Equation de Helmholtz : etude numerique de quelques precondi-tionnements pour la methode GMRES

Anabelle Zebic

UNITÉ DE RECHERCHE
NRIA-ROCQUENCOURT

Institut National
de Recherche
en Informatique
et en Automatique

Domaine de Voluceau
Rocquencourt
B.P 105
78153 Le Chesnay Cedex
France
Tél. (1) 39 63 55 11

Rapports de Recherche

1 9 9 2

25ème anniversaire

N° 1802

Programme 6
Calcul Scientifique, Modélisation et
Logiciels numériques

EQUATION DE HELMHOLTZ : ÉTUDE NUMÉRIQUE DE QUELQUES PRÉCONDITIONNEMENTS POUR LA METHODE GMRES

Anabelle ZEBIC

Décembre 1992

RR_1802

EQUATION DE HELMHOLTZ : ETUDE NUMERIQUE DE QUELQUES PRECONDITIONNEMENTS POUR LA METHODE GMRES

Anabelle ZEBIC

INRIA, Projet MENUSIN, B.P. 105, Rocquencourt,

78153 Le Chesnay Cedex, France

Résumé :

Nous étudions la résolution numérique d'un problème de diffraction d'une onde harmonique par un obstacle bidimensionnel.

L'originalité de notre travail vient d'une part, du fait que nous utilisons une condition aux limites totalement transparente sur la frontière artificielle, d'autre part de la résolution numérique de l'approximation par éléments finis par une méthode itérative dans des sous-espaces de Krylov (GMRES linéaire).

Nous présentons deux approches différentes pour résoudre notre système linéaire et, dans chacun des cas, nous testons plusieurs préconditionneurs qui visent à accélérer la convergence de notre solveur.

Mots-clé : Equation de Helmholtz - Diffraction - Conditions aux limites artificielles - Méthode d'éléments finis - Méthode GMRES linéaire - Opérateur de préconditionnement - Factorisation de CROUT (complète et incomplète).

HELMHOLTZ EQUATION : NUMERICAL STUDY OF SOME PRECONDITIONERS FOR THE GMRES METHOD

Abstract :

We study in this paper the numerical solution of a bidimensional static scattering problem for harmonic waves.

The originality of our work is an implementation of a completely transparent boundary condition associated with an iterative solution of the finite element approximation by a Krylov subspace method (linear GMRES).

We present two different approaches to solve our linear system and, in both cases, we test several preconditioners in order to accelerate the rate of convergence of our solver.

Key words : Helmholtz equation - Scattering - Artificial boundary conditions - Finite element method - Linear GMRES method - Preconditioning operator - (complete and incomplete) CROUT factorization.

Table des Matières

1. Introduction

Les études présentées ici sont motivées par leurs nombreuses applications au sein de l'industrie. En effet, l'étude de la diffraction d'ondes est présente dans de multiples domaines tels que la géophysique (prospection pétrolière), l'océanographie (diffraction de la houle), ou encore la furtivité radar des avions dont le but est de minimiser leur réponse aux ondes incidentes.

La plupart de ces phénomènes de diffraction sont régis par les équations de l'électromagnétisme (équations de Maxwell) et en particulier par l'équation de Helmholtz lorsque l'on se place en régime harmonique (c'est-à-dire à pulsation ω donnée).

Dans le présent travail, nous nous intéressons aux problèmes statiques de diffraction d'une onde harmonique par un obstacle en deux dimensions. Nous choisissons le régime harmonique car c'est très souvent la situation rencontrée dans la pratique et car toute onde peut se décomposer en une somme d'ondes monochromatiques.

Dans un premier temps et comme pour toute résolution de problème extérieur, il est nécessaire d'introduire une frontière artificielle autour de l'obstacle. Nous devons y imposer une condition aux limites qui remplace la condition de radiation à l'infini et doit minimiser, voire annihiler les réflexions artificielles. De nombreuses études ont été menées à ce sujet (cf. par exemple [EM], [BT], [Fen], [KG], [G]) mais peu ont conduit à des conditions aux limites totalement transparentes. Nous avons retenu une de ces conditions (cf. [KG]) : la condition Dirichlet to Neumann (DtN). C'est une condition exacte non locale qui s'est souvent avérée très robuste et très efficace (cf. [KG]).

La suite de notre travail consiste en l'approximation par une méthode d'éléments finis $\mathcal{P}^1$ du système d'équations obtenu. Le système linéaire auquel nous aboutissons est un système complexe creux dont la matrice ne possède pas de propriétés particulières. Pour le résoudre, nous analysons deux approches différentes. La première consiste à transformer le système en un "double" système réel avant de le résoudre et la deuxième à conserver le système complexe tel quel. Dans les deux cas, nous utilisons la méthode GMRES linéaire (cf. [SS]), dont nous présentons par ailleurs les principales propriétés.

En vue d'applications industrielles nécessitant la résolution de grands systèmes, nous nous sommes particulièrement intéressés à accélérer la convergence de notre

solveur. Ainsi, nous développons et comparons plusieurs préconditionneurs que nous testons numériquement sur le système réel et sur le système complexe. En outre, plusieurs résultats numériques illustrent le comportement du préconditionneur retenu et du solveur lorsqu'un paramètre, tel que le nombre de points par longueur d'onde, varie.

Nous avons organisé notre travail de la façon suivante :

Le chapitre 2 est constitué de rappels concernant l'établissement des équations qui régissent les phénomènes électromagnétiques.

Le chapitre 3 est consacré aux principales justifications théoriques de la méthode. Dans un premier temps, nous donnons quelques unes des conditions aux limites les plus souvent utilisées sur la frontière artificielle et montrons comment obtenir la conditon DtN. Nous établissons finalement la formulation variationnelle du problème tronqué et le résultat d'existence et d'unicité qui lui est associé.

Le chapitre 4 traite de la résolution numérique du problème par une méthode d'éléments finis $\mathcal{P}^1$. Nous donnons deux analyses différentes pour résoudre le système linéaire obtenu et présentons la méthode de résolution choisie, la méthode GMRES linéaire. Pour finir nous comparons, dans les deux cas, les différents opérateurs de préconditionnement utilisés.

Le chapitre 5 présente quelques résultats numériques qui montrent le comportement du préconditionneur choisi ainsi que la validité de la méthode.

Le chapitre 6 expose les conclusions de cette étude.

2. Rappels sur les équations fondamentales de l'électromagnétisme

Avant de poser les problèmes de diffraction qui nous intéressent, rappelons brièvement les équations de base de l'électromagnétisme. Le lecteur désirant plus de détails peut se référer entres autres à [R], [Fen], [DL] ou [Fou].

2.1. Position du problème physique

L'étude électromagnétique d'un milieu continu conduit à déterminer un champ électromagnétique constitué par quatre champs vectoriels :

. $\vec{E}$ le champ électrique,

. $\vec{D}$ l'induction électrique,

. $\vec{H}$ le champ magnétique,

. $\vec{B}$ l'induction magnétique.

Ces champs vérifient les équations de Maxwell données par :

$$(1)\qquad \begin{cases} \overrightarrow{rot}\,\vec{E} + \dfrac{\partial \vec{B}}{\partial t} = 0, & \text{(loi de l'induction)} \qquad (1.a) \\ \overrightarrow{rot}\,\vec{H} - \dfrac{\partial \vec{D}}{\partial t} = \vec{J}, & \text{(loi d'Ampère)} \qquad (1.b) \\ div\vec{D} = \rho, & \text{(loi de Gauss électrique)} \qquad (1.c) \\ div\vec{B} = 0, & \text{(loi de Gauss magnétique)} \qquad (1.d) \end{cases}$$

avec :

. $\vec{J}$ la densité de courant électrique,

. ρ la densité de charge électrique.

L'équation dite de conservation ou de continuité relie les densité $\vec{J}$ et ρ.
Elle s'écrit :

$$(2)\qquad div\vec{J} + \frac{\partial \rho}{\partial t} = 0.$$

Remarque 1 :

Les deux lois de Gauss sont des conséquences immédiates de (1.a), (1.b) et (2).
Les milieux parfaits (qui sont linéaires, isotropes et homogènes) vérifient :

$$(3)\qquad \begin{cases} \vec{D} = \varepsilon \vec{E}, & (3.a) \\ \vec{B} = \mu \vec{H}, & (3.b) \end{cases}$$

avec

. ε la permittivité diélectrique,

. μ la perméabilité magnétique. □

Remarque 2 :

Les milieux vérifiant (3.a) et (3.b) sont appelés respectivement diélectriques parfaits et magnétiques parfaits. □

Les milieux conducteurs vérifient la loi d'Ohm :

$$\vec{J} = \sigma \vec{E}, \tag{4}$$

avec σ la conductivité du milieu.

Dans le vide, $\varepsilon = \varepsilon_o$ et $\mu = \mu_o$ où les constantes ε_o et μ_o sont déterminées par le système d'unité. Dans le système U.S.I. :

$$\begin{cases} \varepsilon_o = \dfrac{1}{36\pi} 10^{-9} F/m, \\ \mu_o = 4\pi 10^{-7} H/m, \\ \varepsilon_o \mu_o c_o^2 = 1, \end{cases} \tag{5}$$

avec $c_o = 3.10^8 m/s$ la vitesse de la lumière dans le vide.

A l'aide des équations (1), (2), (3) et (4), on montre que $\vec{E}$ et $\vec{H}$ vérifient les équations :

$$\begin{cases} \varepsilon \dfrac{\partial^2 \vec{E}}{\partial t^2} + \sigma \dfrac{\partial \vec{E}}{\partial t} + \overrightarrow{rot}\,(\dfrac{1}{\mu} \overrightarrow{rot}\,\vec{E}) = 0, \\ \varepsilon \dfrac{\partial^2 \vec{H}}{\partial t^2} + \sigma \dfrac{\partial \vec{H}}{\partial t} + \overrightarrow{rot}\,(\dfrac{1}{\mu} \overrightarrow{rot}\,\vec{H}) = 0. \end{cases} \tag{6}$$

En milieu homogène non conducteur, on obtient les équations des ondes pour $\vec{E}$ et $\vec{H}$:

$$\begin{cases} \varepsilon \mu \dfrac{\partial^2 \vec{E}}{\partial t^2} - \Delta \vec{E} = 0, \\ \varepsilon \mu \dfrac{\partial^2 \vec{H}}{\partial t^2} - \Delta \vec{H} = 0, \end{cases} \tag{7}$$

avec une vitesse de propagation donnée par $c = \frac{1}{\sqrt{\varepsilon \mu}}$ qui est la vitesse de la lumière dans le milieu.

En régime harmonique, avec une dépendance temporelle en e^{-iwt}, les équations (1), (2), (3) et (4) donnent :

$$(8) \qquad \begin{cases} \overrightarrow{rot}\ \vec{E} - iw\mu\vec{H} = 0, \\ \overrightarrow{rot}\ \vec{H} + iw\varepsilon\vec{E} = \vec{J}, \\ div\vec{J} - iw\rho = 0, \\ \vec{J} = \sigma\vec{E}. \end{cases}$$

2.2. Un problème statique de diffraction 3D

Nous considérons un obstacle borné Ω_o de $\mathcal{R}^3$ de bord Γ, parfaitement conducteur, percuté par une onde électromagnétique incidente connue, monochromatique de fréquence $f = \frac{\omega}{2\pi}$.

L'onde se propage dans un milieu extérieur Ω_e que nous choisissons homogène, isolant, diélectrique et magnétique parfait, de constantes ε et μ.

Nous nous intéressons au champ électromagnétique statique diffracté par l'obstacle.

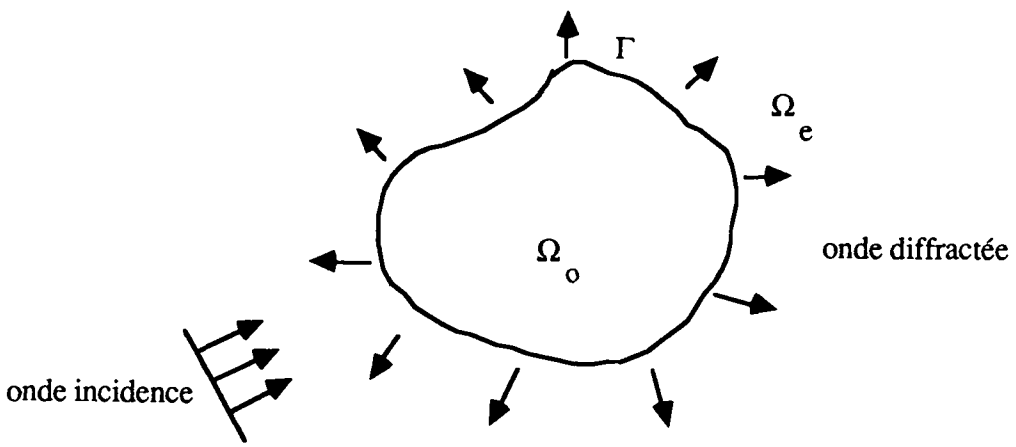


Figure 1 : Diffraction par un obstacle.

Le milieu extérieur Ω_e étant non conducteur, nous avons :

$$\vec{J} = 0, \quad \rho = 0 \quad \text{dans} \quad \Omega_e.$$

L'obstacle étant parfaitement conducteur, nous avons en outre :

$$\vec{E} = \vec{H} = 0 \quad \text{dans} \quad \Omega_o.$$

De plus, la composante normale de $\vec{B}$ et les composantes tangentielles de $\vec{E}$ sont continues à la traversée de Γ. Nous obtenons donc les conditions aux limites :

$$\vec{E} \wedge \vec{n} = 0 \quad \text{sur} \quad \Gamma,$$

$$\vec{H}.\vec{n} = 0 \quad \text{sur} \quad \Gamma.$$

Nous choisissons un champ incident vérifiant les équations de Maxwell dans $\mathcal{R}^3$. Le champ harmonique diffracté est donc solution du problème aux limites :

$$(9) \quad \begin{cases} \overrightarrow{rot}\vec{E}^d - iw\mu\vec{H}^d = 0 \quad \text{dans} \quad \Omega_e, & (9.a) \\ \overrightarrow{rot}\vec{H}^d + iw\epsilon\vec{E}^d = 0 \quad \text{dans} \quad \Omega_e, & (9.b) \\ \vec{E}^d \wedge \vec{n} = -\vec{E}^{inc} \wedge \vec{n} \quad \text{sur} \quad \Gamma, & (9.c) \\ \vec{H}^d.\vec{n} = -\vec{H}^{inc}.\vec{n} \quad \text{sur} \quad \Gamma. & (9.d) \end{cases}$$

Remarque 3 :

Les équations (9.a) et (9.c) impliquent (9.d) car

$$\vec{E} \wedge \vec{n} = 0 \Rightarrow \overrightarrow{rot}\vec{E}.\vec{n} = 0. \qquad \square$$

Il est bien connu que les équations du système (9) ne suffisent pas à déterminer de façon unique les champs $\vec{E}^d$ et $\vec{H}^d$, et que pour lever l'indétermination il faut leur adjoindre une condition de radiation traduisant la propagation de l'énergie vers l'infini. L'une des conditions les plus naturelles pour ce système est la condition de Silver-Müller (cf. [Sa]) :

$$\lim_{r \to +\infty} r(\sqrt{\mu}\vec{H}^d \wedge \frac{\vec{r}}{r} - \sqrt{\epsilon}\vec{E}^d) = 0,$$

uniformément par rapport à la direction $\frac{\vec{r}}{r}$, r étant la distance euclidienne à l'origine.

Par ailleurs, l'équation (9.a) donne :

$$\vec{H}^d = -\frac{i}{\omega\mu}\overrightarrow{rot}\vec{E}^d.$$

Ceci permet d'éliminer $\vec{H}^d$ dans le système (9) et d'obtenir le système différentiel en $\vec{E}^d$. Il s'écrit :

$$(10) \quad \begin{cases} \overrightarrow{rot}\overrightarrow{rot}\vec{E}^d - k^2\vec{E}^d = 0 \quad \text{dans} \quad \Omega_e, \\ \vec{E}^d \wedge \vec{n} = -\vec{E}^{inc} \wedge \vec{n} \quad \text{sur} \quad \Gamma, \\ \lim_{r \to +\infty} r(\overrightarrow{rot}\vec{E}^d \wedge \frac{\vec{r}}{r} - ik\vec{E}^d) = 0, \end{cases}$$

où $k = \omega\sqrt{\varepsilon\mu}$ désigne le nombre d'onde.

Remarque 4 :

Une autre condition traduisant le fait que l'onde diffractée est "sortante" est la condition de radiation de Sommerfeld :

$$\lim_{r\to+\infty} r(\frac{\partial \vec{E}^d}{\partial r} - ik\vec{E}^d) = 0,$$

uniformément par rapport à la direction $\frac{\vec{r}}{r}$.
Nous savons (cf. [Sa]) que cette condition (comme plusieurs autres conditions de radiation) est équivalente à celle de Silver-Müller. □

L'équation (9.b) donne :

$$\vec{E}^d = \frac{i}{\omega\varepsilon}\overrightarrow{rot}\vec{H}^d.$$

Ainsi, de même que pour $\vec{E}^d$, nous obtenons le système différentiel en $\vec{H}^d$. Il s'écrit:

$$(11) \qquad \begin{cases} \overrightarrow{rot}\overrightarrow{rot}\vec{H}^d - k^2\vec{H}^d = 0 \quad \text{dans} \quad \Omega_e, \\ \vec{H}^d.\vec{n} = -\vec{H}^{inc}.\vec{n} \quad \text{sur} \quad \Gamma, \\ \lim_{r\to+\infty} r(\overrightarrow{rot}\vec{H}^d + ik\vec{H}^d \wedge \frac{\vec{r}}{r}) = 0. \end{cases}$$

2.3. Un problème statique de diffraction 2D

Pour nous ramener à l'étude d'un problème 2D, nous prenons un cylindre infini pour obstacle et une onde incidente plane dont le vecteur d'onde est dans le plan orthogonal aux génératrices du cylindre

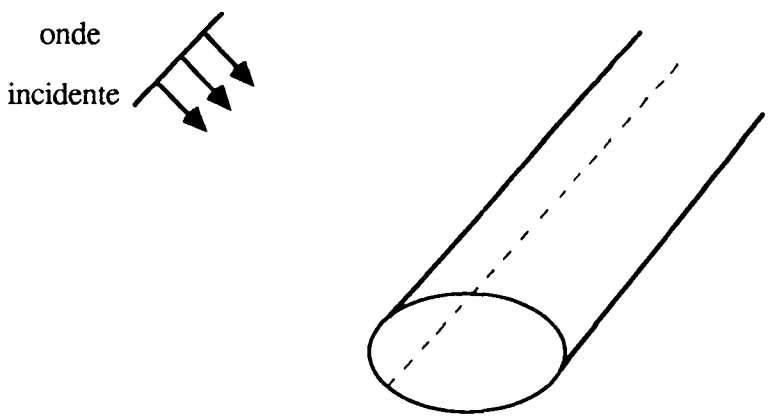


Figure 2 : Diffraction 2D par un cylindre infini.

