THE FAST MULTIPOLE METHOD I: ERROR ANALYSIS AND ASYMPTOTIC COMPLEXITY

ERIC DARVE†

Abstract. This paper is concerned with the application of the fast multipole method (FMM) to the Maxwell equations. This application differs in many aspects from other applications such as the N-body problem, Laplace equation, and quantum chemistry, etc. The FMM leads to a significant speed-up in CPU time with a major reduction in the amount of computer memory needed when performing matrix-vector products. This leads to faster resolution of scattering of harmonic plane waves from perfectly conducting obstacles. Emphasis here is on a rigorous mathematical approach to the problem. We focus on the estimation of the error introduced by the FMM and a rigorous analysis of the complexity \( O(n \log n) \) of the algorithm. We show that error estimates reported previously are not entirely satisfactory and provide sharper and more precise estimates.

Key words. fast multipole method, electromagnetic theory, scattering, iterative method, matrix compression algorithms, computational aspects

AMS subject classifications. 65F10, 65R20, 65Y20, 65D05, 78A45, 78A99

PII. S0036142999330379

Introduction. The fast multipole method (FMM) is a tool that has been widely applied in several fields with great success: N-body problem, Laplace equation, quantum chemistry (see [14], [15], [7, 8, 6]). Another field of application for this method is the compression of dense matrices for the Maxwell equation in an integral formulation.

This formulation consists of the computation of the currents on the surface of the scattering object. A Galerkin method leads to a linear system with a dense matrix. This system can be solved with an iterative method. The FMM will allow fast matrix vector products and an optimal compression of the matrix. This matrix, after compression, will usually be stored in core instead of out of core, as is done if one uses a Gauss method or computes all the terms in the matrix. Thus the FMM allows for the resolution of problems of unprecedented size.

However, the FMM is a complex tool especially when applied to the Maxwell case and implies several approximations. The error that is introduced in the computation has to be a priori evaluated. Moreover we will prove in this paper that there is an essential link between the error estimations and the asymptotic complexity of the algorithm. This can be easily understood since the precision is linked to the number of terms used in the multipole expansion, which in turn will affect the total complexity of the code.

Few studies were devoted so far to a serious estimation of the complexity of the FMM. As far as the author knows, regarding theoretical aspects of the FMM applied to the Maxwell equation, there are two articles: the first by Epton and Dembart [12] and the second by Rahola [20] (also an early paper by Rokhlin [21]). However, those two papers do not entirely solve the problem of the error estimation as we will see in section 7.

The article by Epton and Dembart gives very detailed formulae regarding the translation operators used in the FMM and the way to diagonalize them. They give...
the key formulae for the FMM and we will rely on this article for their proof. They involve the Wigner 3-j symbol which appears as the value of the integration of a spherical harmonic against the product of two other spherical harmonics. They also express those translation operators in a convolutional form. However, this article does not address the problem of the error introduced by the FMM in the solver nor does it address the problem of the complexity of the algorithm.

The second article by Rahola is much closer to our own. The formalism used is totally equivalent to ours. He gives two error estimations corresponding to Theorems 1 and 2. It should be mentioned that the first estimation given by Rahola (Lemma 4.1) can be found in a few other papers: for example, see [21, Theorem 3.3] and also [18, sections 3.1.1 and 3.2]. The original Rokhlin’s paper [21], which was the first to discover the single step FMM (for Maxwell), included straightforward error estimations. All those papers are based on the asymptotic behavior of the Bessel functions

\[ j_l(z) \sim \frac{1}{2l+1} \sqrt{\frac{e}{2}} \left( \frac{ez}{2l+1} \right)^l, \]

\[ y_l(z) \sim -\frac{1}{z} \sqrt{\frac{2}{e}} \left( \frac{ez}{2l+1} \right)^{-l}, \]

as \( l \to +\infty \) (formula (9.3.1) in [1]). Using those formulae has two serious drawbacks. First they lead to inequalities which are valid when \( l \) is large compared to \( |z| \). They are not adapted to our case where we need estimates for \( l \sim |z| \) to study the convergence. Moreover they lead to bounds which involve constants depending on \( z \). Our results involve true constants and lead to bounds and estimates which can be used in a real life code, for all configurations (shape, size) of the scattering object. This implies a total control of the error and ensures that the code has an optimal complexity. We believe it to be a considerable improvement for real life applications.

