Abstract

The aim of this paper is to propose a numerical strategy for computing the solution of two-dimensional time-harmonic acoustic multiple scattering problems at high-frequency. The scatterers are assumed to be circular, leading therefore to semi-analytical representation formulae of the scattered field through the solution of a large linear system of equations. Taking advantage of the special block Toeplitz structure of the matrix of the linear system, a fast iterative and preconditioned numerical method yielding large memory savings is proposed. Several numerical experiments for general configurations are presented to show the efficiency of the numerical method.

© 2007 Elsevier Inc. All rights reserved.

Keywords: Multiple scattering; High frequency; Acoustic; Iterative solver; Fast algorithm

1. Introduction

Multiple scattering problems find their origins in many applications related to different areas of applied sciences: acoustics, electromagnetism, elasticity and water waves. For such problems, the scattered field appears as the superposition of elementary scattered fields resulting from the interaction between the incident wave and the scatterers on one hand, and between the scatterers on the other hand. A better understanding of such problems requires a precise knowledge of the influence of the different physical and geometrical parameters of the problem on these interactions. Due to the complexity of this problem, computing a numerical solution requires a special care, especially in the case where the number of objects is large and/or for high frequencies. For a recent and complete overview of these issues with extended bibliographical references, we refer the reader to the textbook [17] by Martin.
In this paper, we investigate the two-dimensional Helmholtz equation for circular cylinders. Following Záv-iska and next Row [24,18,17], a multipole expansion formulation of the solution can be obtained. Unlike the case of one object where the Mie series solution expressed by the Fourier coefficients is explicitly obtained through the inversion of a diagonal matrix, the case of multiple scattering is more complex and requires the use of numerical methods. Indeed, it can be shown (see Section 2 or [17] for more details) that the scattered field admits an expansion as a superposition of Fourier series. The set of Fourier coefficients is then solution of a non-diagonal and dense complex linear system which can be extremely large for many scatterers at high-frequency. The numerical solution of this system has been already obtained by some authors [15,16,20–23,25] (see also [14] for scattering by spheres). However, this remains restricted to low-moderate wavenumbers. Alternative solutions using numerical methods based on Dirichlet-to-Neumann non-reflecting boundary conditions have been recently developed by Grote and Kirsch [13]. This technique has the advantage of allowing general shaped scatterers but, on the other hand, it requires to mesh the finite computational domain. In particular, this leads to strong limitations (as in the single scattering problem) as the wavenumber increases since e.g. small element sizes are required.

To the best of the authors’ knowledge, the only papers dealing with high frequencies for multiple scattering problems are the ones by Bruno and co-authors [1,6,7,12] where scattering by three convex and non-convex objects is analyzed for two- and three-dimensional problems. The authors consider an integral equation formulation of the Dirichlet problem and derive a fast high-order method for computing the multiple scattered field. This robust solver therefore allows to consider general shaped obstacles. While the technique is efficient and robust, it is very technical and requires specific developments in the background of integral equations. The present paper proposes to come back to the scattering problem by circular cylinders using the multipole expansion and to develop a reliable and robust accurate numerical solution for large frequencies and many scatterers.

The outline of the paper is the following. In Section 2, we recall some classical results on the problem of multiple scattering by circular cylinders and on its resolution using multipole expansion formulation. In particular, we will see that the Fourier coefficients of the local scattered field solve an infinite linear system (see (19)). Section 3 explains how to suitably truncate the above infinite dimensional system to get an accurate numerical solution. Section 4 deals with the numerical resolution of the complex, dense and large size linear system obtained. We take advantage of the special block Toeplitz structure of the linear system to reduce the memory storage and accelerate the computation of the matrix. Next, we propose a fast iterative solution of the linear system using a Krylov solver in conjunction with an original geometrical preconditioner. Several examples, corresponding to many geometrical configurations, are provided to understand the characteristics of the algorithm for high wavenumbers and for a large number of scatterers. Finally, in Section 5, we conclude by summing up our results and outlining some future directions of work.

2. Multiple scattering by $M$ circular cylinders

Consider a homogeneous acoustic medium filling the whole space $\mathbb{R}^2$ and containing $M$ disjoint scatterers $B_1, \ldots, B_M$. We assume that each scatterer $B_p$, $p = 1, \ldots, M$, is a bounded subdomain of $\mathbb{R}^2$ of boundary $\partial B_p$. We denote by $\Omega_\omega = \bigcup_{p=1}^M B_p$ the domain occupied by the obstacles. We consider the scattering problem of an incident plane wave $u^{inc}(r) = e^{i k r}$ of direction $\vec{b}$ by $\Omega_\omega$ (where the time dependence is assumed to be of the form $e^{-i \omega t}$ and where the wavenumber $k$ is real). In other words, we want to determine the scattered wave $u$ solving the exterior boundary value problem

$$
\begin{cases}
\Delta u + k^2 u = 0 & (\mathbb{R}^2 \setminus \overline{\Omega}_\omega), \\
Au = -\Lambda u^{inc} & (\partial \Omega_\omega), \\
\lim_{|x| \to +\infty} |x|^{1/2} \left( \nabla u \cdot \frac{x}{|x|} - i k u \right) = 0.
\end{cases}
$$

To model sound-soft and sound-hard obstacles, the boundary operator $A$ appearing in the above relations denotes either the trace or the normal trace operator on $\partial \Omega_\omega$. A natural idea to solve numerically the multiple scattering problem $(\mathcal{E})$ is to reduce it to a family of single scattering problems. By linearity, this can be
achieved by introducing $M$ fictitious scattered waves $u^1, \ldots, u^M$, where each field $u^p$ corresponds to the wave reflected only by the scatterer $p$ when it is illuminated by the incident wave and the scattered waves $u^q$, for $q = 1, \ldots, M$ with $q \neq p$. More precisely, we have the following result (see [4] for the proof).