Nous savons que ce problème de diffraction conduit à deux types de solutions $\{\vec{E}^d, \vec{H}^d\}$ indépendantes et orthogonales, chacun issu d'un problème de diffraction scalaire 2D. En effet, soit $(\vec{e}^{x_1}, \vec{e}^{x_2}, \vec{e}^{x_3})$ une base orthonormée cartésienne avec $\vec{e}^{x_3}$ suivant l'axe du cylindre. Prenons l'onde incidente telle que $\vec{H}^{inc}.\vec{e}^{x_3} = 0$. De même, le champ magnétique diffracté est tel que $\vec{H}^d.\vec{e}^{x_3} = 0$.

Comme $\vec{E}^d$ et $\vec{H}^d$ sont toujours orthogonaux, on a en fait $\vec{E}^d = (0, 0, E_3^d)$ et $\vec{H}^d = (H_1^d, H_2^d, 0)$ (avec E_3^d, H_1^d, H_2^d indépendants de la variable x_3). On dit alors que l'onde est transverse magnétique (TM) car elle correspond à un champ magnétique purement transverse.

De la même façon, si $\vec{E}^{inc}.\vec{e}^{x_3} = 0$, on a $\vec{E}^d = (E_1^d, E_2^d, 0)$ et $\vec{H}^d = (0, 0, H_3^d)$ et l'onde est dite transverse électrique (TE).

En se plaçant dans un plan quelconque orthogonal à l'axe du cylindre, et en reprenant les sytèmes (10) et (11) dans les deux cas précédents, nous obtenons les deux systèmes d'équations scalaires 2D qui déterminent E_3^d et H_3^d :

Ondes (TM)

$$(12) \qquad \begin{cases} \Delta E_3^d + k^2 E_3^d = 0 \quad \text{dans} \quad \Omega_e, \\ E_3^d = -E_3^{inc} \quad \text{sur} \quad \Gamma, \\ \lim\limits_{r \to +\infty} \sqrt{r}\left(\dfrac{\partial E_3^d}{\partial r} - ikE_3^d\right) = 0. \end{cases}$$

Ondes (TE)

$$(13) \qquad \begin{cases} \Delta H_3^d + k^2 H_3^d = 0 \quad \text{dans} \quad \Omega_e, \\ \dfrac{\partial H_3^d}{\partial n} = -\dfrac{\partial H_3^{inc}}{\partial n} \quad \text{sur} \quad \Gamma, \\ \lim\limits_{r \to +\infty} \sqrt{r}\left(\dfrac{\partial H_3^d}{\partial r} - ikH_3^d\right) = 0, \end{cases}$$

avec $k = \omega\sqrt{\varepsilon\mu}$.

3. Présentation mathématique du problème

Nous nous intéressons à la diffraction d'une onde électromagnétique connue, "venant de l'infini", sur un obstacle borné Ω_o de $\mathcal{R}^2$ de bord Γ, et parfaitement conducteur. L'onde se propage dans un milieu extérieur Ω_e homogène, isolant, diélectrique et magnétique parfait.

Le problème à résoudre est le suivant :

$$(14) \qquad \begin{cases} \text{Trouver } u \text{ tel que :} \\ \qquad \Delta u + k^2 u + f = 0 \quad \text{dans} \quad \Omega_e, \\ \qquad\qquad u = g \quad \text{sur} \quad \Gamma, \\ \lim\limits_{r \to +\infty} \sqrt{r}\left(\dfrac{\partial u}{\partial r} - iku\right) = 0, \end{cases}$$

u désignant l'onde diffractée par Γ et f une fonction source telle que $f = 0$ à l'extérieur d'un cercle de rayon R.

Notons que pour simplifier les notations, nous avons remplacé $-u^{inc}$ par g.

Le résultat d'existence et d'unicité relatif à (14) est :

Proposition 1 :

Si $f \in L^2(\Omega_e)$ à support compact et $g \in H^{\frac{1}{2}}(\Gamma)$, le problème (14) admet une unique solution $u \in H^1_{loc}(\Omega_e)$. □

On trouvera dans [Be] la démonstration pour le système 3D dont provient (14).

En vue de résoudre numériquement le problème (14), nous bornons le domaine de calcul en introduisant une frontière artificielle Σ, située à l'extérieur du support de f. Nous imposons alors sur Σ une certaine condition aux limites qui remplace la condition de radiation à l'infini et doit minimiser les réflexions artificielles.

Notons Ω le domaine borné intérieurement par Γ et extérieurement par Σ.

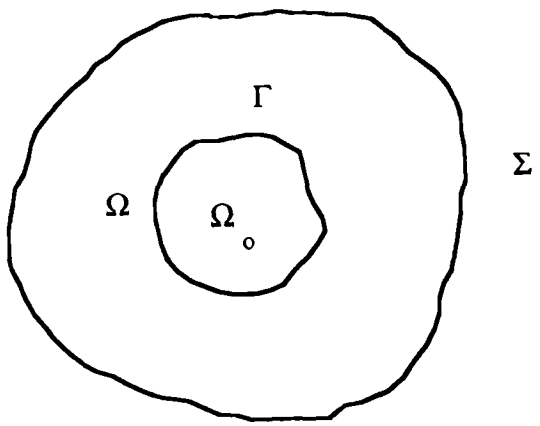


Figure 3 : Troncature du domaine de calcul.

3.1. Conditions aux limites sur la frontière artificielle

Le choix de la condition à imposer sur Σ est très important car il détermine la distance à laquelle on peut rapprocher Σ du support de f ou de Γ lorsque f est nulle. De plus, il doit aussi conduire à un problème bien posé.

Nous avons choisi une condition non locale exacte, appelée condition Dirichlet to Neuman (DtN), traduisant que Σ ne produit aucune réflexion artificielle. Elle s'écrit :

$$\frac{\partial u}{\partial n} = \frac{k}{\pi} \sum_{n=0}^{\infty}{}' \int_0^{2\pi} \frac{H_n^{(1)'}(kR)}{H_n^{(1)}(kR)} cos(n(\theta - \theta')) u(R, \theta') d\theta' \quad \text{sur} \quad \Sigma, \tag{15}$$

où

- Σ est le cercle de centre O et de rayon R, O désignant le centre géométrique de Ω_o,
- $H_n^{(1)}$ désigne la fonction de Hankel d'ordre n et du premier type,
- la notation Σ' signifie qu'un facteur $\frac{1}{2}$ multiplie le terme correspondant à $n = 0$.

Nous montrerons ultérieurement comment obtenir la condition (15).

Ainsi, la solution du problème aux limites posé dans Ω est la restriction à Ω de la solution du problème (14) posé dans tout Ω_e. On dit alors que l'on a mis sur Σ une condition aux limites transparente (C.L.T.). D'autres méthodes consistent à ne prendre que des approximations de la C.L.T. : on parle alors de conditions aux limites absorbantes (C.L.A.). Ainsi, on ne rencontre pas les difficultés numériques dues à la non-localité des C.L.T.. Cependant, on n'évite pas les réflexions artificielles sur Σ.

L'idée la plus simple est d'imposer :

$$\frac{\partial u}{\partial n} = iku \quad \text{sur} \quad \Sigma,$$

car cette condition est de la même forme que la condition de radiation de Sommerfeld. Mais il est connu que cela donne de mauvais résultats.

Pour diminuer les réflexions sur Σ, certains auteurs ont proposé des conditions aux limites d'ordre supérieur.
Nous n'en citons que quelques uns (pour plus de détails, le lecteur peut se référer à [KG] et [G]).

. Engquist et Majda (cf. [EM]) ont développé une technique, basée sur la théorie des opérateurs pseudo-différentiels qui leur permet d'obtenir une suite de C.L.A. locales d'ordre croissant.

Dans le cas où Σ est un cercle de rayon R, les deux premières conditions sont :

$$E_1 u = (\frac{\partial}{\partial r} - ik + \frac{1}{2R})u = 0 \quad \text{sur} \quad \Sigma,$$
$$E_2 u = (\frac{\partial}{\partial r} - ik + \frac{1}{2R} - \frac{i}{2kR^2}\frac{\partial^2}{\partial \theta^2} - \frac{1}{2k^2R^3}\frac{\partial^2}{\partial \theta^2})u = 0 \quad \text{sur} \quad \Sigma.$$

. L'idée de Bayliss et Turkel (cf. [BT]) est d'utiliser un développement asymptotique en $\frac{1}{r}$ de la solution u. Ainsi, ils forment une suite d'opérateurs différentiels locaux dont chaque terme E_m s'obtient en éliminant les m premiers termes du développement.

Dans le cas où Σ est un cercle de rayon R, chaque opérateur E_m est défini par :

$$E_m u = (\prod_{j=1}^{m}(-ik + \frac{\partial}{\partial r} + \frac{4j-3}{R}))u = 0 \quad \text{sur} \quad \Sigma.$$

. Feng (cf. [Fen]) obtient une condition exacte non-locale en dérivant une relation intégrale sur Σ utilisant la fonction de Green appropriée. Il approche finalement cette condition par une suite de C.L.A. locales.

Dans le cas où Σ est un cercle de rayon R, les quatre premières conditions sont :

$$F_o u = (\frac{\partial}{\partial r} - ik)u = 0 \quad \text{sur} \quad \Sigma,$$

$$F_1 u = (\frac{\partial}{\partial r} - ik + \frac{1}{2R})u = 0 \quad \text{sur} \quad \Sigma,$$

$$F_2 u = (\frac{\partial}{\partial r} - ik + \frac{1}{2R} - \frac{i}{8kR^2} - \frac{i}{2kR^2}\frac{\partial^2}{\partial\theta^2})u = 0 \quad \text{sur} \quad \Sigma,$$

$$F_3 u = (\frac{\partial}{\partial r} - ik + \frac{1}{2R} - \frac{i}{8kR^2} - \frac{1}{8k^2R^3} - \frac{i}{2kR^2}\frac{\partial^2}{\partial\theta^2} - \frac{1}{2k^2R^3}\frac{\partial^2}{\partial\theta^2})u = 0 \quad \text{sur} \quad \Sigma.$$

Cependant, nous savons (cf. [G]) que la plupart des C.L.A. ne conduisent à de faibles réflexions sur la frontière artificielle que pour certains angles d'incidence ou certaines fréquences. C'est pourquoi nous nous sommes intéressés aux C.L.T. souvent plus robustes et plus précises.

Par ailleurs, d'après Keller et Givoli (cf. [KG]), la condition DtN s'est avérée bien meilleure que la plupart des conditions locales données précédemment.

La condition DtN

Dans cette section, nous montrons comment établir la condition DtN.

Prenons pour Σ le cercle de centre O et de rayon R.

On note D le domaine extérieur à $\Omega \cup \Omega_o$ et $\vec{n}$ la normale unitaire à Σ orientée vers l'extérieur de $\Omega \cup \Omega_o$.

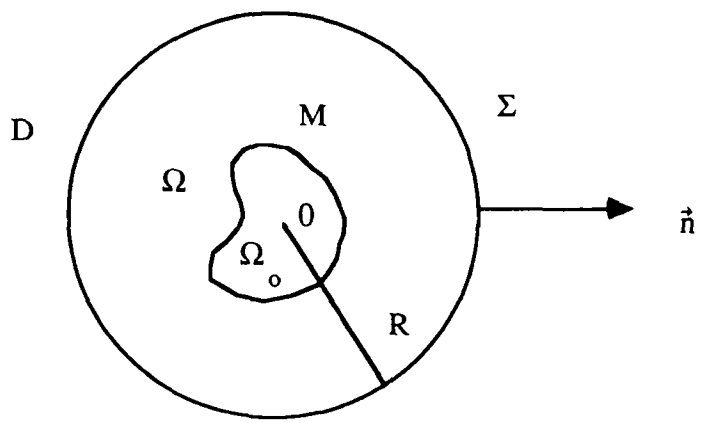


Figure 4 : Domaine de calcul de la condition DtN.

Considérons le problème suivant :

$$(16)\quad \begin{cases} \text{Trouver } u \text{ tel que :} \\ \Delta u + k^2 u = 0 \quad \text{dans} \quad D, & (16.a) \\ u = u(R,\theta) = g_R \quad \text{sur} \quad \Sigma, & (16.b) \\ \lim\limits_{r\to+\infty} \sqrt{r}(\dfrac{\partial u}{\partial r} - iku) = 0. & (16.c) \end{cases}$$

Nous savons (cf. Proposition 1) que ce problème admet une unique solution $u \in H^1_{loc}(D)$. Nous cherchons à expliciter u.

Tout d'abord, écrivons le développement en série de Fourier de u en coordonnées polaires :

$$u(r,\theta) = \sum_{n\in\mathcal{Z}} \hat{u}_n(r) e^{in\theta}, \qquad \forall r \in [R, +\infty[, \forall \theta \in [0, 2\pi],$$

où $\hat{u}_n$ désigne le coefficient de Fourier d'ordre n de u.

L'équation $\Delta u + k^2 u = 0$ s'écrit en coordonnées polaires

$$\frac{\partial^2 u}{\partial r^2} + \frac{1}{r}\frac{\partial u}{\partial r} + \frac{1}{r^2}\frac{\partial^2 u}{\partial \theta^2} + k^2 u = 0,$$

d'où :

$$\sum_{n\in\mathcal{Z}} (\frac{d^2\hat{u}_n}{dr^2} + \frac{1}{r}\frac{d\hat{u}_n}{dr} - \frac{n^2}{r^2}\hat{u}_n + k^2\hat{u}_n) e^{in\theta} = 0,$$

d'où encore, la suite d'équations différentielles pour $n \in \mathcal{Z}$

$$r^2\frac{d^2\hat{u}_n}{dr^2} + r\frac{d\hat{u}_n}{dr} + (k^2r^2 - n^2)\hat{u}_n = 0.$$

Posons $s = kr$ et appelons $\hat{v}_n$ la fonction définie sur $\mathcal{R}$ et à valeurs dans $\mathcal{C}$ telle que $\hat{v}_n(s) = \hat{u}_n(r)$.
Nous obtenons alors :

$$(17)\qquad s^2\frac{d^2\hat{v}_n}{ds^2} + s\frac{d\hat{v}_n}{ds} + (s^2 - n^2)\hat{v}_n = 0, \qquad \forall n \in \mathcal{Z}.$$

Nous savons (cf. [AS]) que les solutions de l'équation (17) sont les fonctions de Bessel d'ordre n du premier type $J_{\pm n}$, du second type Y_n et du troisième type $H_n^{(1)}, H_n^{(2)}$ (ces dernières étant aussi appelées fonctions de Hankel d'ordre n du premier et deuxième type). Toutefois, seule $H_n^{(1)}$ vérifie la condition de radiation à l'infini (16.c)

Ainsi, $\hat{u}_n$ s'écrit :

$$\hat{u}_n(r) = \alpha_n H_n^{(1)}(kr), \qquad \forall n \in \mathcal{Z},\ \forall r \in [R, +\infty[,$$

où α_n est une constante complexe indépendante de r.
Par sommation, il vient :

$$u(r,\theta) = \sum_{n\in\mathcal{Z}} \alpha_n H_n^{(1)}(kr) e^{in\theta}, \qquad \forall r \in [R, +\infty[, \quad \forall \theta \in [0, 2\pi]. \tag{18}$$

Par définition, nous avons

$$\hat{u}_n(r) = \frac{1}{2\pi} \int_0^{2\pi} u(r,\theta') e^{-in\theta'} d\theta', \qquad \forall n \in \mathcal{Z}, \quad \forall r \in [R, +\infty[,$$

d'où

$$u(r,\theta) = \frac{1}{\pi} {\sum_{n=0}^{+\infty}}' \int_0^{2\pi} u(r,\theta') cos(n(\theta - \theta')) d\theta', \qquad \forall r \in [R, +\infty[, \quad \forall \theta \in [0, 2\pi], \tag{19}$$

où nous rappelons que la notation Σ' signifie qu'un facteur $\frac{1}{2}$ multiplie le terme correspondant à $n = 0$.
En écrivant $u(r,\theta')$ sous la forme (18) et en le reportant dans (19), nous obtenons :

$$u(r,\theta) = \frac{1}{\pi} {\sum_{n=0}^{+\infty}}' \sum_{n'\in\mathcal{Z}} \alpha_{n'} H_{n'}^{(1)}(kr) \int_0^{2\pi} e^{in'\theta'} cos(n(\theta - \theta')) d\theta',$$

ce qui s'écrit

$$u(r,\theta) = \frac{1}{\pi} {\sum_{n=0}^{+\infty}}' [\pi H_n^{(1)}(kr)(\alpha_n e^{in\theta} + e^{in\pi} \alpha_{-n} e^{-in\theta})]. \tag{20}$$

En écrivant $u(R,\theta')$ sous la forme (18), nous obtenons :

$$\begin{cases} {\sum_{n=0}^{+\infty}}' \dfrac{H_n^{(1)}(kr)}{H_n^{(1)}(kR)} \displaystyle\int_0^{2\pi} cos(n(\theta - \theta')) u(R,\theta') d\theta' \\ = {\displaystyle\sum_{n=0}^{+\infty}}' \dfrac{H_n^{(1)}(kr)}{H_n^{(1)}(kR)} [\pi H_n^{(1)}(kR)(\alpha_n e^{in\theta} + e^{in\pi} \alpha_{-n} e^{-in\theta})], \end{cases}$$

ce qui nous donne finalement (via (20)) :

$$\begin{cases} u(r,\theta) = \dfrac{1}{\pi} {\displaystyle\sum_{n=0}^{+\infty}}' \displaystyle\int_0^{2\pi} \frac{H_n^{(1)}(kr)}{H_n^{(1)}(kR)} cos(n(\theta - \theta')) u(R,\theta') d\theta', \\ \forall r \in [R, +\infty[, \forall \theta \in [0, 2\pi]. \end{cases} \tag{21}$$

Soit M l'opérateur de Calderon inverse associée au problème (16)

$$M : H^{\frac{1}{2}}(\Sigma) \longrightarrow H^{-\frac{1}{2}}(\Sigma)$$
$$g_R \longrightarrow -\frac{\partial u}{\partial n}|_\Sigma,$$

avec u solution de (16).

