32)$ $(h = \pi/48)$
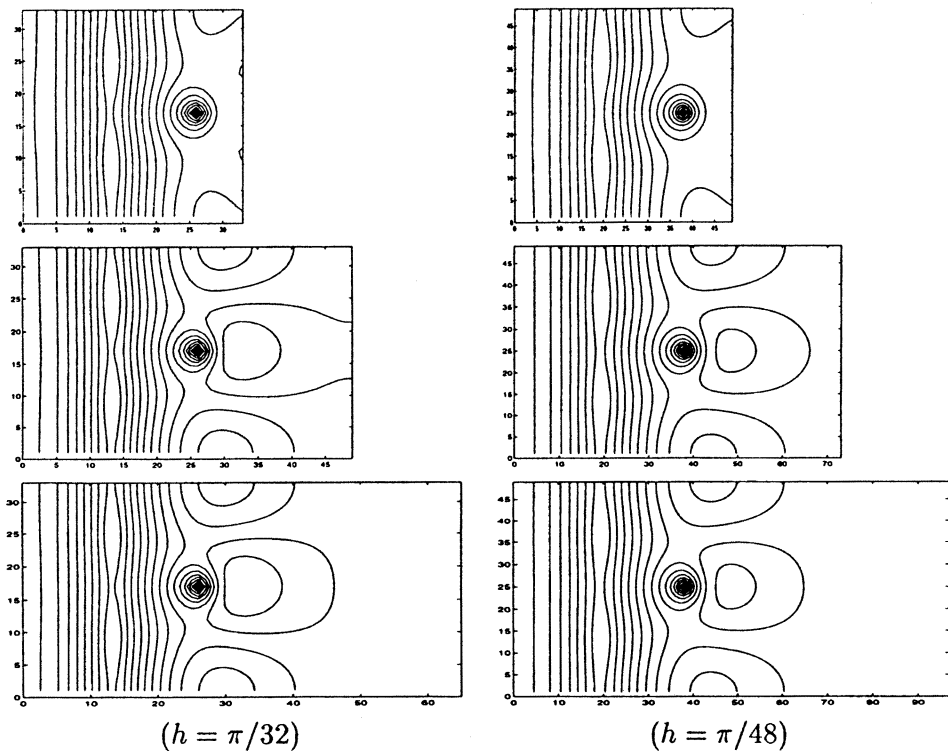
FIGURE 2. Isovalues curves of $|u|$ for $L = 0$, $L = \pi/2$ and $L = \pi$

5.3.2. *A second example.* The second example consists of a two-dimensional infinite waveguide $\Omega' = \mathbb{R} \times\ ]0,\ \pi[$, again with a homogeneous Neumann boundary condition. The eigenvalues are the same as in the previous example. Now we consider four propagating modes corresponding to $k = 3.3$. Two cross-sections are fixed to limit the waveguide:

$$S_0 = \{0\} \times\ ]0,\ \pi[\quad \text{and} \quad S_\pi = \{\pi\} \times\ ]0,\ \pi[.$$

The waveguide is fed on the interface S_0 by the first four modes ε_n^+, $n = 1$, 2, 3, 4, one by one. The energy of each mode reflected by the opposite section is given in tables 2 and 3 corresponding to the successive mesh lengths $h = \pi/10$, $\pi/20$, $\pi/40$ and $\pi/80$. These tables demonstrate the stability of the quasi-local formulation. The low-level energy of the reflected modes gives an indication of the great accuracy reached by this method.

TABLE 2. Reflected energy, $h = \pi/10$ and $h = \pi/20$

Incident wave	ε_1^+	ε_2^+	ε_3^+	ε_4^+	Incident wave	ε_1^+	ε_2^+	ε_3^+	ε_4^+
$|u_1^-|^2$ (dB)	-32	$-\infty$	-53	$-\infty$	$|u_1^-|^2$ (dB)	-41	-280	-72	$-\infty$
$|u_2^-|^2$ (dB)	-280	-42	$-\infty$	-38	$|u_2^-|^2$ (dB)	-268	-49	$-\infty$	-57
$|u_3^-|^2$ (dB)	-49	$-\infty$	-23	$-\infty$	$|u_3^-|^2$ (dB)	-69	-274	-35	-280
$|u_4^-|^2$ (dB)	-280	-43	$-\infty$	-11	$|u_4^-|^2$ (dB)	-258	-62	-268	-22