**Theorem 1.** Let $u$ be the solution of the multiple scattering problem $(\mathcal{E})$. Then, the family of $M$ coupled single scattering problems for $p = 1, \ldots, M$:

$$
\begin{cases}
\Delta u^p + k^2 u^p = 0 & (\mathbb{R}^2 \setminus \overline{B}_p), \\
u^p = -A \left( u^{\text{inc}} + \sum_{q=1, q \neq p}^{M} u^q \right) & (\partial B_p), \\
\lim_{|x| \to +\infty} |x|^{1/2} \left( \nabla u \cdot \frac{x}{|x|} - i ku \right) = 0,
\end{cases}
$$

admits a unique solution $(u^1, \ldots, u^M)$. Furthermore, the following decomposition holds:

$$u = \sum_{p=1}^{M} u^p. \tag{1}$$

In this paper, we focus on the case where the scattering obstacles $B_p$ are circular cylinders of radius $a_p$ centered at the points $O_p = (x_p, y_p)$ of a given orthonormal system of coordinates $(\overline{Ox}, \overline{Oy})$.

We set for all $p = 1, \ldots, M$ (see Fig. 1):

$$b_p = \overline{O O_p}, \quad b_p = |b_p|, \quad \alpha_p = \text{Angle}(Ox, b_p)$$

and for all $q = 1, \ldots, M$, with $q \neq p$:

$$b_{pq} = \overline{O_q O_p}, \quad b_{pq} = |b_{pq}|, \quad \alpha_{pq} = \text{Angle}(Ox, b_{pq}).$$

Any point $M$ of the plane will be described by its cartesian coordinates $(x, y)$ or by its polar coordinates:

$$r = \overline{O M}, \quad r = |r|, \quad \theta = \text{Angle}(Ox, r).$$

We will also use in the sequel the local polar coordinates of the point $M$ in the orthonormal system of coordinates associated to the scatterer $p$:

$$r_p = \overline{O_p M}, \quad r_p = |r_p|, \quad \theta_p = \text{Angle}(Ox, r_p).$$

![Fig. 1. A view of two typical cylinders.](image)
Let us introduce for all \( m \in \mathbb{Z} \) the following cylindrical wavefunctions, which are particular solutions of Helmholtz equations for \( r > 0 \):

\[
\begin{align*}
\psi_m(r) &= H_m^{(1)}(kr)e^{im\theta}, \\
\hat{\psi}_m(r) &= J_m(kr)e^{im\theta},
\end{align*}
\]

(2)

where \( J_n \) is the \( n \)th order Bessel function and \( H_n^{(1)} \) is the \( n \)th order Hankel function of the first kind. Note that for all \( m \in \mathbb{Z} \), the function \( \psi_m \) satisfies the outgoing Sommerfeld radiation condition. We also define for all \( m \in \mathbb{Z} \) the local cylindrical wavefunctions associated with the scatterer \( p \), for \( p = 1, \ldots, M \), by setting:

\[
\begin{align*}
\psi^p_m(r) &= \psi_m(r_p) = H_m^{(1)}(kr_p)e^{im\theta_p}, \\
\hat{\psi}_m^p(r) &= J_m(kr_p)e^{im\theta_p}.
\end{align*}
\]

(3)

Since each field \( u^p \) is an outgoing solution of a single scattering problem outside a disk, it admits the following modal decomposition in the local cylindrical outgoing wavefunctions:

\[
u^p(r) = \sum_{m \in \mathbb{Z}} e^{im\theta} \psi^p_m(r) \quad \forall p = 1, \ldots, M, \forall r_p > a_p,
\]

(4)

where the complex coefficients \( (c^p_m)_{m \in \mathbb{Z}} \) are determined by imposing the boundary condition on the boundary of the scatterer \( p \):

\[ Au^p = -Au^{inc} - \sum_{q=1}^M A u^q \quad \text{on } \partial B_p. \]

(5)

In order to explicit this equation, we have to express the incident field \( u^{inc} \) and the fictitious scattered fields \( u^q \), for \( q \neq p \), in the local system of coordinates of the scatterer \( p \). For the incident plane wave of direction \( \beta = (\cos \beta, \sin \beta) \) we recall that (cf. \([17, \text{p. 125}]\))

\[ u^{inc}(r) = \sum_{m \in \mathbb{Z}} d_m^p \hat{\psi}_m^p(r), \]

(6)

where \( d_m^p = e^{ikr_p b_p} e^{im(\pi - \beta)} \). Concerning the fields \( u^q \), we make use of the separation theorem (see for instance \([17, \text{Theorem 2.12}]\)).

**Theorem 2.** Let \( 1 \leq p,q \leq M \), with \( p \neq q \). Then, we have the following relations:

\[ \psi^q_m(r) = \begin{cases} 
\sum_{n \in \mathbb{Z}} S_{mn}(b_{pq}) \hat{\psi}^q_n(r) & \text{for } r_p < b_{pq}, \\
\sum_{n \in \mathbb{Z}} \tilde{S}_{mn}(b_{pq}) \psi^q_n(r) & \text{for } r_p > b_{pq},
\end{cases} \quad \forall m \in \mathbb{Z}, \]

(7)

where we have set

\[ S_{mn}(b_{pq}) = \hat{\psi}_{m-n}(b_{pq}), \quad \tilde{S}_{mn}(b_{pq}) = \hat{\psi}_{m-n}(b_{pq}). \]

(8)

The infinite matrices \( \mathcal{S}^{pq} = (S_{mn}(b_{pq}))_{m,n \in \mathbb{Z}} \) and \( \tilde{\mathcal{S}}^{pq} = (\tilde{S}_{mn}(b_{pq}))_{m,n \in \mathbb{Z}} \) are called separation (or transfer) matrices.