En différentiant (21) par rapport à r et en prenant $r = R$, nous obtenons explicitement cet opérateur M.

Nous établissons ainsi une condition aux limites transparente sur Σ, en écrivant que la solution du problème (14) vérifie (16) : c'est la condition DtN. Elle s'écrit :

$$\begin{cases} \dfrac{\partial u}{\partial n} = -Mu \quad \text{sur} \quad \Sigma, \quad \text{avec} \\ Mu(R,\theta) = -\dfrac{k}{\pi}{\sum_{n=0}^{\infty}}' \displaystyle\int_0^{2\pi} \frac{H_n^{(1)'}(kR)}{H_n^{(1)}(kR)} cos(n(\theta-\theta'))u(R,\theta')d\theta', \quad \forall\theta \in [0,2\pi]. \end{cases}$$

3.2. Formulation variationnelle du problème tronqué

Nous sommes arrivés au problème tronqué suivant :

$$(22) \qquad \begin{cases} \text{Trouver } u \text{ tel que :} \\ \Delta u + k^2 u + f = 0 \quad \text{dans} \quad \Omega, \\ u = g \quad \text{sur} \quad \Gamma, \\ \dfrac{\partial u}{\partial n} = -Mu \quad \text{sur} \quad \Sigma. \end{cases}$$

Si on montre que ce problème admet une solution unique, alors on décompose la trace de cette solution sur Σ en série de Fourier, et le prolongement défini par (21) donne la solution du problème (14).

Pour définir la formulation faible du problème (22), nous introduisons les deux espaces fonctionnels suivants :

$$.V = H^1(\Omega),$$
$$.V_o = \{v \ ; \ v \in H^1(\Omega) \ ; \ v = 0 \quad \text{sur} \quad \Gamma\}.$$

Par ailleurs, si $g \in H^{\frac{1}{2}}(\Gamma)$, l'application trace $\gamma_o : H^1(\Omega) \to H^{\frac{1}{2}}(\Gamma)$ ayant un relèvement continu, il existe une fonction $\tilde{g} \in H^1(\Omega)$ telle que $\gamma_o \tilde{g} = g$.

Nous avons alors :

Proposition 2 :

Soit $f \in L^2(\Omega)$ à support compact, $g \in H^{\frac{1}{2}}(\Gamma)$ et $\tilde{g} \in H^1(\Omega)$ tel que $\gamma_o \tilde{g} = g$. Toute solution $u \in H^1(\Omega)$ du problème (22) est solution du problème variationnel suivant :

$$(23) \qquad \begin{cases} \text{Trouver } u \in V \text{ tel que : } (u - \tilde{g}) \in V_o \quad \text{et} \\ \int_\Omega (\vec{\nabla} u . \vec{\nabla} \bar{v} - k^2 u \bar{v}) dx + \int_\Sigma M u \bar{v} d\sigma = \int_\Omega f \bar{v} dx, \qquad \forall v \in V_o. \end{cases} \qquad \square$$

Démonstration :

On suppose tout d'abord que f et u sont très régulières, ce qui nous permet d'écrire:

$$\int_\Omega (\Delta u + k^2 u) \bar{v} dx = - \int_\Omega f \bar{v} dx, \qquad \forall v \in \mathcal{D}(\bar{\Omega}),$$

avec $\mathcal{D}(\bar{\Omega})$ la restriction à $\bar{\Omega}$ de l'ensemble des fonctions infiniment dérivables sur $\mathcal{R}^2$. D'où, en utilisant une formule de Green :

$$\int_\Omega (\vec{\nabla} u . \vec{\nabla} \bar{v} - k^2 u \bar{v}) dx - \int_\Gamma \frac{\partial u}{\partial n} \bar{v} d\gamma - \int_\Sigma \frac{\partial u}{\partial n} \bar{v} d\sigma = \int_\Omega f \bar{v} dx, \qquad \forall v \in \mathcal{D}(\bar{\Omega}).$$

Par densité, cette égalité reste vraie dans V_o. De plus, $\dfrac{\partial u}{\partial n} = -Mu$ sur Σ.
D'où :

$$\int_\Omega (\vec{\nabla} u . \vec{\nabla} \bar{v} - k^2 u \bar{v}) dx + \int_\Sigma M u \bar{v} d\sigma = \int_\Omega f \bar{v} dx, \qquad \forall v \in V_o. \qquad \square$$

Nous donnons désormais le résultat d'existence et d'unicité relatif au problème (23):

Théorème 1 :

Le problème :

$$(24) \qquad \begin{cases} f \in L^2(\Omega) \text{ à support compact, } g \in H^{\frac{1}{2}}(\Gamma) \text{ données,} \\ \text{trouver } u \in V \text{ tel que : } (u - \tilde{g}) \in V_o \quad \text{et} \\ \int_\Omega (\vec{\nabla} u . \vec{\nabla} \bar{v} - k^2 u \bar{v}) dx + \int_\Sigma M u \bar{v} d\sigma = \int_\Omega f \bar{v} dx, \qquad \forall v \in V_o, \end{cases}$$

admet une solution unique. $\square$

Démonstration :

Soient $u, v \in V$ avec $u - \tilde{g} \in V_o$.
Soient a la forme sesquilinéaire définie par :

$$a(u,v) = \int_\Omega (\vec{\nabla} u . \vec{\nabla} \bar{v} - k^2 u \bar{v}) dx + \int_\Sigma M u \bar{v} d\sigma,$$

et L la forme linéaire continue définie par :

$$L(v) = \int_\Omega f \bar{v} dx.$$

Dans un premier temps, nous cherchons à écrire $a(u,v)$ sous la forme :

$$a(u,v) = ((I+T)u, v)_{H^1(\Omega)},$$

où T est un opérateur linéaire de $H^1(\Omega)$ dans $H^1(\Omega)$ à définir.
Commençons par mettre $a(u,v)$ sous la forme :

$$a(u,v) = \int_\Omega (\vec{\nabla} u . \vec{\nabla} \bar{v} + u \bar{v}) dx - (k^2+1) \int_\Omega u \bar{v} dx + \int_\Sigma M u \bar{v} d\sigma.$$

La forme sesquilinéaire définie par :

$$(u,v) \to -(k^2+1) \int_\Omega u \bar{v} dx$$

étant continue, le théorème de Riesz nous dit qu'il existe un opérateur linéaire T_1 continu de $H^1(\Omega)$ dans $H^1(\Omega)$ défini par :

$$(a) \qquad (T_1 u, v)_{H^1(\Omega)} = -(k^2+1) \int_\Omega u \bar{v} dx.$$

Soit W l'opérateur linéaire de $H^1(\Omega)$ dans $H^1(\Omega)$ défini par :

$$(Wu, v)_{H^1(\Omega)} = \int_\Sigma M u \bar{v} d\sigma,$$

ce qui s'écrit également :

$$(Wu, v)_{H^1(\Omega)} = -\frac{kR}{2\pi} \sum_{n \in \mathcal{Z}} \frac{H_n^{(1)'}(kR)}{H_n^{(1)}(kR)} \tau_n(u) \bar{\tau}_n(v),$$

avec $\tau_n(u) = \int_{-\pi}^{+\pi} u(R,\theta) e^{-in\theta} d\theta$.

Soit $z = kR$, nous savons que (cf. [Ab]) :

$$\frac{H_n^{(1)'}(z)}{H_n^{(1)}(z)} = \frac{H_{|n|}^{(1)'}(z)}{H_{|n|}^{(1)}(z)} = \frac{-H_{|n+1|}^{(1)}(z) + \frac{|n|}{z} H_{|n|}^{(1)}(z)}{H_{|n|}^{(1)}(z)}$$

$$= -\frac{H_{|n+1|}^{(1)}(z)}{H_{|n|}^{(1)}(z)} + \frac{|n|}{z}, \quad \forall n \in \mathcal{Z}.$$

Par ailleurs, $\frac{H_{|n+1|}^{(1)}(z)}{H_{|n|}^{(1)}(z)} \simeq 2\frac{|n|}{z}$, pour $|n| \to +\infty$.

D'où :

$$\frac{H_n^{(1)'}(z)}{H_n^{(1)}(z)} \simeq -\frac{|n|}{z}, \text{ pour } |n| \to +\infty, \quad \text{et} \quad z \text{ fixé dans } \mathcal{R}_*^+.$$

Ainsi,

(b) $\quad \forall z \in R_*^+, \exists N \in \mathcal{N}$ tel que : $\forall |n| \geq N, \frac{H_n^{(1)'}(z)}{H_n^{(1)}(z)} \simeq -\frac{|n|}{z} < 0.$

Nous pouvons donc écrire $(Wu, v)_{H^1(\Omega)}$ sous la forme :

$$(Wu, v)_{H^1(\Omega)} = (W_1, u, v)_{H^1(\Omega)} + (W_2 u, v)_{H^1(\Omega)},$$

avec :

$$.(c) \quad (W_1 u, v)_{H^1(\Omega)} = -\frac{kR}{2\pi} \sum_{|n|<N} \frac{H_n^{(1)'}(kR)}{H_n^{(1)}(kR)} \tau_n(u)\overline{\tau_n(v)},$$

$$.(d) \quad (W_2 u, v)_{H^1(\Omega)} = -\frac{kR}{2\pi} \sum_{|n|\geq N} \frac{H_n^{(1)'}(kR)}{H_n^{(1)}(kR)} \tau_n(u)\overline{\tau_n(v)},$$

Nous avons finalement :

$$a(u, v) = ((I + T)u, v)_{H^1(\Omega)},$$

avec :

$.T = T_1 + W_1 + W_2,$

.Les opérateurs T_1, W_1 et W_2 définis respectivement par

$(a), (c)$ et (d).

En utilisant le théorème de Riesz, nous pouvons également définir ϕ et écrire :

$$L(v) = (\phi, v)_{H^1(\Omega)}.$$

Ainsi, résoudre

$$a(u, v) = L(v)$$

revient à résoudre dans $H^1(\Omega)$

$$(I+T)u = \phi.$$

En vue d'utiliser l'alternative de Fredholm (cf. [Br]), nous allons désormais montrer que $(I+T)$ est la somme d'un opérateur coercif et d'un opérateur compact.
Pour cela, nous allons successivement montrer que :

.W_2 est un opérateur positif,

.T_1 est un opérateur compact,

.W_1 est un opérateur compact.

. D'après (b), nous savons que :

$$\forall |n| \geq N, \frac{H_n^{(1)'}(kR)}{H_n^{(1)}(kR)} < 0,$$

or

$$(W_2 u, u)_{H^1(\Omega)} = -\frac{kR}{2\pi} \sum_{|n|\geq N} \frac{H_n^{(1)'}(kR)}{H_n^{(1)}(kR)} |\tau_n(u)|^2,$$

donc $(W_2 u, u)_{H^1(\Omega)} > 0$.

. Soit (u_k) une suite bornée dans $H^1(\Omega)$. Nous allons montrer que l'on peut en extraire une sous-suite (u_{k_l}) telle que $(T_1 u_{k_l})$ converge dans $H^1(\Omega)$.

Nous savons que :

$$\|T_1 u_k - T_1 u\|^2_{H^1(\Omega)} = -(k^2+1)(u_k - u, T_1(u_k) - T_1 u)_{L^2(\Omega)}.$$

L'injection de $H^1(\Omega)$ dans $L^2(\Omega)$ étant compacte (théorème de Rellich), il existe une sous-suite (u_{k_p}) de (u_k) telle que :

$$u_{k_p} \to u \quad L^2(\Omega).$$

Comme (u_{k_p}) est aussi une suite bornée dans $H^1(\Omega)$ et que T_1 est un opérateur linéaire continu de $H^1(\Omega)$ dans $H^1(\Omega)$, $T_1(u_{k_p})$ reste bornée dans $H^1(\Omega)$.
En utilisant à nouveau le théorème de Rellich, on extrait une sous-suite (u_{k_l}) de (u_{k_p}) telle que $T_1(u_{k_l}) \to T_1(u) \quad L^2(\Omega)$.
Ainsi,

$$T_1 u_{k_l} \to T_1 u \quad H^1(\Omega).$$

. Soit (u_k) une suite bornée dans $H^1(\Omega)$.

Nous savons qu'il existe une sous-suite (u_{k_l}) de (u_k) telle que

$$u_{k_l} \rightharpoonup u \quad H^1(\Omega).$$

Comme l'opérateur trace est continu de $H^1(\Omega)$ dans $L^2(\Sigma)$, nous avons également :

$$u_{k_l}|_\Sigma \rightharpoonup u|_\Sigma \quad L^2(\Sigma),$$

et donc

$$\tau_n(u_{k_l}) \to \tau_n(u), \quad \forall |n| < N.$$

Montrons désormais que $\tau_n(W_1 u_{k_l}) \to \tau_n(W_1 u)$.

Dans un premier temps, montrons que W_1 est un opérateur continu de $H^1(\Omega)$ dans $H^1(\Omega)$.

Pour cela, il suffit de montrer qu'il existe une constante $C > 0$ telle que :

$$|(W_1 u, v)_{H^1(\Omega)}| \leq C \|u\|_{H^1(\Omega)} \|v\|_{H^1(\Omega)}, \quad \forall u, v \in H^1(\Omega).$$

Or,

$$\begin{aligned}
|(W_1 u, v)_{H^1(\Omega)}| &\leq C \sum_{|n|<N} \left| \frac{H_n^{(1)'}(kR)}{H_n^{(1)}(kR)} \right| |\tau_n(u)||\tau_n(v)|, \\
&\leq C \sum_{|n|<N} \left| \frac{H_n^{(1)'}(kR)}{H_n^{(1)}(kR)} \right| \|u|_\Sigma\|_{L^2(\Sigma)} \|v|_\Sigma\|_{L^2(\Sigma)}, \\
&\leq C \sum_{|n|<N} \left| \frac{H_n^{(1)'}(kR)}{H_n^{(1)}(kR)} \right| \|u\|_{H^1(\Omega)} \|v\|_{H^1(\Omega)}, \\
&\leq C(2N-1) \sup_{|n|<N} \left| \frac{H_n^{(1)'}(kR)}{H_n^{(1)}(kR)} \right| \|u\|_{H^1(\Omega)} \|v\|_{H^1(\Omega)}, \\
&\leq C \|u\|_{H^1(\Omega)} \|v\|_{H^1(\Omega)}.
\end{aligned}$$

Comme W_1 est continu de $H^1(\Omega)$ dans $H^1(\Omega)$, nous avons aussi :

$$W_1(u_{k_l}) \rightharpoonup W_1(u) \quad H^1(\Omega),$$

et donc

$$W_1(u_{k_l})|_\Sigma \rightharpoonup W_1(u)|_\Sigma \quad L^2(\Sigma).$$

Ainsi $\tau_n(W_1(u_{k_l})) \to \tau_n(W_1(u))$.

Par ailleurs,

$$\|W_1 u_{k_l} - W_1 u\|^2_{H^1(\Omega)} = -\frac{kR}{2\pi} \sum_{|n|<N} \frac{H_n^{(1)'}(kR)}{H_n^{(1)}(kR)} \tau_n(u_{k_l} - u)\overline{\tau_n(W_1 u_{k_l} - W_1 u)}.$$

Donc

$$W_1 u_{k_l} \to W_1 u \quad H^1(\Omega).$$

L'alternative de Fredholm nous dit alors que $(I+T)$ est inversible si et seulement si $(I+T)$ est injectif.

Nous allons en fait montrer que $(I+T)$ est injectif sur l'espace X défini par :

$$X = \{u \in V \ ; \ (u - \tilde{g}) \in V_o\}.$$

Soit $u_1, u_2 \in W$ deux solutions de (24).
En remplaçant u et v par $U = u_1 - u_2$ dans la formulation variationnelle (24), nous avons :

$$\int_\Omega (|\vec{\nabla} U|^2 - k^2|U|^2)dx + \int_\Sigma MU\bar{U}d\sigma = 0.$$

Par ailleurs,

$$\int_\Sigma MU\bar{U}d\sigma = Re((WU,U)_{H^1(\Omega)}) + iIm((WU,U)_{H^1(\Omega)}),$$

d'où :

$$.(e) \quad \int_\Omega (|\vec{\nabla} U|^2 - k^2|U|^2)dx + Re((WU,U)_{H^1(\Omega)}) = 0,$$
$$.(f) \quad Im((WU,U)_{H^1(\Omega)}) = 0.$$

En outre, nous pouvons montrer que l'équation (f) donne $U = 0$ sur Σ.

En effet, elle s'écrit :

$$\sum_{n\in\mathcal{Z}} Im\left(\frac{H_n^{(1)'}(kR)}{H_n^{(1)}(kR)}\right)|\tau_n(U)|^2 = 0.$$

De plus, nous savons que (cf. [Ab]) :

$$\frac{H_n^{(1)'}(z)}{H_n^{(1)}(z)} = -\frac{H_{n+1}^{(1)}(z)}{H_n^{(1)}(z)} + \frac{n}{z}, \quad \forall n \in \mathcal{Z}, \ \forall z \in \mathcal{R}_*.$$

Par ailleurs,

$$H_{n+1}^{(1)}(z) + H_{n-1}^{(1)}(z) = \frac{2n}{z} H_n^{(1)}(z), \quad \forall n \in \mathcal{Z},\ \forall z \in \mathcal{R}_*.$$

D'où :

$$Im\Big(\frac{H_{n+1}^{(1)}(z)}{H_n^{(1)}(z)}\Big) = \Big|\frac{H_{n-1}^{(1)}(z)}{H_n^{(1)}(z)}\Big|^2 Im\Big(\frac{H_n^{(1)}(z)}{H_{n-1}^{(1)}(z)}\Big), \quad \forall n \in \mathcal{Z},\ \forall z \in \mathcal{R}_*.$$

En faisant un raisonnement par récurrence, nous obtenons :

$$Im\Big(\frac{H_{n+1}^{(1)}(z)}{H_n^{(1)}(z)}\Big) > 0, \quad \forall n \in \mathcal{Z},\ \forall z \in \mathcal{R}_*,$$

et donc :

$$Im\Big(\frac{H_n^{(1)'}(z)}{H_n^{(1)}(z)}\Big) < 0, \quad \forall n \in \mathcal{Z},\ \forall z \in \mathcal{R}_*.$$

Ainsi, l'équation (f) ne peut être vérifiée que si $\tau_n(U) = 0,\ \forall n \in \mathcal{Z}$.
D'où :

$$U = 0 \quad \text{sur} \quad \Sigma.$$

En conséquence, l'équation (e) donne :

$$\int_\Omega (|\vec{\nabla} U|^2 - k^2 |U|^2) dx = 0.$$

La fonction U vérifie donc :

$$\begin{cases} \Delta U + k^2 U = 0 \quad \text{dans} \quad \Omega, \\ U = 0 \quad \text{sur} \quad \Gamma, \\ \dfrac{\partial U}{\partial n} = -MU = U = 0 \quad \text{sur} \quad \Sigma. \end{cases}$$

A l'aide d'un théorème de prolongement unique (cf. [H]), nous obtenons :

$$U = 0 \quad \text{dans} \quad \Omega.$$

Finalement, nous avons existence et unicité de la solution du problème (24). □

Remarque 5 :

Par des arguments classiques, on montre que la solution de (24) est solution de (22). □

Nous donnons finalement le résultat d'existence et d'unicité relatif au problème (22).