TABLE 3. Reflected energy, $h = \pi/40$ and $h = \pi/80$

Incident wave	ε_1^+	ε_2^+	ε_3^+	ε_4^+	Incident wave	ε_1^+	ε_2^+	ε_3^+	ε_4^+
$|u_1^-|^2$ (dB)	-53	-270	-91	-280	$|u_1^-|^2$ (dB)	-65	-256	-110	-268
$|u_2^-|^2$ (dB)	-258	-59	-274	-76	$|u_2^-|^2$ (dB)	-242	-71	-261	-94
$|u_3^-|^2$ (dB)	-88	-264	-48	-266	$|u_3^-|^2$ (dB)	-106	-253	-60	-257
$|u_4^-|^2$ (dB)	-253	-81	-256	-33	$|u_4^-|^2$ (dB)	-238	-99	-250	-46

5.3.3. *A third example.* The two examples above are based on the boundary condition analyzed in this paper. The rational function R interpolates the N real values k_n at the N points λ_n^2, $1 \leq n \leq N$. The function R is real and satisfies the assumption $R(\lambda_n)^2 > 0$ for $n \geq 1$. However, it seems that the later assumption is unnecessary, and only the interpolation of the constants of propagation k_n is important in the approximation process, even if some of these constants are complex. The following example gives an illustration of this fact.

We use here the semi-infinite waveguide described in the first example in the case of one propagating mode: $k = 0.5$. We have tested the four different boundary conditions

$$\partial_\mathbf{n} u - T_j u = 0, \qquad 1 \leq j \leq 4,$$

corresponding to the four functions $R_j(t)$, $1 \leq j \leq 4$, respectively defined by

$$
\begin{aligned}
R_1(t) &= b_1, \\
R_2(t) &= a_2 t + b_2, \\
R_3(t) &= a_3 t + b_3 + \frac{c_3}{1+t}, \\
R_4(t) &= a_4 t + b_4 + \frac{c_4}{1+t} + \frac{d_4}{2+t},
\end{aligned}
$$

where the coefficients a_j, b_j, c_j, d_j are chosen so that $R_j(\lambda_n^2) = k_n$, $1 \leq n \leq j \leq 4$. The first function leads to the well-known boundary condition $\partial_\mathbf{n} u - i k_1 u = 0$, extensively used in electromagnetic calculations [21]. Except for this one, since $R_j(t)$ is not a real-valued function for $j \geq 2$, the theoretical results of this paper cannot be applied to ensure convergence of the finite-element scheme. However, taking into account evanescent modes in the formulation leads to accurate results even when only a small part of the waveguide is included in the domain of computation. Tables 4 and 5 show the results that are obtained, with the same notation as in table 1. Column j corresponds to the boundary condition $\partial_\mathbf{n} u - T_j u = 0$.

TABLE 4. Error $|u_L - u_E|/|u_E|$ in dB, comparison of the four boundary conditions

L	T_1	T_2	T_3	T_4	L	T_1	T_2	T_3	T_4
0	-28.4	-37.3	-56.0	-55.1	0	-33.4	-41.7	-67.0	-71.8
$\pi/4$	-53.4	-63.2	-64.9	-64.9	$\pi/4$	-59.1	-67.8	-74.4	-74.4
$\pi/2$	-64.1	-64.2	-64.2	-64.2	$\pi/2$	-73.9	-74.1	-74.0	-74.0
$3\pi/4$	-66.3	-66.3	-66.3	-66.3	$3\pi/4$	-76.9	-76.9	-76.9	-76.9
π	-72.1	-72.1	-72.1	-72.1	π	-85.1	-85.1	-85.1	-85.1
		$(h = \pi/8)$					$(h = \pi/16)$		

TABLE 5. Error $|u_L - u_E|/|u_E|$ in dB, comparison of the four boundary conditions