Using relations (1), (4), (5) and the first equation in (7), straightforward computations show that the unknown Fourier coefficients solve the following equations:

\[
\begin{align*}
&c^p_m + \frac{J_m(ka_p)}{H_m^{(1)}(ka_p)} \sum_{q=1}^M \sum_{n \in \mathbb{Z}} S_{mn}(b_{pq}) c^q_n = \frac{-J_m(ka_p)}{H_m^{(1)}(ka_p)} d_m^p \quad \forall m \in \mathbb{Z} \quad \forall p = 1, \ldots, M
\end{align*}
\]

(9)

in the sound-soft case and

\[
\begin{align*}
&c^p_m + \frac{J'_m(ka_p)}{H_m^{(1)}(ka_p)} \sum_{q=1}^M \sum_{n \in \mathbb{Z}} S_{mn}(b_{pq}) c^q_n = \frac{-J'_m(ka_p)}{H_m^{(1)}(ka_p)} d_m^p \quad \forall m \in \mathbb{Z} \quad \forall p = 1, \ldots, M
\end{align*}
\]

(10)
in the sound-hard case (note that \( H_n^{(1)}(ka_p) \) and \( H_n^{(1)}(ka_p) \) never vanish for real wavenumbers). The infinite linear systems (9) and (10) can be written in the more compact vector form

\[
C^p + D^p \sum_{q=1,q \neq p}^M (S^{p,q})^T C^q = B^p \quad \forall p = 1, \ldots, M,
\]

where

- \( C^p = (c^p_n)_{n \in \mathbb{Z}} \) is the infinite vector containing the coefficients of the cylindrical decomposition (4) of \( \ell^p \),
- \( (S^{p,q})^T \) denotes the transpose of the separation matrix \( S^{p,q} \) between the obstacles \( B_p \) and \( B_q \) defined by \( S^{p,q} = \sum_{m \in \mathbb{Z}} \psi_m(b_{pq}) \).
- \( D^p = (d^p_{mn})_{m,n \in \mathbb{Z}} \) is the diagonal infinite matrix, with diagonal terms
  \[
  D^p_{m,m} = \begin{cases} 
  J_n(ka_p) & \text{for sound-soft obstacles}, \\
  J_n(ka_p) & \text{for sound-hard obstacles}.
  \end{cases}
\]
- \( B^p = -D^p d^p \), where \( d^p = (d^p_{m,n})_{m,n \in \mathbb{Z}} \) is the infinite vector containing the coefficients of the cylindrical decomposition (6) of the incident wave.

The \( M \) infinite linear systems (11) can equivalently be written in the abstract form

\[
A_C = B,
\]

where

\[
A = \begin{bmatrix}
I & D^1(S^{1,2})^T & \ldots & D^1(S^{1,M})^T \\
D^2(S^{2,1})^T & I & \ldots & D^2(S^{2,M})^T \\
\vdots & & \ddots & \vdots \\
D^M(S^{M,1})^T & D^M(S^{M,2})^T & \ldots & I
\end{bmatrix},
\]

\[
C = \begin{bmatrix}
C^1 \\
C^2 \\
\vdots \\
C^M
\end{bmatrix}, \quad B = \begin{bmatrix}
B^1 \\
B^2 \\
\vdots \\
B^M
\end{bmatrix}
\]

and where \( I \) denotes the identity operator on \( \ell^2(\mathbb{C}) \).

Let us emphasize that many physical quantities of interest can be deduced from the solution of (13). The scattered field \( u \) can be obtained in the neighborhood of \( B_p \) (namely, for \( r_p < \min_{q \neq p} b_{pq} \)) by the relation

\[
u(r) = \sum_{m \in \mathbb{Z}} c^p_m \psi_m^p(r) + \sum_{m \in \mathbb{Z}} \left( \sum_{q=1,q \neq p}^M \sum_{n \in \mathbb{Z}} S_{mn}(b_{pq}) c^q_n \right) \tilde{\psi}_m^p(r).
\]  

Using \( r_p = r - b_p \cos(\theta - \alpha_p) + O(1/r) \), the scattering amplitude \( a(\theta) \) defined by

\[
u(r) = \frac{e^{ikr}}{\sqrt{r}} a(\theta) + O\left( \frac{1}{r} \right) \quad \text{as } r \to +\infty
\]

is obtained by

\[
a(\theta) = e^{-i\pi/4} \sqrt{\frac{2}{\pi k}} \sum_{p=1}^M e^{-ib_p k \cos(\theta - \alpha_p)} \left( \sum_{n \in \mathbb{Z}} c_n^{i(n-\frac{3}{2})} e^{i\alpha_p n} \right).
\]
For the Dirichlet multiple scattering problem, one can be interested in computing the normal derivative \( \partial_n U^p \) of the scattered field on the boundary of the scatterer \( p \). This can be achieved by taking the normal derivative of (15). Following [15], let us notice that the resulting formula can be simplified (to eliminate the double sum) using (13) and the Wronskian relation for Bessel functions. The final expression reads

\[
\partial_n U^p = -\partial_n U^{\text{inc}} + \frac{2i}{\pi a_p} J^{-p} C^p, \quad p = 1, \ldots, M,
\]

where \( \partial_n U^p \) (respectively \( \partial_n U^{\text{inc}} \)) is the infinite dimensional vector containing the Fourier coefficients of the normal derivative of the scattered (respectively incident) field on \( \partial B_p \) and \( J^{-p} \) is the infinite diagonal matrix with diagonal terms \( 1/J_n'(ka_p) \). A similar formula for the trace can be derived in the case of Neumann boundary conditions. Note that Eq. (17) is particularly well adapted for the numerical computations since it only involves diagonal matrices.