Théorème 2 :

Soit $f \in L^2(\Omega)$ à support compact et $g \in H^{\frac{1}{2}}(\Gamma)$. Alors, le problème (22) admet une unique solution $u \in H^1(\Omega)$. □

Démonstration :

Ce résultat découle de la proposition 2 et du théorème 1. □

4. Résolution numérique par une méthode d'éléments finis

Le but de cette section est de résoudre par une méthode d'éléments finis le problème initial. Comme nous l'avons vu, ce dernier est équivalent au problème tronqué (22) dont la formulation variationnelle est donnée par (23).

4.1. Discrétisation

Nous commençons par approcher $\bar{\Omega}$ par un domaine polygonal $\bar{\Omega}_h$, les sommets des frontières Γ_h et Σ_h de $\bar{\Omega}_h$ étant des points des frontières Γ et Σ de $\bar{\Omega}$.

Soit $\mathcal{T}_h$ une triangulation de $\bar{\Omega}_h$ à l'aide de triangles k de diamètre $h_k \leq h = \max_{k\in\mathcal{T}_h} h_k$.
Soit P_1 l'espace des polynômes de deux variables de degré inférieur ou égal à un.
On note :

. N le nombre de noeuds de $\mathcal{T}_h$,
. N_T le nombre de triangles de $\mathcal{T}_h$,
. N_o le nombre de noeuds de $\mathcal{T}_h$ appartenant à Γ,
. N_1 le nombre de noeuds de $\mathcal{T}_h$ n'appartenant pas à Γ,
. N_2 le nombre de noeuds de $\mathcal{T}_h$ appartenant à Σ,
. η l'ensemble des noeuds de $\mathcal{T}_h$,
. η_o l'ensemble des noeuds de $\mathcal{T}_h$ appartenant à Γ,
. η_1 l'ensemble des noeuds de $\mathcal{T}_h$ n'appartenant pas à Γ,
. η_2 l'ensemble des noeuds de $\mathcal{T}_h$ appartenant à Σ.

Nous voulons construire, à l'aide d'une méthode d'éléments finis $\mathcal{P}^1$ une solution approchée u_h.
Pour cela, nous introduisons deux espaces de dimension finie V_h et V_{oh} respectivement sous-espaces de V et V_o :

$$.V_h = \{v \in \mathcal{C}^o(\bar{\Omega}),\ ;\ \forall k \in \mathcal{T}_h\ v|_k \in \mathcal{P}^1\},$$
$$.V_{oh} = \{v \in \mathcal{C}^o(\bar{\Omega})\ ;\ v = 0 \quad \text{sur} \quad \Gamma\ ;\ \forall k \in \mathcal{T}_h\ v|_k \in \mathcal{P}^1\}.$$

La théorie de l'approximation variationnelle conduit alors à remplacer le problème (23) par le problème suivant :

$$(25) \qquad \begin{cases} \text{Trouver } u_h \in V_h \text{ tel que :} \\ \displaystyle\int_\Omega (\vec{\nabla} u_h . \vec{\nabla} \bar{v}_h - k^2 u_h \bar{v}_h) dx + \int_\Sigma M u_h \bar{v}_h d\sigma = \int_\Omega f \bar{v}_h dx, \qquad \forall v_h \in V_{oh}. \end{cases}$$

Notons φ_i la fonction de base associée à un noeud i de η et φ_{oi} la fonction de base associée à un noeud i de η_o.
u_h et v_h s'écrivent alors :

$$u_h = \sum_{i \in \eta_1} u_h(a_i)\varphi_i + \sum_{i \in \eta_o} u_h(a_i)\varphi_{oi},$$

$$v_h = \sum_{i \in \eta_1} v_h(a_i)\varphi_i,$$

a_i désignant le point physique au noeud i.

En reportant u_h et v_h dans la formulation de (25), nous obtenons finalement un système linéaire dans $\mathcal{C}^{N_1}$:

$$AU = F, \tag{26}$$

où :

. A est la matrice carrée, complexe, symétrique définie par

$$A = B + C,$$

avec :

. $B_{ij} = \int_\Omega (\vec{\nabla}\varphi_i.\vec{\nabla}\varphi_j - k^2\varphi_i\varphi_j)dx,$

. $C_{ij} = \int_\Sigma M_t\varphi_i\varphi_j d\sigma.$

. M_t l'opérateur DtN dans lequel on a tronqué la série à l'indice N_t,

. $U = (u_h(a_j))_{j \in \eta_1}$,

. $F = (F_{hi})_{i \in \eta_1}$,

avec :

$$F_{hi} = \int_\Omega f\varphi_i dx - \sum_{k \in \eta_o} u_h(a_k)(\int_\Omega (\vec{\nabla}\varphi_i.\vec{\nabla}\varphi_{ok} - k^2\varphi_i\varphi_{ok}dx).$$

Remarque 6 :

Il apparait que l'effet de la condition DtN dans le schéma éléments finis standard n'est que l'inclusion de C dans la matrice globale A. De plus, C_{ij} est nul dès que i ou j n'appartient pas à Σ.
Habituellement, comme deux noeuds n'appartenant pas au même segment n'interagissent pas l'un avec l'autre, on peut à l'aide d'un maillage structuré bien numéroté se ramener à une matrice bande et faciliter ainsi le stockage.
Ici, la condition DtN étant non locale, C_{ij} est non nul pour tous les noeuds i et j appartenant à Σ et cela peut détruire la structure bande de la matrice. Toutefois, il est possible de l'éviter en imposant au maillage une numérotation particulière. La matrice étant symétrique, on peut alors utiliser un stockage profil, toujours bien adapté aux structures bandes.

4.2. Calcul de la matrice provenant de la condition aux limites sur la frontière artificielle

Nous avons vu que la matrice A du système linéaire à résoudre s'écrit :

$$A = B + C,$$

avec :

. $B_{ij} = \int_\Omega (\vec{\nabla}\varphi_i . \vec{\nabla}\varphi_j - k^2 \varphi_i \varphi_j) dx, \qquad \forall i, j \in \eta_1,$

. $C_{ij} = \int_\Sigma M_t \varphi_i \varphi_j d\sigma, \qquad \forall i, j \in \eta_2.$

Nous proposons ici une méthode de calcul pour la matrice C, la matrice B se calculant de façon classique.

Pour tout $i, j \in \eta_2, C_{ij}$ s'écrit :

$$C_{ij} = -\frac{kR}{\pi} \sum_{n=0}^{N_t}{}' \frac{H_n^{(1)'}(kR)}{H_n^{(1)}(kR)} \int_0^{2\pi} \int_0^{2\pi} cos(n(\theta - \theta'))\varphi_i(R, \theta')\varphi_j(R, \theta) d\theta d\theta',$$

d'où

$$C_{ij} = -\frac{kR}{\pi} \sum_{n=0}^{N_t}{}' \frac{H_n^{(1)'}(kR)}{H_n^{(1)}(kR)} I_{ij}^n,$$

où

$$\begin{cases} I_{ij}^n = (\int_0^{2\pi} cos(n\theta)\varphi_i(R, \theta) d\theta)(\int_0^{2\pi} cos(n\theta)\varphi_j(R, \theta) d\theta) \\ \qquad + (\int_0^{2\pi} sin(n\theta)\varphi_i(R, \theta) d\theta)(\int_0^{2\pi} sin(n\theta)\varphi_j(R, \theta) d\theta). \end{cases}$$

Nous avons donc deux types d'intégrales à calculer :

$$. \int_0^{2\pi} cos(n\theta)\varphi_i(R, \theta) d\theta,$$

$$. \int_0^{2\pi} sin(n\theta)\varphi_i(R, \theta) d\theta.$$

Soit $(0, \vec{e}_x, \vec{e}_y)$ un repère orthonormé cartésien avec O le centre géométrique de l'obstacle. Pour tout noeud $i \in \eta_2$, a_i désigne le point physique correspondant et θ_i l'angle que forme le vecteur $\overrightarrow{Oa_i}$ avec l'axe des x positifs, $\theta_i \in [0, 2\pi]$.

Soit $\{1, 2, ..., N_2\}$ la numérotation des noeuds de Σ telle que :

$$0 \leq \theta_1 < \theta_2 \ldots < \theta_{N_2} < 2\pi.$$

On définit aussi

$$\begin{cases} a_o & = a_N, \\ \theta_o & = \theta_N - 2\pi, \end{cases}$$

$$\begin{cases} a_{N+1} = a_1, \\ \theta_{N+1} = \theta_1 + 2\pi. \end{cases}$$

En se plaçant dans la nouvelle numérotation, nous avons donc à calculer :

$$\int_{\theta_{i-1}}^{\theta_{i+1}} cos(n\theta)\varphi_i(\theta)d\theta, \quad \forall i \in \{1, 2, \ldots, N_2\},$$

$$\int_{\theta_{i-1}}^{\theta_{i+1}} sin(n\theta)\varphi_i(\theta)d\theta, \quad \forall i \in \{1, 2, \ldots, N_2\}.$$

Nous allons voir que pour $(\theta_{k+1} - \theta_{k-1})$ suffisamment petit, φ_i peut être parfaitement approchée par une fonction ψ_i affine par morceaux en θ, définie par :

$$\psi_i(\theta) = \begin{cases} \dfrac{\theta - \theta_{i-1}}{\theta_i - \theta_{i-1}} & \text{si} \quad \theta \in [\theta_{i-1}, \theta_i], \\ \dfrac{\theta_{i+1} - \theta}{\theta_{i+1} - \theta_i} & \text{si} \quad \theta \in [\theta_i, \theta_{i+1}]. \end{cases} \tag{27}$$

Sur $[\theta_{i-1}, \theta_{i+1}]$, φ_i est définie de façon exacte par :

$$\varphi_i(\theta) = \begin{cases} \dfrac{-cos(\theta_{i-1}) - \frac{cos\theta sin(\theta_{i-1} - \theta_i)}{sin(\theta - \theta_{i-1}) - sin(\theta - \theta_i)}}{cos(\theta_{i-1}) - cos(\theta_i)} & \text{si} \quad \theta \in [\theta_{i-1}, \theta_i], \\ \dfrac{-cos(\theta_i) - \frac{cos\theta sin(\theta_i - \theta_{i+1})}{sin(\theta - \theta_i) - sin(\theta - \theta_{i+1})}}{cos(\theta_i) - cos(\theta_{i+1})} & \text{si} \quad \theta \in [\theta_i, \theta_{i+1}]. \end{cases}$$

Les courbes des figures 5, 6 et 7 comparent les fonctions φ_i et ψ_i sur $[\theta_{i-1}; \theta_i]$ pour $(\theta_i - \theta_{i-1})$ de plus en plus petit.

En particulier, la figure 7 montre que pour $(\theta_i - \theta_{i-1}) \leq 0,5$ (ce qui est toujours le cas pour un maillage pas trop grossier !),ψ_i approche parfaitement la fonction de base φ_i.

On montre de même sur $[\theta_i; \theta_{i+1}]$, pour $(\theta_{i+1} - \theta_i) \leq 0,5$.

1 : φ_i

2 : ψ_i

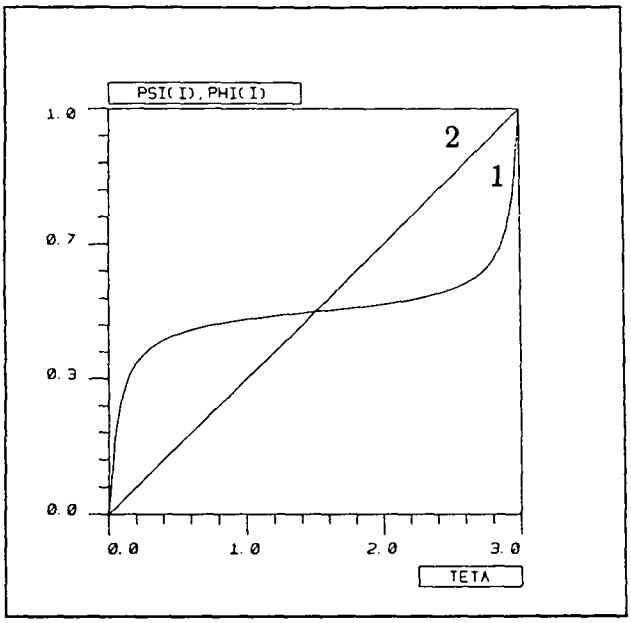


Figure 5 : Comparaison de φ_i et ψ_i sur $[\theta_{i-1}; \theta_i] = [0; 3]$

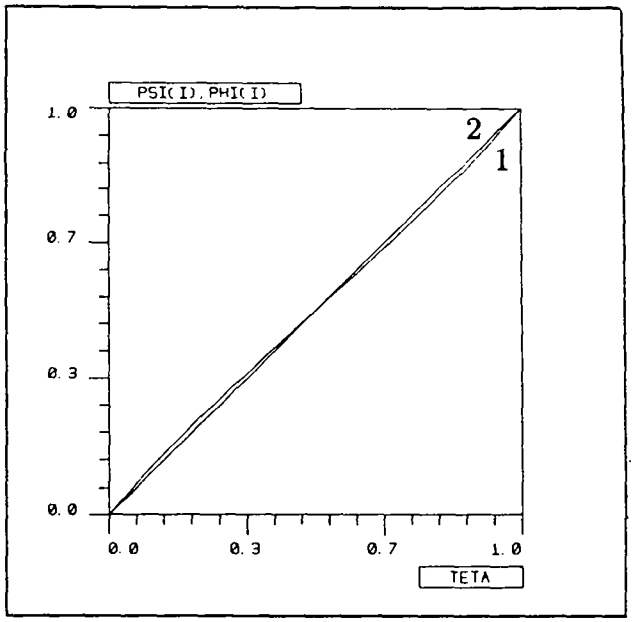


Figure 6 : Comparaison de φ_i et ψ_i sur $[\theta_{i-1}; \theta_i] = [0; 1]$

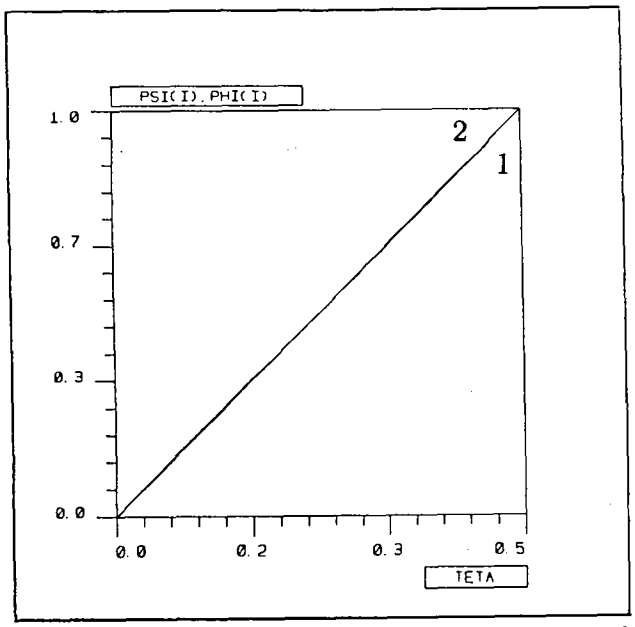


Figure 7 : Comparaison de φ_i et ψ_i sur $[\theta_{i-1}; \theta_i] = [0; 0,5]$

Nous sommes donc ramené à calculer :

$$\cdot\ P_i^n = \int_{\theta_{i-1}}^{\theta_{i+1}} cos(n\theta)\psi_i(\theta)d\theta,$$

$$\cdot\ Q_i^n = \int_{\theta_{i-1}}^{\theta_{i+1}} sin(n\theta)\psi_i(\theta)d\theta,$$

où ψ_i est la fonction affine par morceaux définie par (27).

Nous obtenons alors :

$$\cdot\ P_i^o = \frac{\theta_{i+1} - \theta_{i-1}}{2},$$

$$\cdot\ Q_i^o = 0,$$

$$\cdot\ P_i^n = \frac{1}{n^2}\left[\frac{cos(n\theta_i) - cos(n\theta_{i-1})}{\theta_i - \theta_{i-1}} - \frac{cos(n\theta_{i+1}) - cos(n\theta_i)}{\theta_{i+1} - \theta_i}\right], \quad \forall n \geq 1,$$

$$\cdot\ Q_i^n = \frac{1}{n^2}\left[\frac{sin(n\theta_i) - sin(n\theta_{i-1})}{\theta_i - \theta_{i-1}} - \frac{sin(n\theta_{i+1}) - sin(n\theta_i)}{\theta_{i+1} - \theta_i}\right], \quad \forall n \geq 1.$$

4.3. Résolution du système linéaire : analyse réelle

Par une méthode de prolongement classique, nous ramenons le système (26) à un système linéaire dans $\mathcal{C}^N$ que nous notons aussi :

$$Au = F. \tag{28}$$

Nous savons que A est une matrice carrée, complexe, creuse, symétrique, non hermitienne, à diagonale non dominante et de partie hermitienne non définie positive.

Nous allons présenter deux méthodes de résolution du système (28).

4.3.1. Réduction à un système réel

Nous choisissons ici de nous ramener à un système linéaire réel.
Pour cela, nous écrivons :

$$.A = Re(A) + iIm(A),$$
$$.u = Re(u) + iIm(u),$$
$$.F = Re(F) + iIm(F).$$

$Au = F$ nous donne alors un système linéaire dans $\mathcal{R}^{2N}$:

$$Mv = b, \tag{29}$$

avec :

. M la matrice carrée, réelle, creuse, non symétrique, à diagonale non dominante et de partie symétrique non définie positive définie par :

$$M = \begin{pmatrix} Re(A) & \vdots & -Im(A) \\ \dots & \dots & \dots \\ Im(A) & \vdots & Re(A) \end{pmatrix},$$

. $v = (Re(u_1), Re(u_2), \dots, Re(u_N), Im(u_1), Im(u_2), \dots, Im(u_N))^t$,

. $b = (Re(F_1), Re(F_2), \dots, Re(F_N), Im(F_1), Im(F_2), \dots, Im(F_N))^t$.