L	T_1	T_2	T_3	T_4	L	T_1	T_2	T_3	T_4
0	-35.6	-43.9	-72.1	-84.6	0	-36.4	-44.7	-73.6	-88.3
$\pi/4$	-61.8	-70.3	-86.1	-86.0	$\pi/4$	-62.7	-71.1	-93.1	-93.1
$\pi/2$	-85.5	-85.9	-85.9	-85.8	$\pi/2$	-88.9	-93.0	-92.9	-92.9
$3\pi/4$	-89.4	-89.4	-89.4	-89.4	$3\pi/4$	-96.8	-96.8	-96.8	-96.8
π	-101.3	-101.3	-101.3	-101.3	π	-111.4	-111.4	-111.4	-111.4
	$(h = \pi/32)$					$(h = \pi/48)$			

Of particular interest is the second condition $\partial_\mathbf{n} u - T_2 u = 0$. This condition is completely local, hence very easy to implement on an existing code based on the first usual condition. The improvement is then of about $10\,\mathrm{dB}$ when the cross-section S_L is near to the source.

It is worth noting that when the cross-section S_L is far from the source (last lines of the tables), the four boundary conditions give the same results.

6. FINAL COMMENTS

The same technique can be applied to the three-dimensional Maxwell equations and will be presented in a forthcoming paper. It would be interesting to apply the quasi-local boundary condition to the time-domain formulations of waveguide problems.

ACKNOWLEDGMENT

REFERENCES

[1] S. Agmon, A. Douglis and L. Nirenberg, *Estimates near the boundary for solutions of elliptic partial differential equations satisfying general boundary conditions. Part II*, Comm. Pure Appl. Math. **27** (1964), 35–92. MR **28**:5252

[2] A. Bayliss and E. Turkel, *Radiation boundary conditions for wave-like equations*, Comm. Pure Appl. Math., **33** (1980), 707–725. MR **82b**:65091

[3] A. Bendali, *Numerical analysis of the exterior boundary value problem for the time-harmonic Maxwell equations by a boundary finite element method. Part 2: the discrete problem*, Math. of Comp., **43** 167 (1984), 47-68. MR **86i**:65071b

[4] A. Bendali and M. Souilah, *Consistency estimates for a double-layer potential and application to the numerical analysis of the boundary-element approximation of acoustic scattering by a penetrable object*, Math. of Comp. **62** (1994), 65–91. MR **94c**:65146

[5] J.-P. Berenger, *A Perfectly Matched Layer for the Absorbtion of Electromagnetic Waves*, Journal of Computational Physics, **114** (1994), 185–200. MR **94c**:78002

[6] C. Bernardi, *Optimal finite-element interpolation on curved domains*, SIAM J. of Num. Anal., **26** (1989), 1212–1240. MR **91a**:65228

[7] J. H. Bramble and V. Thomée, *Discrete time Galerkin methods for a parabolic boundary value problem*, Ann. Mat. Pura Appl., **101** (1974), 115–152. MR **52**:9639

[8] Z. Bi, K. Wu, C. Wu and J. Litva, *A Dispersive Boundary Condition for Microstrip Component Analysis Using The FD-TD Method*, IEEE Transactions on Microwave Theory and Techniques, **40** (1992), 774–777.

[9] J. Chazarain and A. Piriou, *Introduction to the theory of linear partial differential equations*, North-Holland, Amsterdam and New-York, 1982. MR **83j**:35001

[10] P. G. Ciarlet, *The finite element method for elliptic problems.* North-Holland, Amsterdam and New-York, 1978. MR **58**:25001

[11] P. G. Ciarlet, *Basic Error Estimates for Elliptic Problems,* in Handbook of Numerical Analysis, Vol. II, Finite Element Methods (Part 1), (P. G. Ciarlet and J. L. Lions, eds.), North-Holland, Amsterdam, 1991, 17-351. MR **92f**:65001

[12] B. Engquist and A. Majda, *Absorbing Boundary Conditions for the Numerical Simulation of Waves,* Math. of Comp. **31** 139 (1977), 629–651. MR **55**:9555

[13] B. Engquist and A. Majda, *Radiation boundary conditions for acoustic and elastic wave calculations,* Comm. Pure Appl. Math. **32** (1979), 313–357. MR **80e**:76041