### 3. Finite dimensional approximation

Obviously, the infinite linear system (11) (or (13)) must be truncated to be solved numerically. In particular, only a finite number of modes can be used to describe the solution \( u^p \). Let \( 2N_p + 1 \) be the corresponding number of modes used for the scatterer \( p \), i.e. assume that we keep only the modes \( \psi_m^p \) such that \( -N_p \leq m \leq N_p \). Note that we have chosen a possibly different number of modes for each scatterer, in order to take into account geometrical configurations where the obstacles have different radii. With this notation, (11) should be truncated to the finite system

\[
C^p + D^p \sum_{q=1,q\neq p}^M (S^{p,q})^T C^q = B^p \quad \forall p = 1, \ldots, M,
\]

where

- \( C^p = (c_m^p)_{m=-N_p,\ldots,N_p} \) is the finite vector containing approximations of the first \( 2N_p + 1 \) modal coefficients of the cylindrical decomposition (6) of \( u^p \), that we still denote \( c_m^p \) for the sake of clarity.
- \( S^{p,q} \) is the \( (2N_p + 1) \times (2N_q + 1) \) finite dimensional separation matrix taking into account only the interactions between the first modes of the obstacles \( B_p \) and \( B_q \):

\[
S^{p,q} = (S_{mn}^{p,q})_{-N_p \leq m \leq N_p, -N_q \leq n \leq N_q}, \quad S_{mn}^{p,q} = \psi_m^p(\psi_n^q).
\]

- \( D^p = (d_{mn}^p)_{-N_p \leq m \leq N_p, -N_q \leq n \leq N_q} \) is the diagonal finite matrix, with diagonal terms

\[
d_{mn}^p = \begin{cases} J_0(ka_m) & \text{for sound-soft obstacles,} \\ J_1(ka_m) & \text{for sound-hard obstacles.} \end{cases}
\]

- \( B^p = -D^p d^p \), where \( d^p = (d_m^p)_{-N_p \leq m \leq N_p} \) is the finite vector containing the \( 2N_p + 1 \) first coefficients of the cylindrical decomposition (4) of the incident wave.

The \( M \) coupled finite dimensional systems (18) can equivalently be written as

\[
A C = B,
\]

where \( A \in \mathbb{C}^{N \times N} \) is the full complex square matrix of size \( N = \sum_{p=1}^{M} (2N_p + 1) \) defined by

\[
A = \begin{bmatrix}
I & D^1(S^1) & \ldots & D^1(S^M) \\
D^2(S^2) & I & \ldots & D^2(S^M) \\
\vdots & \vdots & \ddots & \vdots \\
D^M(S^M) & D^M(S^M) & \ldots & I^M
\end{bmatrix},
\]

(19)
where $T^p$ denotes the identity matrix of $\mathbb{C}^{2N_p+1}$ and

$$
C = \begin{bmatrix}
C^1 \\
C^2 \\
\vdots \\
C^M
\end{bmatrix}, \quad B = \begin{bmatrix}
B^1 \\
B^2 \\
\vdots \\
B^M
\end{bmatrix}
$$

are to complex-valued vector fields from $\mathbb{C}^N$.

The number of modes $N_p$ to retain in the approximation must be fixed to get an accurate solution. On one hand, $N_p$ must be large enough to capture both the propagating and grazing parts of the solution (typically, $N_p \geq k\alpha_p$). On the other hand, taking too many modes for approximating the solution makes the matrix $A$ ill-conditioned. Therefore, a numerical stagnation occurs when using an iterative algorithm for solving (19) (see Section 4.3.1). This phenomenon is due to the fact that high order spatial modes $|m|$ correspond to the evanescent part of the field and therefore, computing their Fourier coefficients is definitively out of reach using an iterative solver (with a fixed tolerance). For our simulations, we will use the following empirical formula:

$$
N_p = \left[ k\alpha_p + \left( \frac{1}{2\sqrt{2}} \ln(2\sqrt{2\pi k\alpha_p e^{-1}}) \right)^{2/3} (k\alpha_p)^{1/3} + 1 \right],
$$

(21)

where $[x]$ denotes the integer part of a real number $x$, and $\varepsilon$ is the desired error bound on the Fourier coefficients. The above formula has been proposed in the literature in the contexts of single scattering [11] and multipole methods [8]. Nevertheless, according to our numerical results (see Section 4.3.1), it turns out that it can also be successfully used in the framework of multiple scattering investigated in this paper.