Pour le résoudre, nous pourrions utiliser une méthode directe, mais ce serait cher en stockage et en temps de factorisation. Nous nous sommes donc dirigés vers les méthodes itératives adaptées aux matrices creuses et en particulier vers les méthodes de gradient. La méthode du gradient conjugué classique n'est pas utilisable car la matrice n'est pas symétrique définie positive. Par ailleurs, on sait que les méthodes de l'équation normale et de l'erreur minimale dégradent fortement la convergence.

Nous avons donc choisi d'utiliser la méthode GMRES (Generalized Minimal RESidual method), également méthode de gradient, qui est toujours très robuste.

Dans l'annexe A, nous présentons la méthode GMRES, proposée par Y. Saad et M. Schultz, pour résoudre un système linéaire quelconque. Pour les détails, nous renvoyons le lecteur à leur article (cf. [SS]).

4.3.2. Différents préconditionneurs utilisés

Lorsque l'on utilise une méthode itérative, il est indispensable de bien préconditionner le système, à la fois pour la stabilité et pour la vitesse de convergence de la méthode.
Pour cela, on remplace la résolution de (29) $Mv = b$ par celle du système équivalent :

$$S^{-1}Mv = S^{-1}b,$$

la matrice S^{-1} devant être choisi pour que le conditionnement de $S^{-1}M$ soit beaucoup plus petit que le conditionnement de M.
En théorie, le meilleur choix est donc $S^{-1} = M^{-1}$.
En pratique, il faut trouver S^{-1} le plus proche de M^{-1}, sans que les calculs pour S^{-1} soient trop coûteux.
Nous avons analysé plusieurs choix pour la matrice S.

. Préconditionnement diagonal

Nous avons tout d'abord testé le choix le plus simple pour S, c'est-à-dire $S = S_1 = D$ où D est la diagonale de M.

. Préconditionnement de style relaxation

Le deuxième choix que nous avons fait est un préconditionnement de style relaxation car ce type de préconditionnement a l'avantage de n'être pas cher en temps de calcul et de ne pas nécessiter de stockage de matrice. Il dépend par ailleurs d'un paramètre ω, qu'il faut choisir entre 0 et 2 strictement.

Tout d'abord, nous écrivons M sous la forme :

$$M = L + D + U,$$

avec :

. D diagonale de M,

. L partie strictement triangulaire inférieure de M,

. U partie strictement triangulaire supérieure de M.

Nous approchons alors la matrice M par la matrice Q :

$$Q = (D + \omega L)D^{-1}(D + \omega U) = L_\omega U_\omega,$$

avec :

. $0 < \omega < 2$,

. $L_\omega = D + \omega L$,

. $U_\omega = D^{-1}(D + \omega U)$.

Finalement, nous remplaçons le système (29) par :

$$L_\omega^{-1} M U_\omega^{-1}(U_\omega U) = L_\omega^{-1} b,$$

ce qui revient à choisir $S = S_2 = D + \omega L$ comme préconditionneur du système (29).

Remarque 7 :

Nous pouvons noter que le calcul de $S^{-1}Av_k$, à chaque itération $k \in [1, k_m]$ de GMRES (k_m) (cf. algorithme 2 de l'annexe A), se réduit alors à la résolution d'un système triangulaire inférieur. □

. Préconditionnement "CROUT complet"

Nous avons également testé le préconditionneur

$$S = S_3 = \begin{pmatrix} B & 0 \\ 0 & B \end{pmatrix},$$

avec B factorisée par une méthode de CROUT complet.
On rappelle que la matrice B est définie par $B_{ij} = \int_\Omega (\vec{\nabla}\varphi_i . \vec{\nabla}\varphi_j - k^2 \varphi_i \varphi_j) dx$, pour $1 \leq i, j \leq N_1$.
Plus précisément, on écrit :

$$B = LDL^t,$$

avec :

. B matrice donnée par un stockage profil,

. D matrice diagonale,

. L matrice triangulaire inférieure à diagonale unité.

A chaque itération k de l'algorithme GMRES (k_m) on calcule $s_k = S^{-1}Av_k$ en résolvant $Ss_k = Av_k$, par les étapes suivantes :

. $Lt_k = Av_k$,

. $Du_k = t_k$,

. $L^t s_k = u_k$.

Remarque 8 :

Dans les applications industrielles, l'ordre souvent très grand de la matrice B rend le stockage de B et de L en mémoire centrale important voire même impossible. De plus, même si B est une matrice creuse, la matrice L est pleine. Par conséquent, on doit utiliser des unités de mémoire auxiliaire (disques ou bandes) entrainant des transferts de données exagérément coûteux, notamment en temps de calcul. Ainsi, il est intéressant de réaliser ce qu'on appelle une factorisation incomplète.
Dans celle-ci, on ne stocke pas la matrice L au complet, mais seulement un certain nombre d'éléments contenant l'information "représentative" de L. Ceci permet de plus de stocker la matrice B sous forme morse, ce qui est toujours très économique en place mémoire. □

. Préconditionnement "CROUT incomplet"

Notre dernier choix s'est donc porté sur le préconditionneur

$$S = S_4 = \begin{pmatrix} B & 0 \\ 0 & B \end{pmatrix},$$

avec B factorisée par une méthode de CROUT incomplet.

Le principe de la factorisation incomplète de CROUT est de chercher une matrice L triangulaire inférieure à diagonale unité et aussi creuse que possible, ainsi qu'une matrice diagonale D telles que LDL^t soit proche de B, dans un certain sens à définir. Par exemple, on pose $H = B - LDL^t$ et on demande que $\frac{\|H\|}{\|B\|}$ soit petit pour une norme matricielle donnée.

Souvent, on fixe a priori la structure creuse de L, c'est-à-dire l'ensemble

$$X_L = \{(i,j)\ ;\ 1 \leq j \leq i-1, 1 \leq i \leq N,\ l_{ij} \neq 0\}$$

On choisit en général d'imposer à L la même structure non nulle que celle de la partie triangulaire inférieure de B, afin de pouvoir stocker L et D dans la place mémoire réservée pour B.

Pour déterminer les coefficients de L, on suppose dans un premier temps D connu. On calcule alors, de proche en proche, les coefficients L_{ij} en imposant

$$H_{ij} = 0 \quad \text{pour} \quad (i,j) \in X_L.$$

On peut ensuite calculer les coefficients de D, également de proche en proche, par les équations

$$H_{k,k} = 0, \quad 1 \leq k \leq N.$$

4.3.3. Comparaisons entre les préconditionneurs

Dans ce paragraphe, nous comparons les préconditionneurs cités précédemment pour la résolution du système (29) par la méthode GMRES.

Pour cela, nous nous fixons le cas test suivant :

. L'obstacle est un disque de rayon $r_o = 0,5m$,

. L'onde incidente est plane, de longueur d'onde $\lambda = 1$ et se propage suivant l'axe $\overrightarrow{0X_1}$ dans le sens $x_1 > 0$ (cf. figure 8).
Elle s'écrit :

$$u^{inc}(x) = e^{ikx_1},$$

avec $k = \frac{2\pi}{\lambda} = 2\pi$,

. La frontière artificielle est placée à une distance 3λ de l'obstacle,

. Dans GMRES, la dimension de l'espace de Krylov est fixée à 50 et la précision à 10^{-6},

. La triangulation $\mathcal{T}_h$ comporte 4831 noeuds, 9410 triangles et $\frac{\lambda}{h} = 10$.

Nous illustrons les résultats obtenus à l'aide de courbes de convergence.

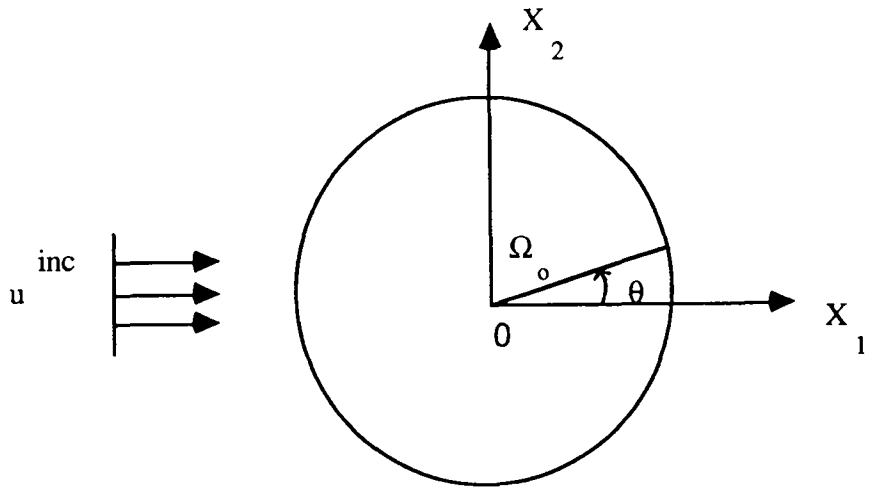


Figure 8 : cas de calcul

Remarque 9 :

(9a) Les courbes de convergence représentent la décroissance du résidu du système linéaire (plus exactement du résidu divisé par le résidu initial), en échelle logarithmique, en fonction du nombre d'itérations et du temps CPU.

(9b) Nous entendons par nombre d'itérations la somme des itérations effectuées dans la boucle de l'étape 1 de l'algorithme GMRES (k_m).

(9c) Le temps CPU est celui écoulé depuis la mise en route du programme. Sur toutes les figures présentées, nous l'indiquons en secondes.

(9d) A titre de référence pour les différents préconditionneurs testés, nous avons également tracé les courbes de convergence correspondant au cas où GMRES n'est pas préconditionné (i.e. : $S = S_o = I$).

(9e) Nous avons pris le terme source f (cf. (22)) égal à zéro.

(9f) Nous avons fixé à 20 l'indice de troncature de la série intervenant dans l'opérateur DtN.

(9g) Les tests ont été effectués sur un APOLLO DN10000. □

. Les courbes des figures 9 et 10 montrent la grande efficacité du préconditionneur "CROUT complet" contre l'inéfficacité totale des autres.
On voit même que non seulement les préconditionneurs S_1, S_2, S_4 n'apportent aucune amélioration au cas sans préconditionnement, mais pire encore ils entraînent plus rapidement "l'arrêt" de la convergence !

. Les courbes des figures 11 et 12 montrent l'influence du paramètre ω dans le cas du préconditionnement de style relaxation.
On constate là aussi que GMRES ne converge jamais.
Nous avons donné les résultats pour quatre ω différents mais beaucoup d'autres entre 0 et 2 ont été testés et aucun n'a donné satisfaction.

Conclusion :

Nous n'avons obtenu de bons résultats que pour un seul préconditionneur, ce qui nous a un peu déçu. Malgré tout, le bilan de ces tests n'est pas totalement négatif. En effet, les courbes des figures 9 et 10 illustrent que l'on a tout de même très nettement accéléré la convergence grâce au préconditionneur "Crout complet".

Toutefois, nous ne devons pas perdre de vue le fait que le coût de ce préconditionneur devient vite très élevé à mesure que la taille du système à résoudre augmente. C'est pourquoi, en vue d'applications industrielles, nous ne pouvons nous en contenter.

Cas réel

0 : sans préconditionnement (S_0)

1 : préconditionnement diagonal (S_1)

2 : préconditionnement de style relaxation (S_2); $\omega = 1,2$

3 : préconditionnement "CROUT complet" (S_3)

4 : préconditionnement "CROUT incomplet" (S_4)

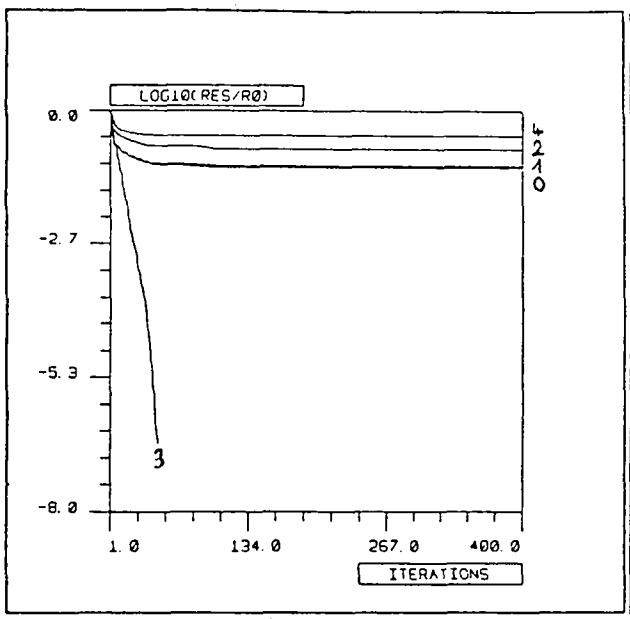


Figure 9

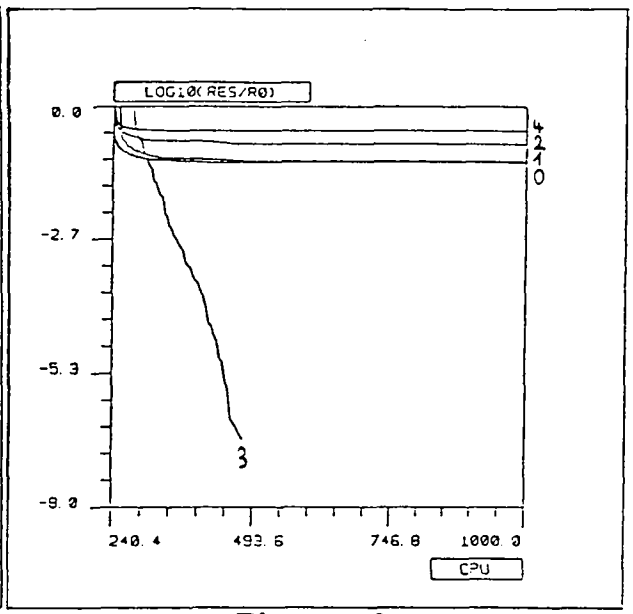


Figure 10

Convergence en fonction du préconditionnement

1 : $\omega = 0.2$

2 : $\omega = 0.9$

3 : $\omega = 1.5$

4 : $\omega = 1.7$

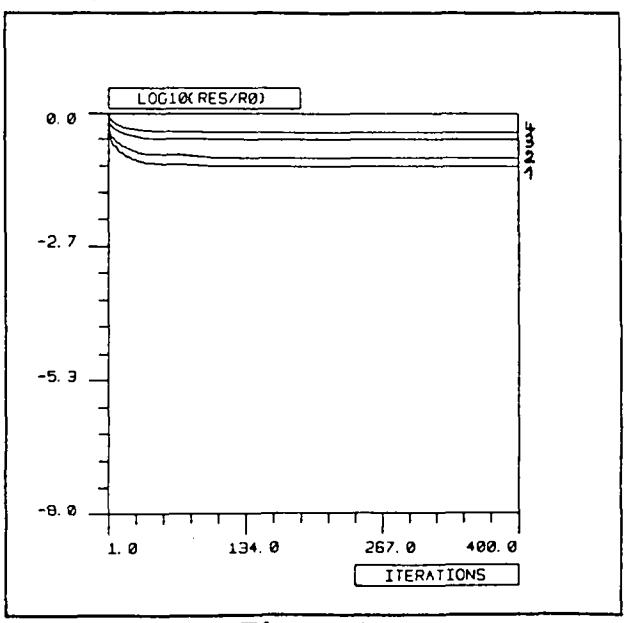


Figure 11

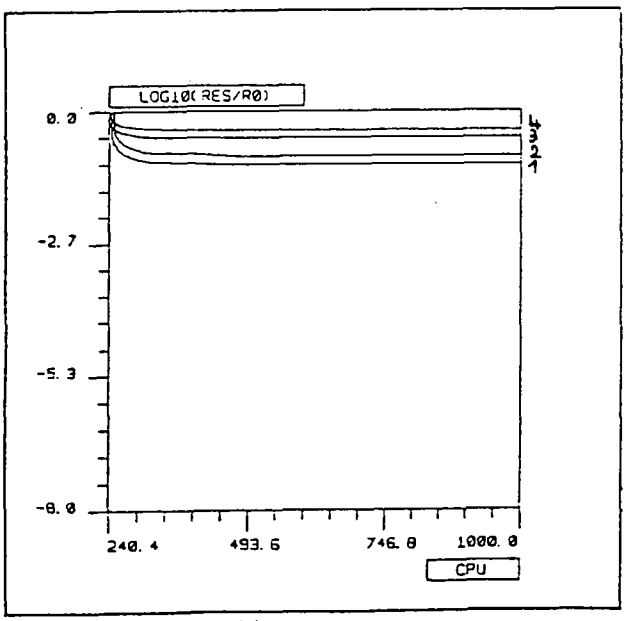


Figure 12

Convergence en fonction du paramètre ω de relaxation

4.4. Résolution du système linéaire : analyse complexe

Désormais, nous allons résoudre le système linéaire complexe d'ordre N :

$$Au = F, \tag{28}$$

sans le transformer en un système réel. Comme nous l'avons vu, A est une matrice carrée, creuse, symétrique, non hermitienne, à diagonale non dominante et de partie hermitienne non définie positive.

Pour les mêmes raisons que dans le cas réel, nous avons décidé de résoudre le système (28) par la méthode GMRES complexe que nous décrivons dans l'annexe A. Là aussi est comme toujours lorsque l'on utilise une méthode itérative, le problème crucial est de bien préconditionner le système. A ce sujet, nous avons décidé de tester les mêmes préconditionneurs que dans le cas réel mais appliqués à la matrice A du système (28).

4.4.1. Différents préconditionneurs utilisés

Les préconditionneurs choisis étant identiques à ceux utilisés dans l'analyse réelle, nous ne les présentons ici que très brièvement.

. Préconditionnement diagonal

$$S = S_1 = D,$$

avec D la diagonale de A.

. Préconditionnement de style relaxation

$$S = S_2 = D + \omega L,$$

avec

. D diagonale de A,
. L partie strictement triangulaire inférieure de A,
. $0 < \omega < 2$.

. Préconditionnement "CROUT complet"

$$S = S_3 = B,$$

avec B factorisée par une méthode de CROUT complet.

. **Préconditionnement "CROUT incomplet"**

$$S = S_4 = B,$$

avec B factorisée par une méthode de CROUT incomplet.