Sections 1 and 2 give a quick general overview of the FMM applied to the Maxwell equation.

Section 3 gives the main theorems and results proved in this paper.

In section 4 we deduce the previous results from a more general theorem which is our key tool to study the convergence of the FMM. It is proved in the two following sections 5 and 6.

1. General presentation of the multipole method.

1.1. Maxwell equation and integral formulation. We want to solve the Maxwell equation with an integral formulation for a monochromatic signal of frequency \( \frac{\omega}{2\pi} \). Several integral formulations can be derived. In the case of a perfectly conducting object the most popular ones are EFIE and MFIE: electric/magnetic integral formulation. With a Galerkin method, the EFIE formulation leads to a linear system \( Ax = b \). We denote the basis functions of the Galerkin method by \( J_l(x) \) for edge \( l \) and \( J_k(x) \) for edge \( k \):

\[ 1 \leq k, l \leq n, \]

\[ b_l \overset{def}{=} \frac{1}{\kappa Z} \int_{\Gamma} J_l(x) E_l^{inc}(x) \, d\gamma(x), \]

where \( E_l^{inc} \) is the tangential part of the incident field
\[ A_{(k,l)} \overset{\text{def}}{=} \int_{\Gamma} \int_{\Gamma} G(x,y)(J_k(x)J_l(y) - 1/\kappa^2 \nabla_{\Gamma} J_k(x) \nabla_{\Gamma} J_l(y)) \, d\gamma(x) \, d\gamma(y), \]

\[ G(x,y) \overset{\text{def}}{=} \frac{\kappa^2}{\kappa^2} h_0^{(1)}(\kappa|x-y|) \overset{\text{def}}{=} \frac{e^{\kappa|x-y|}}{4\pi|x-y|}. \]

Notation.
\[ \iota : \sqrt{-1}, \]
\[ \lambda : \text{wavelength}, \]
\[ \kappa : \text{wave number} \quad \kappa = 2\pi/\lambda, \]
\[ \nabla_{\Gamma} : \text{two-dimensional divergence on the surface}, \]
\[ h_0^{(1)} : \text{spherical Hankel function of the first kind}, \]
\[ Z : Z^{120\pi}. \]

While some integrals must be computed partly analytically when the edges \( k \) and \( l \) are too close, for most pairs \( k,l \) a simple Gauss quadrature gives a very good approximation of \( A_{(k,l)} \).

Using an iterative method (conjugate gradient, GMRES \ldots) and a preconditioner (such as SPAI; see [16] and [2]) the computer intensive part of the resolution of the linear problem reduces to the computation of matrix-vector products of the kind
\[ \sum_j e^{\kappa|x_i-x_j|}/4\pi|x_i-x_j| \sigma_j \overset{\text{def}}{=} \sum_j M_{(i,j)} \sigma_j \]
for all \( i \) such that \(|x_i-x_j|\) is larger than some threshold. The points \( x_i \) are the Gauss points on the surface of the object.

1.2. Fast multipole method.

Fast products. A fast matrix-vector product may be performed if the matrix \( M \) can be approximated, with an error \( \eta \), as
\[ M_{(i,j)} = \sum_{m \leq M} f_i^m g_j^m + \eta. \]
Such a formula leads to a matrix-vector product with \( 2Mn \) floating operations:
\[ \sum_{j \leq n} M_{(i,j)} \sigma_j = \sum_{m \leq M} f_i^m \left( \sum_{j \leq n} g_j^m \sigma_j \right) + \eta O(\sigma). \]