Let us give here a formal derivation of the above formula in the case of simple scattering (see the proof given in [11] for more details). First of all, the convergence of the Fourier series $\sum_{m \in \mathbb{Z}} c_m^p H_m^{(1)}(k\alpha_p)e^{im\theta}$ in $L^2(0,2\pi)$ implies that for $|m|$ large enough, we have

$$
|c_m^p| \leq \frac{1}{|H_m^{(1)}(k\alpha_p)|}.
$$

Taking in the above relation: $|m| = k\alpha_p(1 + \zeta)$, with $\zeta \ll 1$, and using the asymptotics of Hankel functions for large arguments and large orders, one obtains that

$$
|c_m^p| \leq \frac{\sqrt{2\pi k\alpha_p}}{2} \exp \left( -\sqrt{2k\alpha_p} |m|^{3/2} \right).
$$

Therefore, we can ensure an error bound of $\varepsilon$ on the unknown Fourier coefficients by imposing the right-hand side of the above relation to be less than $\varepsilon$. This gives

$$
\zeta = \left( \frac{1}{2\sqrt{2}} \ln(2\sqrt{2\pi k\alpha_p e^{-1}}) \right)^{2/3} (k\alpha_p)^{-2/3}
$$

and thus an order of truncation $N_p = k\alpha_p(1 + \zeta)$ which is exactly formula (21).

4. Implementation and validation of a fast preconditioned iterative numerical method

We want to solve the linear system (19) for realistic complex configurations, including numerous obstacles $M$ and high frequencies $k\alpha_p$. Therefore, according to (21), the number of modes $2N_p + 1$ needed to approximate the solution $u^p$ with a reasonable precision must then be large. This implies that a huge memory storage is a priori required and large computational times are needed to build the dense matrix $A$. Moreover, we are led to solve a large scale complex-valued linear system with size $N \times N$, with $N = \sum_{p=1}^{M} (2N_p + 1)$. Using a direct linear solver would yield a prohibitive computational time, especially for high frequencies. Our solution consists in developing a strategy based on an iterative solver leading to problems related to fast evaluations of
dense matrix–vector products as well as convergence questions. We carefully analyze these delicate problems in the present section after presenting the different geometrical test configurations.

4.1. Geometrical configurations

In the numerical computations, we consider three kinds of geometrical configurations:

- **The single-row configuration:** this structure is composed of $M_x$ equally spaced obstacles aligned along the $x$-axis, the distance between two successive scatterers being denoted $b_x = b_{12}$ (see Fig. 2). Moreover, the row is centered at $x = 0$.

- **The centered uniform square/rectangular lattice:** we consider here a rectangular lattice composed from $M = M_x \times M_y$ circular cylinders (the structure is called square lattice if $M_x = M_y$). For brevity, we restrict our experiments to a rectangular lattice which is composed of $M_y$ uniformly spaced single-rows with respect to $b_y = b_{1(2M_x)}$, each row being composed from $M_x$ equally spaced disks according to $b_x = b_{12}$ (see Fig. 3). Moreover, the array is centered at the origin.

- **The triangular lattice:** this structure is composed of two parallel horizontal single-rows. The first one contains $M_x \geq 2$ equally spaced disks and the second one $M_x - 1$. The horizontal distance between two objects is $b_x$. Next, the first row is repeated vertically $M_y$ times with a uniform distance $b_y = b_{1(2M_x)}$, while the second one is reproduced $M_y \pm 1$ times with again a separation distance $b_y$ (see Fig. 4).

4.2. Storage

Although being full, the matrix $A$ has a particular structure. Indeed, each one of its off-diagonal blocks is obtained (see (20)) by multiplying the diagonal matrix $D^p \in \mathbb{C}^{2N_p+1,2N_q+1}$ by the matrix $(S^{p,q})^T \in \mathbb{C}^{2N_p+1,2N_q+1}$ which has a Toeplitz structure [9] since

$$S^{p,q}_{mn} = \psi_{m-n}(b_{pq}).$$

Consequently, using the notations from [9], the storage of $(S^{p,q})^T$ can be optimized using a compressed version based on the root vector

$$\sigma^{p,q} = (S^{p,q}_{N_q-N_p}, \ldots, S^{p,q}_{N_q+1-N_p}, \ldots, S^{p,q}_{-N_q-N_p}, \ldots, S^{p,q}_{-N_q-N_p})^T.$$  (22)

To take advantage in our future algorithms of the special structure of each off-diagonal block of $A$, we store both the Toeplitz matrix $(S^{p,q})^T$ (through the root-vector (22)) and the diagonal matrix $D^p$. According to our notations, the compressed storage needs $2(2N_p + N_q + 1)$ entries instead of the $(2N_p + 1)(2N_q + 2)$ complex coefficients required for the full version. For $A$, this must be repeated for the $M(M-1)$ off-diagonal blocks by summing over $p$ and $q$. This results in a global storage equal to $3N(M-1)$ entries which must be compared to the $N^2 - NM$ cost of the full storage. Furthermore, the computational time involved in the construction of the global matrix is also reduced according to the memory storage. In the case where we have $a_p = a$ for any $p$, the vector root version of $A$ leads to a memory storage and a CPU time of the order of $O(6kaM^2)$ while it is $O(4k^2a^2M^2)$ for the full version. This is a crucial point for solving a multiple scattering problem for a large wavenumber. To show the improvement induced by the compressed storage version using (22) over the full...
version, we represent on Fig. 5 the logarithm of the CPU time scaled by the computational memory requirement with respect to the wave number \( k a_p \) needed for building the global matrix \( A \). We consider a single-row configuration for \( M = 2 \), \( a_p = 1 \), for \( p = 1, 2 \), and \( b_x = 3 \). As expected, the CPU time for the compressed version is linear according to \( k a_p \) while it is quadratic for the full version.