4.4.2. Comparaisons entre les préconditionneurs

Dans ce paragraphe, nous comparons l'efficacité des préconditionneurs S_1, S_2, S_3 et S_4 sur le même problème test que celui utilisé dans l'analyse réelle.
Là aussi, nous illustrons les résultats obtenus à l'aide de courbes de convergence.

Remarque 10 :

La remarque 9 du §4.3.3. s'applique également ici. □

. Les courbes des figures 13 et 14 montrent que, pour ce problème test, GMRES complexe converge quels que soient les préconditionneurs utilisés et même lorsqu'il n'est pas préconditionné.
On constate tout de même que l'efficacité de ces préconditionneurs est assez inégale et que seuls S_3 et S_4 améliorent nettement la convergence.

. Les courbes des figures 15 et 16 montrent l'influence du paramètre ω dans le cas du préconditionnement de style relaxation sur la convergence de GMRES complexe.
Nous avons par ailleurs effectué d'autres tests sur ω et avons constaté que l'intervalle $[0,5; 0,7]$ est le plus efficace même si les résultats restent assez semblables dans $[0,2; 1,2]$.
Mais surtout, il est ressorti que le préconditionneur S_2 ne peut en aucun cas "rivaliser" avec les préconditionneurs S_3 et S_4.

Conclusion :

Contrairement au cas réel, nous avons toujours obtenu la convergence de GMRES quel que soit le préconditionneur testé. **Apparemment** le préconditionneur "CROUT complet" est le plus efficace car il fait converger GMRES le plus vite. **Toutefois**, il ne faut pas oublier qu'il ne le sera rapidement plus lorsque la taille du système à résoudre deviendra grande. En effet, on sera alors confronté au problème du stockage de la matrice en mémoire centrale, problème nous obligeant à des transferts de données coûteux entre la mémoire centrale et une unité de mémoire auxiliaire.

Par contre, on ne rencontre pas ce problème lors de la factorisation incomplète de CROUT LDL^t car on peut utiliser le stockage morse et on ne stocke qu'une partie de la matrice L.

Ainsi, c'est le préconditionneur "CROUT incomplet" qui nous a paru le plus intéressant car nous savons qu'il restera très performant lors de la résolution de grands systèmes.

Cas complexe

0 : sans préconditionnement (S_n)
1 : préconditionnement diagonal (S_1)
2 : préconditionnement de style relaxation (S_2). $\omega = 1.2$
3 : préconditionnement "CROUT complet" (S_3)
4 : préconditionnement "CROUT incomplet" (S_4)

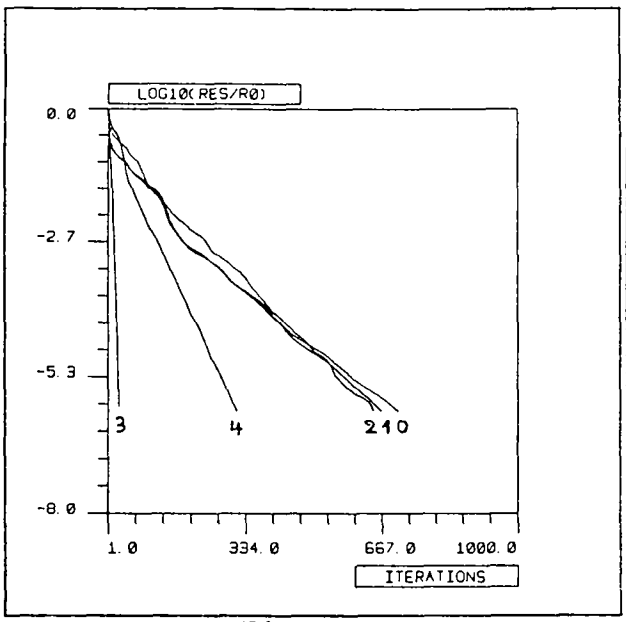


Figure 13

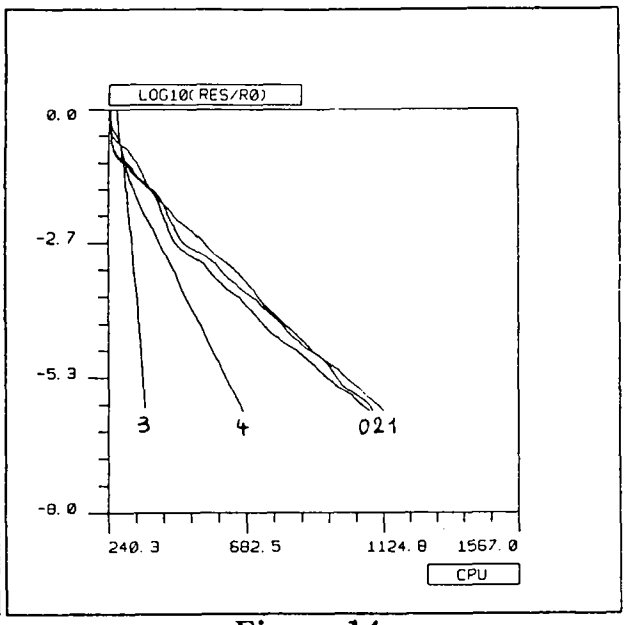


Figure 14

Convergence en fonction du préconditionnement

1 : $\omega = 0.2$
2 : $\omega = 0.9$
3 : $\omega = 1.5$
4 : $\omega = 1.7$

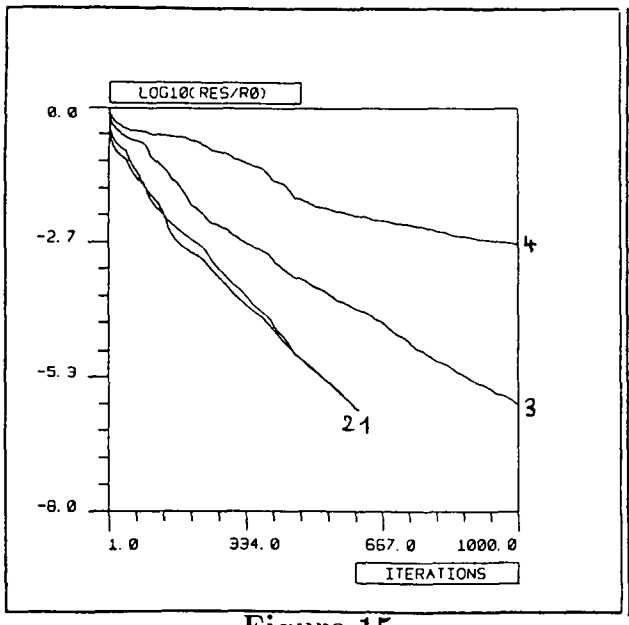


Figure 15

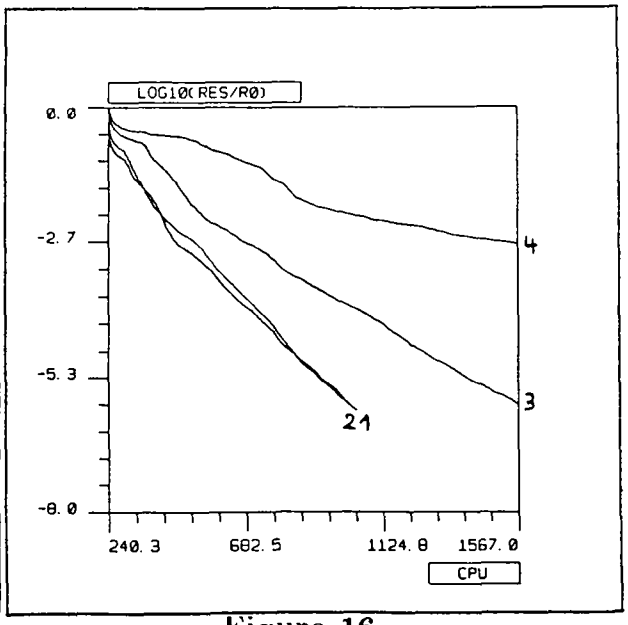


Figure 16

Convergence en fonction du paramètre ω de relaxation

5. Résultats numériques

Remarque préliminaire :

Tous les résultats présentés dans cette section sont relatifs à la résolution du système linéaire complexe (28), par la méthode GMRES utilisant le préconditionneur "CROUT incomplet".

5.1. Influence de la dimension de l'espace de Krylov

Nous étudions ici l'influence de la dimension k_m imposée à l'espace de Krylov sur la vitesse de convergence de GMRES.
Le cas test utilisé est le même que celui décrit au §4.3.3.

Nous illustrons les résultats obtenus à l'aide de la courbe de la figure 17 représentant le temps CPU nécessaire pour converger en fonction de la dimension k_m que l'on impose à l'espace de Krylov.

Il apparaît bien que le choix de k_m est très important : par exemple, la convergence est près de 30% plus rapide pour $k_m = 37$ que pour $k_m = 5$!
Même si 37 semble être le k_m optimal, il est clair que l'on peut choisir k_m dans l'intervalle $[25, 50]$ et conserver une rapidité de convergence très semblable.

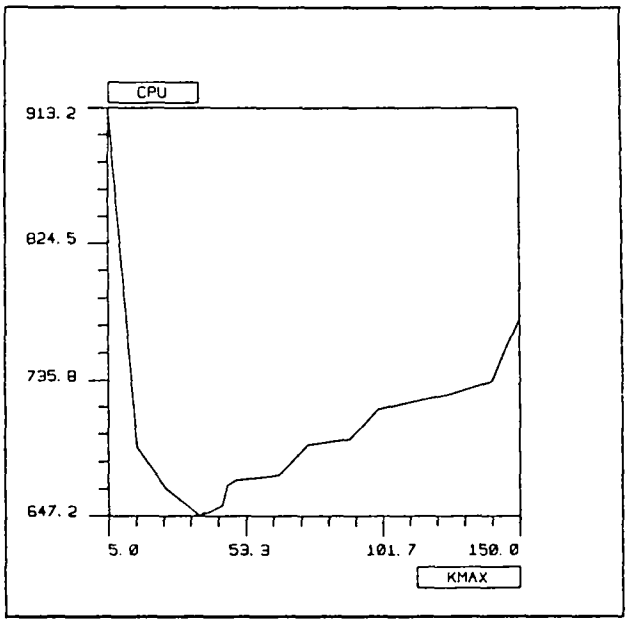


Figure 17 : Influence de la dimension de l'espace de Krylov.

5.2. Influence de la discrétisation

Dans cette sous-section, nous donnons une série de résultats obtenus lorsque l'on utilise successivement une discrétisation en h, $\frac{h}{2}$ et $\frac{h}{4}$.

Le cas test considéré est le suivant :

. L'obstacle est un disque de rayon $r_o = 0,5m$,

. L'onde incidente est plane, de longueur d'onde $\lambda = 1$ et se propage suivant l'axe $\overrightarrow{0X_1}$ dans le sens $x_1 > 0$ (cf. figure 8).

Elle s'écrit :

$$u^{inc}(x) = e^{ikx_1},$$

avec $k = \frac{2\pi}{\lambda} = 2\pi$,

. La frontière artificielle est placée à une distance λ de l'obstacle,

. Dans GMRES, la dimension de l'espace de Krylov est fixée à 50 et la précision à 10^{-6},

. Les trois triangulations utilisées ont les caractéristiques suivantes :

. $\mathcal{T}_h$: 850 noeuds et 1574 triangles, $\frac{\lambda}{h} = 10$,

. $\frac{\mathcal{T}_h}{2}$: 3274 noeuds et 6296 triangles, $\frac{\lambda}{h} = 20$,

. $\frac{\mathcal{T}_h}{4}$: 12844 noeuds et 25184 triangles, $\frac{\lambda}{h} = 40$,

. Les courbes des figures 18 et 19 montrent l'évolution de la convergence lorsque le paramètre h est tel que $\frac{\lambda}{h} = 10$, $\frac{\lambda}{h} = 20$ et $\frac{\lambda}{h} = 40$.

Remarque 11 :

Les remarques (9a), (9b), (9c), (9e), (9f) et (9g) du §4.3.3. s'appliquent également ici. □

. La courbe de la figure 20 représente l'erreur en norme L^2 entre la partie réelle de la solution calculée et de la solution analytique en fonction du pas h de discrétisation. Les échelles étant logarithmiques, il est bien rassurant de trouver une droite qui est de pente 2 !

Remarque 12 :

. La solution analytique de notre problème test est la fonction à valeurs complexes définie par :

$$p(r,\theta) = -\sum_{n\in\mathbb{Z}} i^{|n|} J_{|n|}(kr_o) e^{in\theta} \frac{H_n^{(1)}(kr)}{H_n^{(1)}(kr_o)}, \qquad \forall r \in [r_o, +\infty[, \forall \theta \in [0, 2\pi],$$

où

. J_n désigne la fonction de Bessel d'ordre n du premier type,

. $H_n^{(1)}$ désigne la fonction de Hankel d'ordre n du premier type.

. Tous les résultats de comparaison entre la solution calculée et la solution analytique sont donnés pour la partie réelle de l'onde diffractée. □

. Le tableau 1 donne les erreurs relatives en norme L^2, obtenues sur les trois triangulations $\mathcal{T}_h$, $\frac{\mathcal{T}_h}{2}$ et $\frac{\mathcal{T}_h}{4}$.

. Les courbes de la figure 21 représentent les coupes de la solution calculée et de la solution analytique suivant l'axe $\overrightarrow{0X_1}$ entre Γ et Σ.
Ces coupes sont présentées pour chaque triangulations $\mathcal{T}_h$, $\frac{\mathcal{T}_h}{2}$ et $\frac{\mathcal{T}_h}{4}$.

. La figure 22 visualise la solution calculée et la solution analytique lorsque le domaine de calcul est discrétisé avec $\frac{\mathcal{T}_h}{2}$.

Conclusion :

Ces résultats numériques confirment l'influence de la discrétisation caractérisée par le nombre de points par longueur d'onde $\frac{\lambda}{h}$, sur l'erreur obtenue.
Ils montrent en particulier que la condition DtN employée sur Σ donne d'excellents résultats pour une frontière artificielle très proche de l'obstacle (ici à une distance λ), et ceci même pour $\frac{\lambda}{h} = 10$.

1 : $\frac{\lambda}{h} = 10$
2 : $\frac{\lambda}{h} = 20$
3 : $\frac{\lambda}{h} = 40$

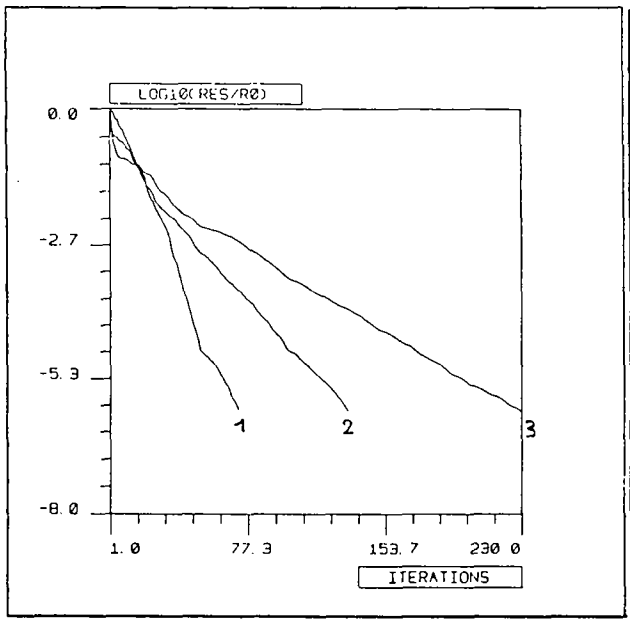


Figure 18

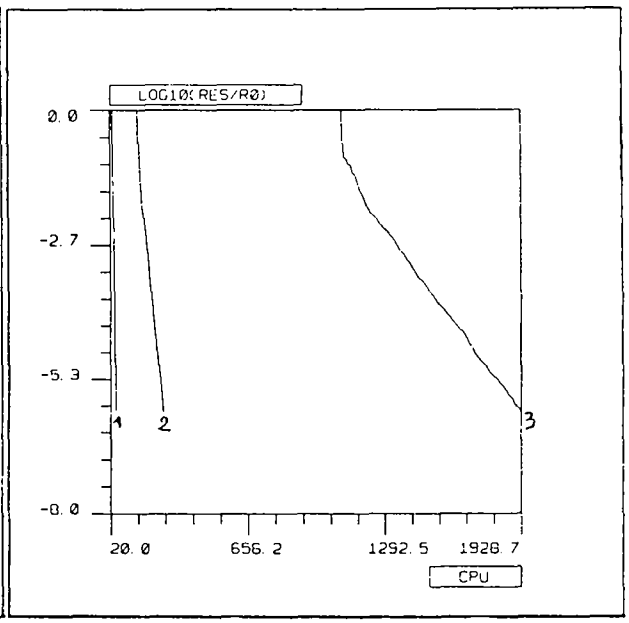


Figure 19

Convergence en fonction de la discrétisation

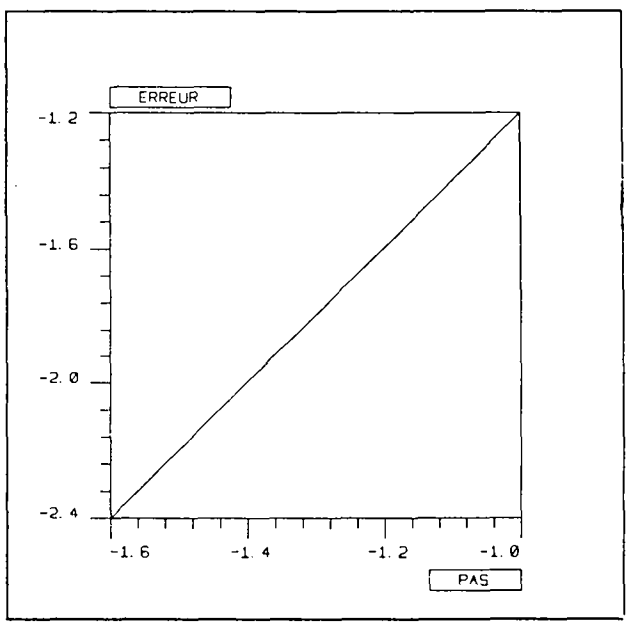


Figure 20

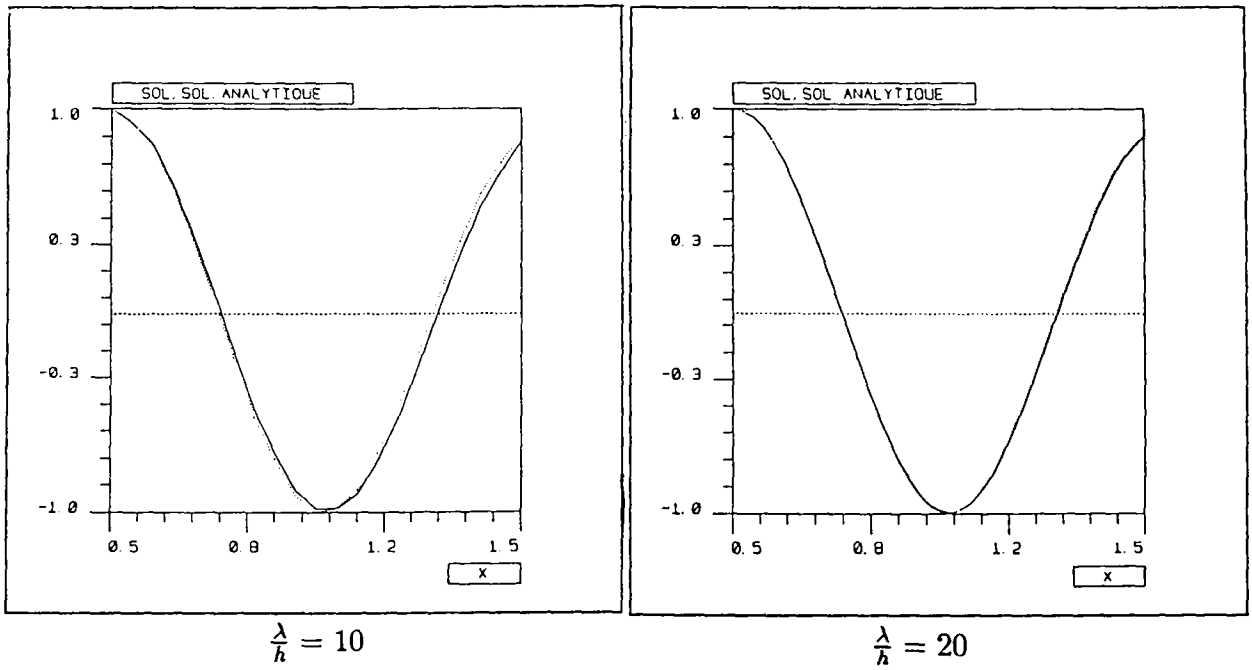
$\frac{\lambda}{h}$	erreur relative
10	5,15%
20	1,33%
40	0,34%