Similar decomposition obtained for the Maxwell case. We consider the kernel \( h_0^{(1)}(\kappa|x-y|) \). We denote by \( P_m \) the Legendre polynom defined by
\[ P_m(x) = \frac{(-1)^m}{2^m m!} \frac{d^m(1-x^2)^m}{dx^m}. \]
We call “Gegenbauer” series (or sequence) the following expansion of \( h_0^{(1)}(\kappa|x-y|) \) (see the addition theorems in [1, Formulae (10.1.45) and (10.1.46)]):
\[ h_0^{(1)}(\kappa|x-y|) = \sum_{m=0}^{+\infty} (2m+1) h_0^{(1)}(u) j_m(v) P_m(w), \]
where \( u = \kappa|U| \quad v = \kappa|V| \quad w = \cos \gamma \)
and where \( U \) and \( V \) are any two vectors such that
\[
x - y = U - V \quad \text{and} \quad |V| < |U|
\]
and \( γ \) is the angle between vector \( U \) and \( V \).

The quadrature points \( x \) and \( y \) are gathered inside parallelepipedic clusters of centers \( X \) and \( Y \). The vectors \( U \) and \( V \) are
\[
V = y - Y + X - x,
\]
\[
U = X - Y.
\]

Thus condition \( v < u \) is satisfied if the two clusters are sufficiently far away from each other. See Figure 1.

We define the truncated Gegenbauer series by
\[
\tilde{h}_0^M(\kappa|x - y|) \overset{\text{def}}{=} \sum_{m=0}^{M} (2m + 1) h_m^{(1)}(u) j_m(v) P_m(w).
\]

We recall the representation of \( \tilde{h}_0^M(\kappa|x - y|) \) on which the FMM is based.

**Lemma 1.1 (representation of \( \tilde{h}_0^M(\kappa|x - y|) \)).** For all \( \epsilon \), there is a set of points \( s_k \) on the unit sphere \( S^2 \), and a set of weights \( 0 < \omega_k < 1, 0 \leq k \leq K \), such that
\[
T_M(s_k, U) \overset{\text{def}}{=} \sum_{m=0}^{M} \frac{2m + 1}{4\pi m} h_m^{(1)}(\kappa|U|)P_m(\cos(s_k, U)),
\]
\[
\tilde{h}_0^M, K(\kappa|x - y|) \overset{\text{def}}{=} \sum_{k=0}^{K} \omega_k e^{i\kappa(s_k, y - Y)} T_M(s_k, U) e^{i\kappa(s_k, X - x)};
\]
\[
\left| \tilde{h}_0^M(\kappa|x - y|) - \tilde{h}_0^M, K(\kappa|x - y|) \right| < \epsilon,
\]

where
\[
\tilde{h}_0^M(\kappa|x - y|) \overset{\text{def}}{=} \sum_{m=0}^{M} (2m + 1) h_m^{(1)}(u) j_m(v) P_m(w).
\]

This is proved using the decomposition of a plane wave as a sum of Bessel functions \( j_m \). Such a formula is given in several papers related to the Maxwell FMM: [21, 22, 10, 19, 23, 17].

Equation (2) is similar to (1) and thus leads to a fast algorithm for matrix-vector products. For a more complete description of the algorithm, see [20].
2. The multilevel algorithm.

2.1. General idea. The multilevel scheme starts with an oct-tree decomposition of the quadrature points on the surface of the scattering object. We give a rough description of the algorithm. It will have to be detailed later.