4.3. Iterative solution

Since we only have access now to the compressed Toeplitz format (22) of the matrix \( A \), the linear system (19) can only be solved through an iterative linear solver. We consider both the GMRES, possibly with a restart parameter \( \eta \) (then denoted by GMRES(\( \eta \))), and the BICGStab algorithms. The tolerance error of the iterative solver is set to \( tol \) and the number of iterations to get this tolerance is denoted by \( n^{\text{iter}} \).

Let us recall that the main CPU cost of the GMRES is due to one matrix–vector product (MVP) per iteration. Moreover, a large amount of memory can be necessary to store the Krylov basis if too many iterations are needed to reach the tolerance. We will see that it leads us to consider a restart parameter if the number of obstacles is large and the frequency high. Two MVPs are required for each iteration of the BICGStab. This finally gives the total cost which is related to each algorithm. We first study numerically the influence of the order of truncation \( N_p \) on the convergence of the iterative solver. In particular, this validates the choice of \( N_p \) given by formula (21). Next, we explain how to reduce the computational cost of a global MVP using a fast algorithm for each Toeplitz subblock. We investigate numerically the dependence of the rate of convergence of

---

Footnote: All the computations were performed on a Power Mac G4 1.67 GHz with 1 Go DDR SDRM. The algorithms are developed under Matlab.
the iterative solver with respect to the geometrical and physical parameters. Finally, we propose and study a geometrically-based preconditioner.

From now on, we consider an angle of incidence $\beta = 0$ and we restrict the presentation to the Dirichlet case (which appears to be harder to solve than the Neumann problem in terms of iterations).

4.3.1. Influence of the order of truncation $N_p$ on the convergence

As noticed in Section 3, the order of truncation $N_p$ must be fixed carefully e.g. through formula (21): $N_p$ must be large enough to compute accurately the solution, but not too large to avoid the stagnation of the iterative solver. To make this statement precise, let us consider a uniform square lattice with $M_x = M_y = 2$ ($M = 4$) and $b_x = b_y = 3$ for a radius $a_p = 1$, $1 \leq p \leq M$, and $k = 100$. The linear system (19) is solved by the GMRES with $tol = 10^{-8}$. For a given value of $tol$ and all along the paper, we fix $\varepsilon = tol$ in formula (21). This gives here $N_p = 120$ (represented by a red dot in Figs. 6 and 8). We report $n^{\text{iter}}$ versus $N_p$ in Fig. 6. One can observe three distinct zones. First, from $N_p = 1$ to $N_p < ka_p + 2$, $n^{\text{iter}}$ increases. This means that the computation of a correct solution requires more harmonics. This is obtained in the second stable zone (for $k a_p + 2 \leq N_p \leq 138$). However, if we include too many harmonics (third zone), typically $N_p \geq 139$ in our example, then we obtain a break down of the GMRES as it can be remarked in Figs. 6–8. Indeed, stagnation

---

Fig. 5. Behaviour of the CPU time according to the wavenumber $ka_p$ for building the matrix $A$ in the case of the scattering by two circular cylinders fixing $N_p$ by formula (21).

Fig. 6. Number of iterations with respect to $N_p$. 
occurs e.g. for $N_p = 150$ while it does not for $N_p = 120$. In particular, the relative error corresponding to the stagnation at $N_p = 150$ is equal to $10^{-2}$ while it is $tol$ for $N_p = 120$.

4.3.2. Fast MVPs for Toeplitz matrices

The direct computation of a MVP $y = Ax$ can be computed by blocks. Let us set: $x = (x_1, \ldots, x_M)$, with $x_p \in \mathbb{C}^{2N_p+1}$, and $y = (y_1, \ldots, y_M)$, with $y_p \in \mathbb{C}^{2N_p+1}$. Then, we directly have from the structure of $A$ given by (20)

$$y_\ell = x_\ell + D^p \sum_{1 \leq p \leq \ell \leq M} z_\ell,$$

$$z_\ell = (S^{p, \ell})^T x_p$$

with $1 \leq \ell \leq M$. The main cost of the above evaluation is linked to the computation of $z_\ell$ which is quadratic according to $(2N_p + 1)$. Moreover, this must be repeated for each sub-block and each component of $y_\ell$. This is very expensive when the frequency is large since the size of $(S^{p, \ell})^T$ is $(2N_\ell + 1) \times (2N_p + 1)$. Another way of computing a Toeplitz MVP for a matrix of size $n \times n$ is to use the fast algorithm explained in ([9, pp. 95–96]) for MVPs involving Toeplitz matrices. The idea consists in building an associated circulant matrix using the Toeplitz matrix and next applying an FFT-based MVP algorithm for circulant matrices. This algorithm is
coded using Matlab FFT function. The resulting total cost in terms of real operations for computing $y'_{\ell}$ is composed from: one complex-valued Toeplitz MVP in $15n\log_2(n) + n$ operations (see [5, p. 193]), with $n = 2N_p + 2N_l + 2$, summing up next on $1 \leq p \neq \ell \leq M$, the computation of the diagonal matrix $D^p$ which is $6N_l + 3$, and finally adding $x_l$. Again, summing up on $\ell = 1, \ldots, M$ gives the total cost. If $a_p = a$, then, this requires asymptotically $O(60(M - 1)^2kalog_2(4ka))$ operations compared to $O(4(M - 1)^2ka^2)$ for a direct MVP. An example is given in Fig. 9 showing the CPU time reduction with respect to $ka$ (with $a = a_p$, $1 \leq p \leq M$) using the fast MVP algorithm compared to the direct algorithm in the single-row configuration with $M = 30$ and $b_x = 3$.