[14] Ch. I. Goldstein, *A Finite Element Method for Solving Helmholtz Type Equations in Waveguides and Other Unbounded Domains,* Math. of Comp. **39** 160 (1982), 309–324. MR **84e**:65112

[15] P. Grisvard, *Elliptic Problems in Non-Smooth Domains,* Pitman, Boston, London and Melbourne, 1985. MR **86m**:35044

[16] P. Grisvard, *Caractérisation de quelques espaces d'interpolation,* Arch. Rational Mech. Anal., **25** (1967), 40–63. MR **35**:4718

[17] J.-C. Guillot and C. H. Wilcox, *Steady-State Wave Propagation in Simple and Compound Acoustic Waveguides,* Math. Z. **160** (1978), 89–102. MR **80c**:35023

[18] Ph. Guillaume and M. Masmoudi, *Solution to the time-harmonic Maxwell's equations in a waveguide, use of higher order derivatives for solving the discrete problem,* SIAM J. Numer. Anal., **34** (1997), 1306–1330. CMP 97:16

[19] L. Halpern and L. N. Trefethen, *Well-posedness of one-way wave equations and absorbing boundary conditions,* Math. of Comp. **47** (1986), 421–435. MR **88b**:65148

[20] R. L. Higdon, *Numerical Absorbing Boundary Conditions for the Wave Equation,* Math. of Comp., **49** 179 (1987), 65–90. MR **88f**:65168.

[21] J. Jin, *The Finite Element Method in Electromagnetics,* John Wiley & Sons, New-York, 1993.

[22] D. S. Jones, *The eigenvalues of $\Delta u + \lambda u = 0$ when the boundary conditions are given on semi-infinite domains,* Poc. Cambridge Philos. Soc., **49** (1953), 668–684. MR **15**:319c

[23] M. Lenoir and A. Tounsi, *The localized finite element method and its application to the two-dimensional sea-keeping problem,* SIAM J. Numer. Anal. **25** (1988), 729–752. MR **89k**:65138

[24] E. L. Lindman, *"Free-space" boundary conditions for the time-dependent wave equation,* J. Comp. Phys. **18** (1975), 66–78.

[25] J.-L. Lions and E. Magenes, *Non-Homogeneous Boundary Value Problems and Applications,* Vol. I, Springer-Verlag, Berlin, Heidelberg and New-York, 1972. MR **50**:2670

[26] W. Pascher and R. Pregla, *Analysis of rectangular waveguide junctions by the method of lines,* IEEE Trans. on Microwaves Theory and Techniques **43** (1995), 2649–2653.

[27] C. E. Reuter, R. M. Joseph, E. T. Thiele, D. S. Katz and A. Taflove, *Ultrawideband Absorbing Boundary Condition for Termination of Waveguiding Structures in FD-TD Simulations,* IEEE Microwave and guided wave letters **4** (1994), 344–346.

[28] G. Strang and G. J. Fix, *An analysis of the finite-element method,* Prentice-Hall, Englewood Cliffs, New Jersey, 1973. MR **56**:1747

[29] A. Taflove, *Computational Electrodynamics, The Finite-Difference Time-Domain Method,* Artech House, Boston and London, 1995. MR **96f**:78001

[30] M. Taylor, *Pseudodifferential operators,* Princeton University Press, Princeton, 1981. MR **82i**:35172

[31] C. P. Wu, *Variational and Iterative Methods for Waveguides and Arrays,* in "Computer Techniques for Electromagnetics", (R. Mittra Ed.), Pergamon, New-York, 1973, 265–304. MR **57**:2211

[32] Z. Wu and J. Fang, *Numerical implementation and performance of perfectly matched layer boundary condition for waveguide structure,* IEEE Trans. on Microwaves Theory and Techniques, **43** (1995), 2676–2683.

DÉPARTEMENT DE GÉNIE MATHÉMATIQUE, INSA TOULOUSE & CNRS-UMR 5640 MIP, AVENUE DE RANGUEIL, 31077 TOULOUSE CEDEX, FRANCE
  E-mail address: `bendali@gmm.insa-tlse.fr`

  E-mail address: `guillaum@gmm.insa-tlse.fr`