Tableau 1

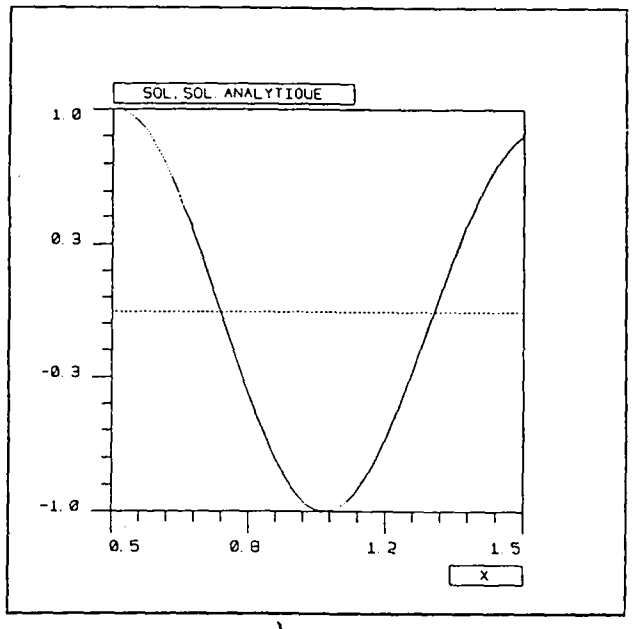
Influence de la discrétisation

——— : solution calculée

..... : solution analytique

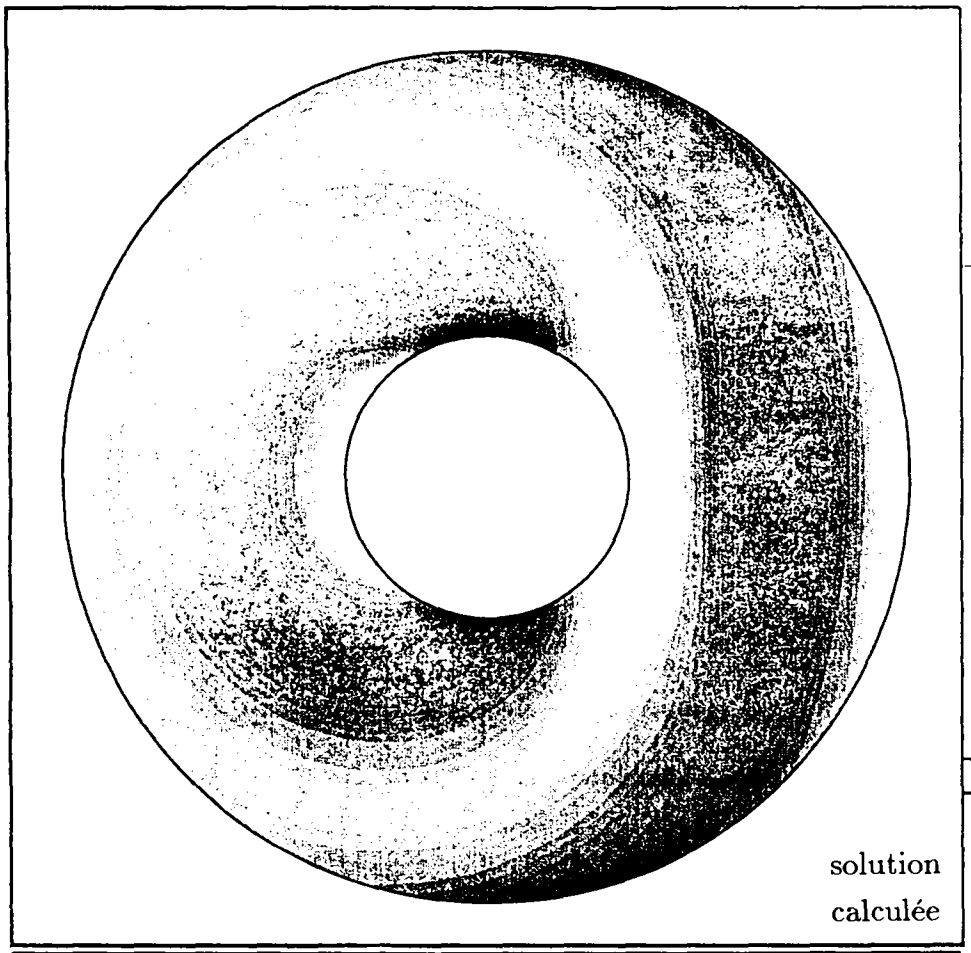


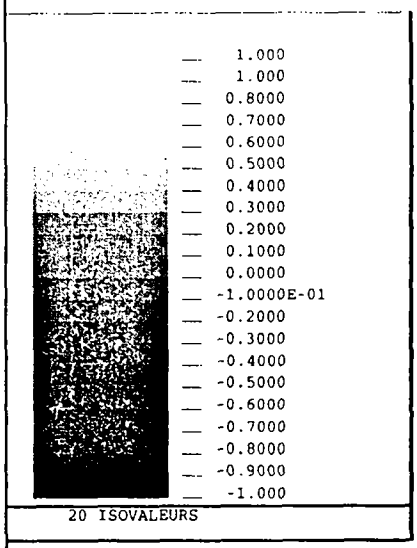
$\frac{\lambda}{h} = 10$ $\frac{\lambda}{h} = 20$

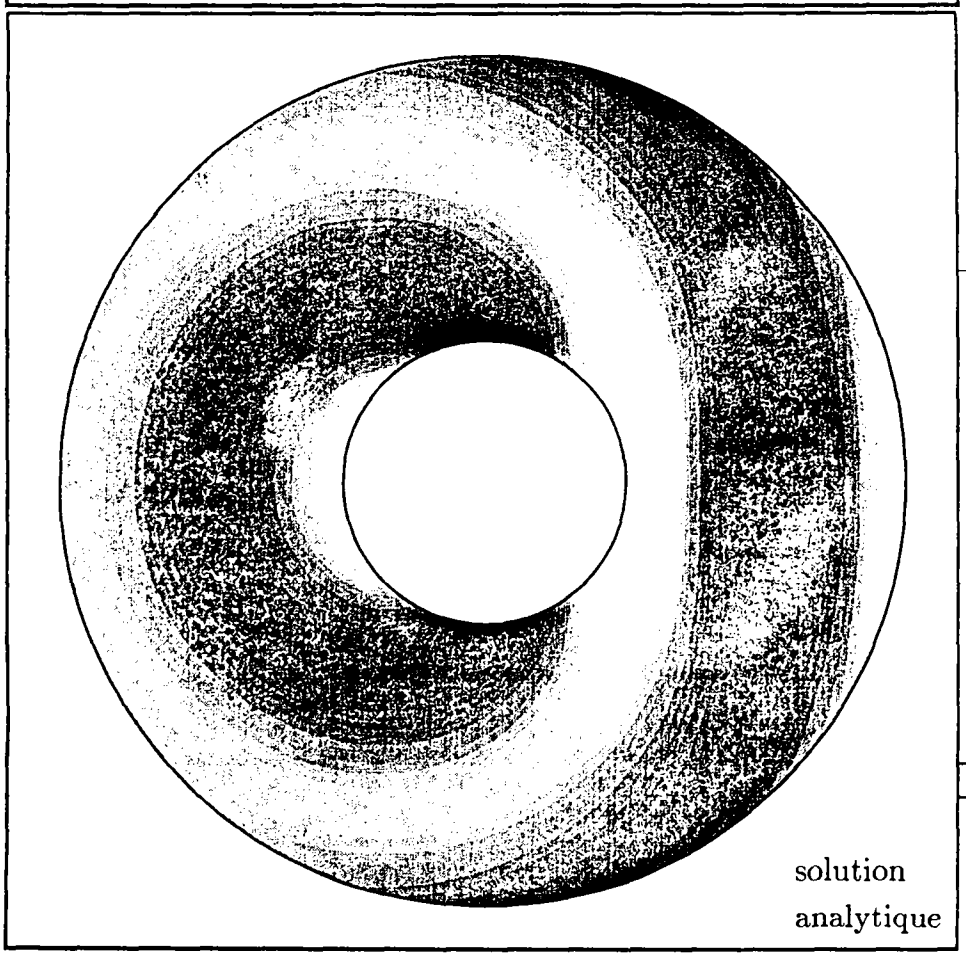


$\frac{\lambda}{h} = 40$

Figure 21: Coupes en fonction de la discrétisation







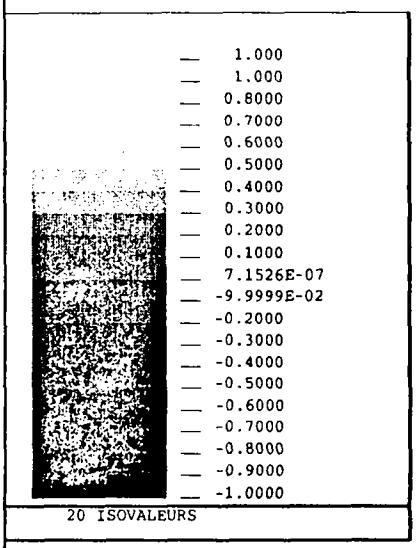


Figure 22 : Ondes diffractées, $\frac{\lambda}{h} = 20$

5.3. Calculs autour d'un profil d'avion

Nous présentons ici quelques résultats obtenus à partir d'un profil d'avion, sur des discrétisations plus fines que précédemment.
Le cas test n° 1 est le suivant :

. L'obstacle est un profil d'avion de longueur égale à 20 m et de hauteur maximale égale à 4,15 m,

. L'onde incidente est plane, de longueur d'onde $\lambda = 5$ et se propage suivant l'axe $\overrightarrow{OX_2}$ dans le sens $x_2 > 0$ (cf. figure 23).

Elle s'écrit :

$$u^{inc}(x) = e^{ikx_2},$$

avec $k = \frac{2\pi}{\lambda} \simeq 1,26$,

. La frontière artificielle est le cercle de centre 0 et de rayon $R = 20m$.

. Dans GMRES, la dimension de l'espace de Krylov est fixée à 500 et la précision à 10^{-6},

. La triangulation $\mathcal{T}_h$ comporte 32350 noeuds et 63900 triangles.

Le cas test n° 2 ne diffère du premier que par deux points :

. La frontière artificielle est le cercle de centre 0 et de rayon $R = 12m$,

. La triangulation $\mathcal{T}_h$ comporte 35666 noeuds et 70332 triangles.

Le cas test n° 3 reprend le précédent, avec une longueur d'onde $\lambda = 3$.

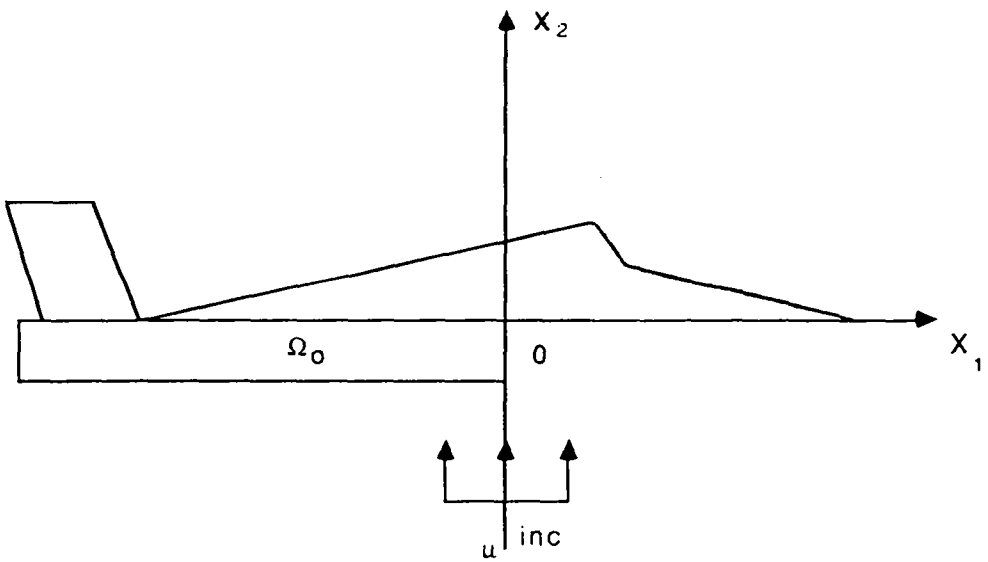


Figure 23 : cas de calcul.

. Le tableau 2 donne, pour chaque cas test, le nombre d'itérations et le temps CPU nécessaires à la convergence de GMRES.

	Itérations	CPU
Cas test n° 1	734	7763
Cas test n° 2	318	5074
Cas test n° 3	927	10485

Tableau 2

Remarque 13 :

. Les remarques (9b), (9c), (9e) et (9f) du §4.3.3 sont également valables dans cette sous-section.

. Les tests ont été effectués sur le CRAY2. □

. Les figures 24 et 25 visualisent respectivement la solution calculée sur le cas test n° 1 et sur le cas test n° 2.

. Les figures 26 et 27 représentent respectivement la solution et la solution totale, calculées sur le cas test n° 3.

Remarque 14 :

Nous entendons par solution calculée la partie réelle de l'onde diffractée par l'obstacle. De même, la solution totale est la somme de la partie réelle de l'onde incidente et de la partie réelle de l'onde diffractée. □

Conclusion :

Ces résultats confirment en particulier ce que nous avons déjà constaté, à savoir qu'à partir d'une distance très petite entre l'obstacle et la frontière artificielle, la solution obtenue ne change quasiment pas.