4.3.3. Numerical study of the convergence rate

Iterative solver: as already mentioned, the linear system (19) must be solved iteratively using either the BICGStab or the GMRES [19]. From a large set of numerical simulations, the GMRES provides the fastest convergence rate (see Fig. 10). However, a more important memory storage is required which can significantly limit the possibility of prospecting high-frequencies for example. We present on Fig. 10 the behaviour in terms of MVPs of the GMRES, BICGStab and GMRES(50) according to $ka$ in the single-row configuration.

Fig. 9. Comparison of the CPU time for computing one MVP using the direct and fast algorithms. We take: $ka = 100$, $b_x = 3$ for the single-row configuration (Dirichlet problem).

Fig. 10. Number of MVPs with respect to the wavenumber $ka$: single-row configuration with 75 obstacles and $b_x = 3$. We use different iterative solvers: GMRES, GMRES(50) and BICGStab. The tolerance is $tol = 10^{-10}$ (Dirichlet problem).
can see that the GMRES breaks down at \( ka = 140 \) due to memory limitations. This is not the case of BICG-Stab and GMRES(50) which keep on working. Moreover, it appears that the restarted GMRES generally leads to similar or better convergence results than the BICGStab. For these reasons, we choose the GMRES(50) in the sequel. Another way of improving the convergence rate consists in preconditioning the linear system. This solution is analyzed in Section 4.3.4 where we build a geometrical preconditioner.

Single-row configuration: Fig. 10 shows the dependence of the number of MVPs according to the wavenumber \( ka \). As seen, for this complex configuration, the number of MVPs increases with \( ka \). For a fixed wavenumber, we also observe a dependence with respect to the number of obstacles \( M \) and the distance \( \delta \) between two obstacles: \( \delta = b_x - a \). We remark in Fig. 11 that the number of MVPs increases linearly with \( M \). Moreover, the slope of the line is more important as \( \delta \) tends toward zero, meaning that we have two closer successive scatterers. This is observed again on Fig. 12 where the number of MVPs is given according to \( \delta \) in logarithmic scale for \( ka = 100 \) and \( M = 10 \). We see that the number of MVPs strongly decreases as the separation distance \( \delta \) tends to infinity, i.e. \( \delta \gg \lambda \). This corresponds to a weaker coupling between the obstacles in the multiple scattering phenomenon. For small values of \( \delta \), \( \delta \ll \lambda \), the number of MVPs strongly increases, because the linear system becomes ill-conditioned. Finally, we observe an intermediate resonance region for \( \delta \approx \lambda \) where we have

![Fig. 11. Number of MVPs with respect to the number of obstacles \( M \): single-row configuration with \( ka = 100 \) and two different values of \( b_x \). The GMRES(50) solver is used fixing \( tol = 10^{-8} \) (Dirichlet problem).](image1)

![Fig. 12. Number of MVPs with respect to the distance \( \delta \) between two obstacles. We fix: \( ka = 100, M = 10 \) obstacles. The solution is obtained with GMRES(50) for \( tol = 10^{-8} \) (Dirichlet problem).](image2)
a few peaks in the number of MVPs. A physical interpretation of this phenomenon is the following. In this regime, an approximate model consists in considering that two close circular scatterers, for \( ka \gg 1 \), behave like two parallel planes separated by a distance \( \delta \). Then resonances occur for such a configuration when
\[
ka \approx \frac{n \pi}{d} \in \mathbb{N}.
\]
This is confirmed for \( n = 1, 2, 3, 4 \), in Fig. 12, corresponding respectively to \( \delta = 3 \times 10^{-2}, 6 \times 10^{-2}, 1.2 \times 10^{-1}, 2.4 \times 10^{-1} \). This generally creates a deterioration of the condition number at these frequencies, due to small eigenvalues in the matrix of the system. A similar problem also arises in the context of integral equations [10,11,2,3].

**Rectangular and triangular lattice configurations:** we now analyze the rectangular lattice configuration. A first test-case is given in Fig. 13. We consider a rectangular lattice with \( M_x = 8 \) and increase the size of layers according to \( M_y = 2, 4, 8 \). We fix \( b_x = 8 \) and \( b_y = 13 \). The number of MVPs required by the GMRES(50) is represented as a function of \( ka \). We observe a stabilization with the frequency but the number of MVPs is larger with \( M_y \), and thus, with \( M \). This is consistent with the previous observations in the single-row case. This situation is, in some sense, not extreme because the distance between the obstacles is sufficiently large.

A more difficult problem is considered in Fig. 14 where we represent the number of GMRES(50) iterations of an \( 5 \times M_y \) lattice, for different values of \( M_y \), according to \( ka \) (for \( b_x = b_y = 3 \)). We observe that the number

![Fig. 13. Number of MVPs with respect to the wavenumber \( ka \) for the rectangular lattice configuration with \( b_x = 8 \) and \( b_y = 13 \). We fix \( M_x = 8 \) and \( M_y = 2, 4, 8 \). The solution is obtained without and with the preconditioned GMRES(50) for \( tol = 10^{-8} \) (Dirichlet problem).](image1)

![Fig. 14. Number of MVPs with respect to the wavenumber \( ka \): rectangular lattice configurations with \( M_x = 5, M_y = 2, \ldots, 6 \). We consider \( b_x = b_y = 3 \) and solve the linear system using the GMRES(50) solver for \( tol = 10^{-8} \) (Dirichlet problem).](image2)
of MVPs is again slightly dependent on $ka$ but strongly varies with the number of layers, characterized by $M_s$. Moreover, some peaks appear at some frequencies, and in particular at $ka = 20$. This is more clearly visible in Fig. 15 where we increase the lattice size according to $M = M_x^2$ at $ka = 50$. We notice that the number of MVPs increases strongly with $M_x$. In particular, on this example, it can be shown that, on the first values, the number of MVPs behaves like $3M^{3.3}$ and not quadratically with $M$. Another way of considering particular frequencies where peaks occur consists in modifying $d$. Some numerical computations, not reported here, confirm this property.