cas test n° 1

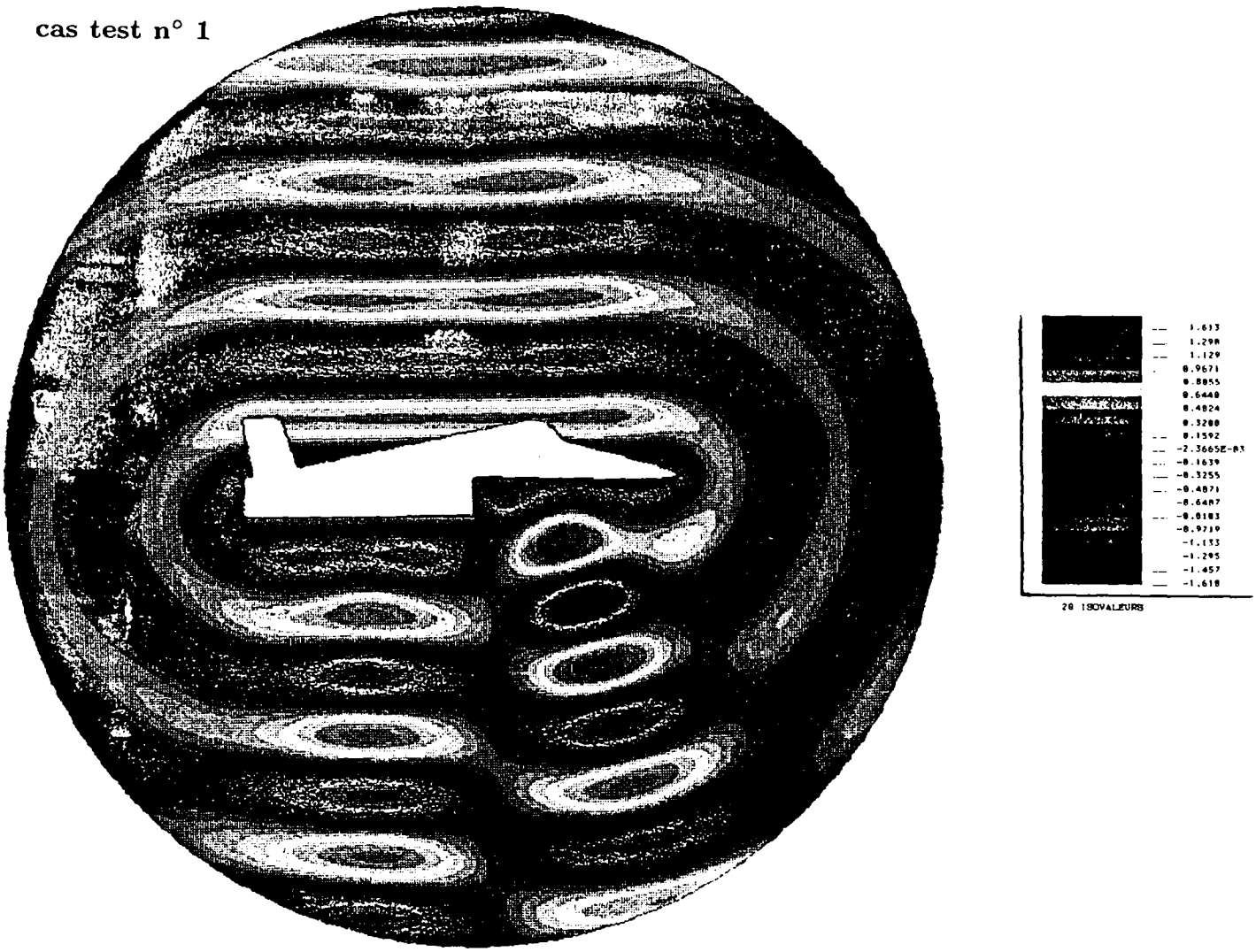


Figure 24 : Onde diffractée, $\lambda = 5, R = 20$

cas test n° 2

Figure 25 : Onde diffractée, $\lambda = 5, R = 12$

cas test n° 3

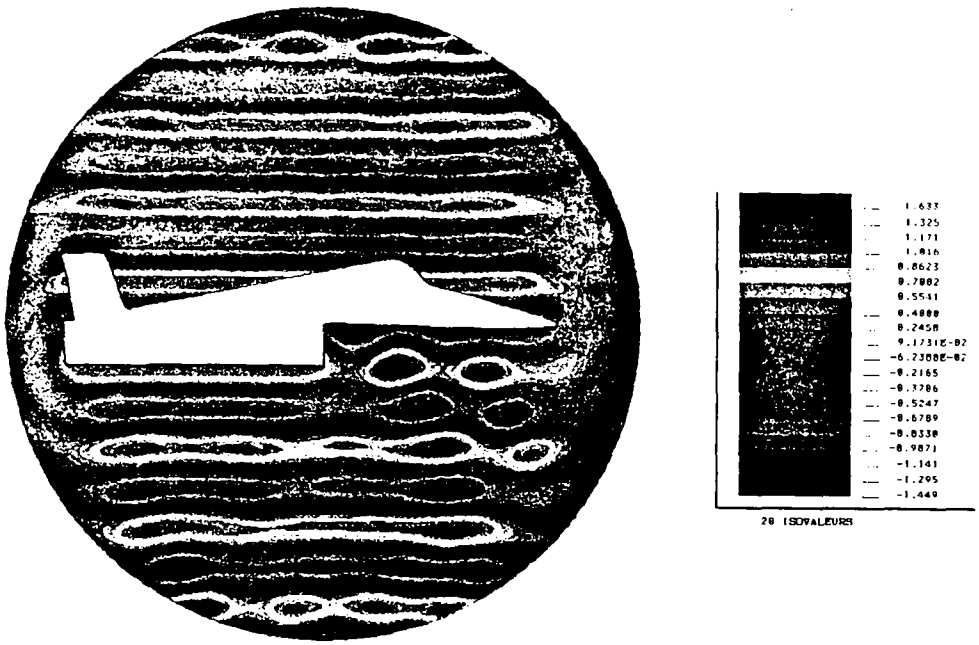


Figure 26 : Onde diffractée, $\lambda = 3, R = 12$

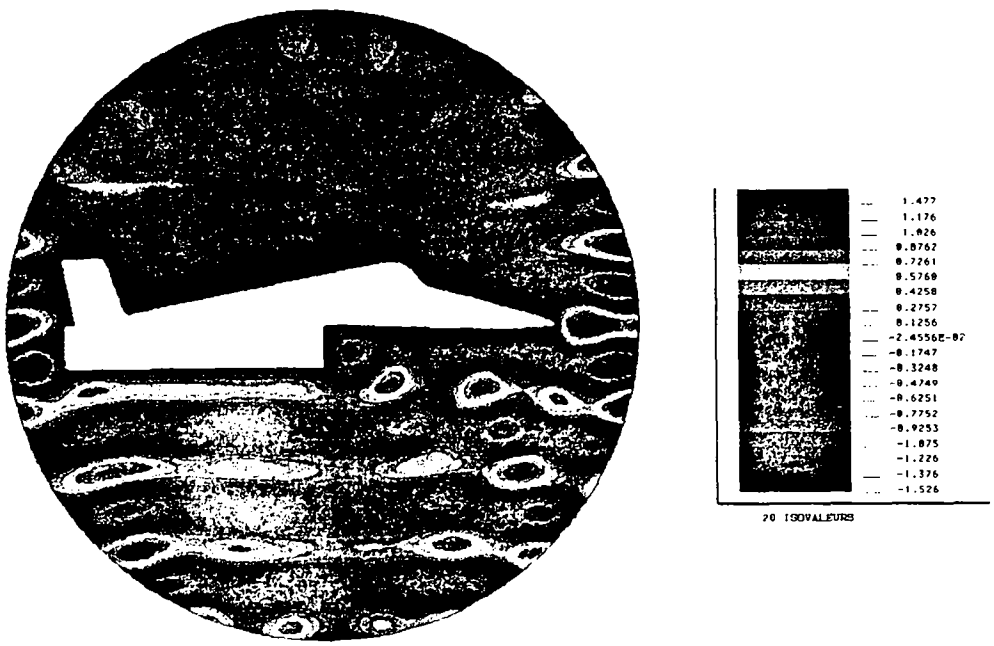


Figure 27 : Onde diffractée, $\lambda = 3, R = 12$

6. Conclusion

Nous avons présenté une méthode numérique pour résoudre l'équation de Helmholtz issue de problèmes bidimensionnels de diffraction d'une onde électromagnétique.

Nous avons analysé deux façons de résoudre le système linéaire obtenu (une résolution dans $\mathcal{R}^{2N}$ et une résolution dans $\mathcal{C}^N$, N désignant le nombre de noeuds du maillage éléments finis utilisé) et nous avons montré que la deuxième était la plus efficace.

Nous avons également décrit la méthode de résolution utilisée dans les deux cas, à savoir la méthode GMRES linéaire. Celle-ci est une méthode de gradient qui s'avère en pratique très robuste pour résoudre des systèmes linéaires sans propriétés particulières.

Nous avons présenté plusieurs façons de préconditionner cette méthode appliquée au système réel et au système complexe.

Dans le premier cas, nous avons constaté que seul le préconditionneur "CROUT complet" conduit à faire converger GMRES en un temps raisonnable.
Dans le cas complexe, nous avons noté une amélioration spectaculaire des performances dues aux préconditionneurs de GMRES. Par ailleurs, c'est le préconditionneur "CROUT incomplet" qui s'est avéré le plus intéressant, car il donne une convergence très rapide de GMRES et n'occasionne pas de problèmes de place mémoire.

Les différentes expériences numériques que nous avons fait ont finalement permis de valider la méthode et le choix du préconditionneur. Elles ont également confirmé ce que nous étions en droit d'espérer, à savoir que la condition DtN donne d'excellents résultats même pour une frontière artificielle très proche de l'obstacle.

Remerciements

Je tiens, en premier lieu, à remercier Monsieur Olivier Pironneau pour m'avoir guidée et conseillée afin de mener à bon terme le présent travail.

Je souhaite également remercier tous ceux qui, de diverses façons, m'ont aidées dans mes travaux. En particulier, j'adresse mes profonds remerciements à Madame Marie-Odile Bristeau qui m'a offert constamment de précieux conseils avec une incroyable gentillesse et disponibilité, ainsi qu'à Monsieur Mikaël Balabane pour son aide amicale et pour les fructueuses discussions que j'ai eues avec lui.

Enfin, je souhaite adresser mes sincères remerciements à Madame Christiane Demars qui a assuré le travail de dactylographie avec beaucoup de patience et de gentillesse.

Annexe A. La méthode GMRES linéaire

Remarque préliminaire :

Pour plus de clarté, nous exposons la méthode dans $\mathcal{R}^N$ mais le principe est le même dans $\mathcal{C}^N$. □

Nous désirons résoudre le système linéaire dans $\mathcal{R}^N$:

$$Ax = b,$$

où :

. A est une matrice $N \times N$ inversible,

. x l'inconnue du système.

L'idée de la méthode est d'écrire l'inconnue x sous la forme $x_o + z$, où x_o est une première estimation de x et où z la nouvelle inconnue appartient à un certain sous-espace affine de $\mathcal{R}^N$, noté K_k et appelé sous-espace de Krylov.
La méthode consiste alors à minimiser, à chaque étape, la norme euclidienne du résidu $b - Ax$ dans K_k.

Ainsi, GMRES cherche $z_k \in K_k$ tel que :

$$\begin{aligned} \|r_k\|_2 = \|b - A(x_o + z_k)\|_2 &= \min_{z \in K_k} \|b - A(x_o + z)\|_2 \\ &= \min_{z \in K_k} \|r_o - Az)\|_2, \end{aligned}$$

K_k étant le sous-espace de Krylov de dimension k associé à la matrice A et au résidu initial $r_o = b - Ax_o$, c'est-à-dire l'espace vectoriel engendré par les vecteurs $r_o, Ar_o, \ldots, A^{k-1}r_o$.

La mise en oeuvre de la méthode commence par la construction, moyennant l'algorithme d'Arnoldi (qui utilise un processus d'orthogonalisation de Gram-Schmidt), d'une base orthonormée de K_k par rapport à la norme euclidienne dans $\mathcal{R}^N$.
Plus précisément, l'algorithme d'Arnoldi génère k vecteurs $v_1, \ldots, v_k$ qui forment une base du sous-espace de Krylov K_k associé au vecteur initial v_1 et à la matrice A, i.e. :

$$< v_1, v_2, \ldots, v_k > = < v_1, Av_1, \ldots, A^{k-1}v_1 > = K_k,$$

où la notation $< v_1, v_2, \ldots, v_k >$ désigne l'espace vectoriel engendré par $v_1, v_2, \ldots, v_k$.

On construit alors une matrice $N \times kV_k$ dont les colonnes sont les vecteurs de base $v_1, v_2, \ldots, v_k$, et une matrice $(k+1) \times k$ $\vec{H}_k$ qui vérifie :

$$AV_k = V_{k+1}\vec{H}_k. \tag{30}$$

On veut donc résoudre le problème :

$$\begin{cases} \text{Trouver } z_k \in K_k \text{ tel que :} \\ \|b - A(x_o + z_k)\|_2 = \min\limits_{z \in K} \|b - A(x_o + z)\|_2 = \min\limits_{z \in K_k} \|r_o - Az\|_2. \end{cases}$$

En écrivant $z = V_k y$, cela revient à minimiser la fonctionnelle

$$J(y) = \|\beta v_1 - AV_k y\|_2,$$

où $\beta = \|r_o\|_2$ et $v_1 = \frac{r_o}{\beta}$.

D'après l'égalité (30), on peut écrire :

$$J(y) = \|\beta v_1 - V_{k+1}\vec{H}_k y\|_2.$$

En introduisant le vecteur $e_1 \in \mathcal{R}^{k+1}$:

$$e_1 = (1, 0, 0, \ldots, O)^t,$$

on a de plus $v_1 = V_{k+1}e_1$.

Comme V_{k+1} est une matrice orthogonale, on obtient finalement

$$J(y) = \|\beta e_1 - \vec{H}_k y\|_2. \tag{31}$$

Ainsi, pour tout k fixé, on peut construire la solution approchée

$$x_k = x_o + V_k y_k,$$

où y_k minimise sur $\mathcal{R}^k$ la fonctionnelle J donnée par la formule (31).

Remarque A :

Lorsque $k = N, x_k$ est la solution exacte x du problème. □

On peut désormais donner l'algorithme complet.

Algorithme 1 : Algorithme GMRES linéaire

Etape 0 : Initialisation

Choisir x_o.
Calculer $r_o = b - Ax_o$.
Calculer $v_1 = \frac{r_o}{\|r_o\|_2}$.
Poser $k = 1$.

Etape 1 :

Faire :
$h_{ik} = (Av_k, v_i)_k, \quad 1 \leq i \leq k,$
$\hat{v}_{k+1} = Av_k - \sum_{i=1}^{k} h_{ik} v_i,$
$h_{k+1,k} \|\hat{v}_{k+1}\|_2,$
$v_{k+1} = \frac{\hat{v}_{k+1}}{h_{k+1,k}}.$

Etape 2 :

Calculer la solution approchée
$x_k = x_o + V_k y_k$ où y_k minimise
$\|\beta e_1 - \tilde{H}_k y\|_2$ sur $\mathcal{R}^k$,
avec $\beta = \|r_o\|_2$.

Etape 3 : Test d'arrêt

Si $\|r_k\|$ est suffisamment petit, s'arrêter.
Sinon, faire $k = k + 1$ et retourner à l'étape 1.

Dans la pratique, on n'utilise pas tout à fait cet algorithme. En effet, il est clair que plus k est grand et plus il devient coûteux car on doit stocker la matrice V_{k+1} (ce qui nécessite une place mémoire $\simeq kN$), le nombre de multiplications augmente en $\frac{k^2 N}{2}$ et le temps de calcul pour le processus d'orthogonalisation est en $k^2 N$.

Pour remédier à cette difficulté, on transforme l'algorithme en une méthode itérative. Tout d'abord, on fixe une dimension maximale, notée k_m, à l'espace de Krylov. Puis, lorsque le nombre d'itérations atteint k_m, on calcule la solution approchée correspondante et on réinitialise l'algorithme avec celle-ci.

On donne maintenant le nouvel algorithme appelé GMRES (k_m).

Algorithme 2 : Algorithme GMRES (k_m) linéaire

Etape 0 : Initialisation

Choisir x_o.
Choisir $k_m \in [1, N-1]$.
Calculer $r_o = b - Ax_o$.
Calculer $v_1 = \frac{r_o}{\|r_o\|_2}$.

Etape 1 :

Pour $k = 1, 2, \ldots, k_m$.
Faire :

$h_{ik} = (Av_k, v_i) \quad 1 \leq i \leq k,$
$\hat{v}_{k+1} = Av_k - \sum_{i=1}^{j} h_{ik} v_i,$
$h_{k+1,k} = \|\hat{v}_{k+1}\|_2,$
$v_{k+1} = \frac{\hat{v}_{k+1}}{h_{k+1,k}}.$

Etape 2 :

Calculer la solution approchée
$x_{k_m} = x_o + V_{k_m} y_{k_m},$
où y_{k_m} minimise $\|\beta e_1 - \bar{H}_{k_m} y\|_2$ sur $\mathcal{R}^{k_m}$,
avec $\beta = \|r_o\|_2$.

Etape 3 : Test d'arrêt

Si $\|r_{k_m}\|$ est suffisamment petit, s'arrêter.
Sinon, faire $x_o = x_{k_m}, v_1 = \frac{r_{k_m}}{\|r_{k_m}\|}$
et retourner à l'étape 1.

Pour être complet dans la description de ces algorithmes, il reste encore à préciser comment minimiser $J(y) = \|\beta c_1 - \bar{H}_k y\|_2$.
L'idée de Y. Saad et M. Schultz (cf. [SS]) a été de factoriser $\bar{H}_k$ en $Q_k R_k$ où Q_k est une matrixe unitaire d'ordre $k+1$ et R_k une matrice triangulaire supérieure d'ordre $(k+1)X_k$ dont la dernière ligne est nulle. (A noter que la structure particulière de $\bar{H}_k$ facilite considèrablement ces calculs).
Minimiser $J(y)$ revient ainsi à résoudre un système triangulaire supérieur d'ordre $k+1$. De plus, cette manière de factoriser donne la norme résiduelle $\|r_k\|_2$ de la solution approchée x_k sans avoir à calculer x_k. On ne calcule x_k qu'à la fin de l'algorithme.

Donnons pour finir quelques propriétés de GMRES.

(1) Dans les algorithmes (1) et (2), on n'utilise la matrice A que pour calculer des produits matrice-vecteur. Ceci permet de simplifier le stockage de A et de traiter de grands systèmes.

(2) La solution x_k obtenue à la k-ième itération est exacte si et seulement si l'une des conditions équivalentes suivantes est vérifiée :

. l'algorithme s'arrête à la k-ième itération,

. $\hat{v}_{k+1} = 0$,

. $h_{k+1,k} = 0$,

. le degré du polynôme minimal associé à r_o est égal à k.

(3) Une conséquence de la propriété (2) est que l'algorithme GMRES (k_m) ne dégénère jamais.

(4) La norme du résidu de l'algorithme GMRES (k_m) ne peut que décroitre, mais la convergence n'est pas toujours assurée. Toutefois, la méthode s'avère en pratique très robuste.

Bibliographie

[AS] M. Abramowitz, I. Stegun, *Handbook of mathematical functions*, Dover Publications, 1968.

[Be] A. Bendali, *Approximation par éléments finis de surface de problèmes de diffraction des ondes électromagnétiques*, Thèse de doctorat d'état, Université de Paris VI, 1984.

[Br] H. Brezis, *Analyse fonctionnelle - Théorie et applications*, Masson, 1983.

[BT] A. Bayliss, E. Turkel, *Communication on Pure and Applied Mathematics*, Vol. XXXIII, pp. 707-725, 1980.

[DL] R. Dautray, J.L. Lions, *Analyse mathématique et calcul numérique pour les sciences et techniques*, Masson, 1984.

[EM] B. Engquist, A. Majda, *Absorbing boundary conditions for the numerical simulation of waves*, Math. of Comp., 1977.

[Fen] K. Feng, in Proceedings, International Congress of Mathematicians, Warsaw, Poland, p. 1439, 1983.

[Fey] R. Feynman, *Cours de physique, Electromagnétisme*, Addison Wesley, Londres, p. 969.

[Fou] G. Fournet, *Electromagnétisme à partir des équations locales*, Masson, 1985.

[G] D. Givoli, Journal of Computational Physics, Vol. 94, n° 1, 1991.

[H] L. Hörmander, The analysis of linear partial differential operators, Springer, 1983.

[KG] J.B. Keller, D. Givoli, Journal of of Computational Physics, Vol. 82, n° 172, 1988.

[R] A.G. Ramm, *Scattering by obstacles*, Reidel Publishing Company, 1986.

[Sa] W.K. Saunders, On solutions of Maxwell's equations in an exterior region, Proc. Nat. Acad. Sci., n° 38, pp. 342-348, 1952.

[SS] Y. Saad, M.H. Schultz, GMRES : *a Generalized Minimal Residual algorithme for solving non-symmetric linear systems*, SIAM J. Sci. Stat. Comp. 7, pp. 856-869, 1986.

ISSN 0249 - 6399