Concerning the triangular lattice, similar conclusions can be drawn. However, it appears that this situation is less dramatic in terms of MVPs compared to the rectangular lattice case.

### 4.3.4. A geometrically-based preconditioner

Since many iterations may be necessary in some situations, one way to improve the rate of convergence is to precondition the linear system. In our context, we cannot directly apply an algebraic strategy like the incomplete LU or SPAI preconditioners [19,9] which would require to reconstruct the full version of the matrix $A$. An alternative direction is to build a geometrically-based preconditioner. We propose here a simple procedure into two steps:

First, let us introduce the notation: $A = I + F$, where $I = \text{diag}((I_j)_{1 \leq j \leq M})$ is the identity diagonal block of $A$ and $F = A - I$ is its complement off-diagonal part. Using the first-order approximation of the Neumann series of $A^{-1}$ gives

$$A^{-1} \approx I - F = P.$$ (24)

In fact, there is no reason to assume that $F$ satisfies $\rho(F) < 1$, where $\rho(F)$ stands for the spectral radius of $F$. However, relation (24) must be viewed as a formal way of building an approximation of the inverse of $A$ and so has also a subjacent limitation range. It could be possible to choose more terms in the approximation. However, extensive numerical computations show that this is not a good strategy. It can even lead to the divergence of the method.

Since $P$ is still a matrix taking all the interactions between the obstacles, it is interesting to reduce its application cost by considering only the closest interactions. This can be done through a second approximation by introducing a parameter $d > 0$ representing a maximal coupling interaction distance. Then, the preconditioner, denoted by $P_d$, only considers the interactions between obstacles with indices $1 \leq p,q \leq M$ satisfying $b_{pq} < d$. We must notice that the construction of $P_d$ is implicit from $A$ and does not require any extra cost. From intensive numerical experiments, it appears that $d = b_x$ is an optimal choice for the single-row configuration while $d = \max(b_x,b_y)$ is the best choice for both the regular rectangular and triangular lattices. Taking a smaller or

---

**Fig. 15.** Number of MVPs with respect to $M_x = \sqrt{M}$ (setting $M_y = M_x$): rectangular lattice configuration with $ka = 50$ and $b_x = b_y = 3$. The GMRES(50) solver is used fixing $\text{tol} = 10^{-8}$ (Dirichlet problem).
larger value yields a slower convergence or sometimes divergence. With this choice, the application of the preconditioner requires a negligible additional cost compared to the unpreconditioned version of the solver.

To show the improvement induced by the proposed preconditioner, we present in Fig. 16 the number of MVPs for a single-row configuration with \( M = 75 \) obstacles. We observe the improvement in terms of convergence rate if we compare these results to the ones obtained in Fig. 10. The preconditioner always improves the convergence even for very close scatterers. The situation is more delicate when considering the rectangular or triangular lattices. We report in Fig. 13 the results obtained for different layers \( M_y = 2, 4, 8 \), setting \( M_x = 8 \). The distance between the obstacles is \( b_x = 8 \) and \( b_y = 13 \) (and so \( d = 13 \)). We see an interesting gain in terms of reduction of iterations. However, it appears that for situations where the scatterers are close \(( b_x - 2a \approx b_y - 2a \leq a = \min\{a_p\}_{1 \leq p \leq M} \), the preconditioner is not efficient and can even lead to a deterioration of the convergence. This means that, for this kind of configuration, more efforts must be done for building a suitable preconditioner.

We conclude this Section by analyzing the performance of our numerical method for an unstructured geometrical configuration. We consider 60 unit circular cylinders (see Fig. 17) which are supposed to be distant enough. More precisely, we assume that \( b_{\min} = \inf_{1 \leq p < q \leq M} b_{pq} \geq 3 \).
Fig. 18 shows the convergence rate of our algorithm without preconditioner and with the preconditioner $P_d$ for $d = 1.5b_{\text{min}}$. Once again, we note a faster convergence for the preconditioned algorithm.

5. Conclusion and future work

In this paper, we investigated the numerical simulation of high frequency multiple scattering by circular cylinders. The main difficulty that arises in this context is due to the fact that the complex dense linear system to be solved is very large and ill-conditioned. This is in particular true when the number of scatterers is large and/or for high frequencies. Taking advantage of the particular block Toeplitz structure of the matrix of the linear system, we proposed an adapted storage of the system and an iterative algorithm of resolution, based on a fast MVP computation. We realized a thorough numerical study of the convergence rate with respect to different geometrical parameters of the problem (the reduced wavenumber, the distance between the scatterers, the number of scatterers). Finally, we proposed and tested the efficiency of a geometrically-based preconditioner, obtained by taking into account close interactions.

Considering the results obtained in this paper, we plan in the future to tackle the following questions:

- **Series truncation**: the heuristic formulae (21) defining the number of harmonics to be taken into account to obtain an accurate approximation needs to be proved rigorously.
- **Preconditioner**: the efficiency of the preconditioner we proposed seems from our numerical experiments to be limited in some situations (very close scatterers for instance). This is due to the fact that the interactions become stronger in this case.
- **Extensions**: a natural continuation of our work is to investigate other scatterers with simple shapes, the three-dimensional Helmholtz equation and Maxwell’s equations.

References