- Step 1: We denote by $P$ one of the clusters at the lowest level, and compute its radiation function, defined by
  \[ F^P_k \overset{\text{def}}{=} \sum_{x_i \in P} \sigma_i e^{i \kappa \langle s_k, X_P - x_i \rangle}, \]
  where $X_P$ is the center of cluster $P$.
- Step 2: All those contributions are added to the father, one level up in the oct-tree, i.e., if cluster $Q$ is the father of $P$ then we perform
  \[ F^Q_k = F^Q_k + F^P_k e^{i \kappa \langle s_k, X_Q - X_P \rangle}. \]
- Step 3: The “transfers” are performed, i.e., if clusters $P$ and $Q$ are such that they are not neighbors but their fathers are neighbors, we perform
  \[ G^P_k = G^P_k + F^Q_k T_M(s_k, X_Q - X_P). \]
- Step 4: All those contributions are added to the sons, i.e., if $Q$ is the son of $P$ we perform
  \[ G^Q_k = G^Q_k + G^P_k e^{i \kappa \langle s_k, X_Q - X_P \rangle}. \]
- Step 5: We sum on $S^2$ to obtain the final value of the matrix vector product for each point $x_i$:
  \[ \sum_k \omega_k e^{i \kappa \langle s_k, x_i - X_Q \rangle} G^Q_k. \]

2.2. Fast interpolation methods. As it will be proved later the functions $F^P_k$ contain more and more oscillatory modes as we go up in the tree. This is due to the multiplication by the factors $e^{i \kappa \langle s_k, X_Q - X_P \rangle}$ of frequency roughly $\kappa |X_Q - X_P|$. Thus to represent accurately those functions more and more points must be taken on $S^2$ to account for this oscillatory behavior. An interpolation must be performed on $F^P_k$ to compute the value of the radiation function at nodes $s'_k, 1 \leq k \leq K'$ from its value at nodes $s_k, 1 \leq k \leq K$. It will be proved later that the number of significant frequencies in $F^P_k$ is proportional to the size of the cluster. We will not write down the constant in the following for clarity.

Let $|P|$ be equal to $\kappa$ times the diameter of the cluster. Among the four steps described in the previous subsection, two steps have to be modified:

- Step 2: The radiation function is sampled at $|P|^2$ points on $S^2$. We have to compute its value at $|Q|^2$ points on $S^2$, which is the number of points used to represent the function $F^Q_k$. In order to do so, we compute the Fourier coefficients on $S^2$ for the $|P|^2$ first frequencies. A backward Fourier transform on $S^2$ is then performed to obtain $|Q|^2$ terms on the unit sphere. This is done prior to the multiplication by $e^{i \kappa \langle s_k, X_Q - X_P \rangle}$.
- Step 4: In a similar fashion, $G^P_k$ is sampled at $|P|^2$ points after step 3 of the FMM. However, after the multiplication by $e^{i \kappa \langle s_k, X_Q - X_P \rangle}$ we can represent
accurately the function $G^Q$ with fewer sample points. This subsampling is done in order to reduce the CPU. The higher frequencies can be eliminated, as we go down the tree from one level to the other. This is done by performing a forward Fourier transform on $S^2$ and then a backward Fourier transform using the first $|Q|^2$ frequencies only.

In practical implementations $|Q|$ is equal to $2|P|$, for example, if we partition the surface with an oct-tree. With this assumption we now give an algorithm to perform those transformations (interpolation or subsampling) with a complexity of $O(|P|^2 \log |P|)$. It consists of the following operations.

The reference for the algorithm can be found in [3].

Let us consider a discrete set of values $F(s_k)$. The set $s_k$ is such that there exists angles $\phi_i$ and $\theta_j$ such that $s_k = (\phi_i, \theta_j)$ in spherical coordinates.

- Perform a fast Fourier transform with $\phi_i$, $1 \leq i \leq I$ and obtain $F_m(\theta_j) = \sum_i F(\phi_i, \theta_j)e^{-im\phi_i}$.

- Forward and backward $\theta$ transform. We denote by $\bar{P}_{lm}$ the associated Legendre function with the proper normalization:

$$
\bar{P}_{lm}(x) = \sqrt{(2l+1)(l+m)!(l-m)!} P_{lm}(x) = (-1)^m \sqrt{(2l+1)(l+m)!(l-m)!} (1-x^2)^{m/2} d^{m} P_{l}(x)
$$

and by $\epsilon_{lm}$ the following sequence:

$$
\epsilon_{lm} = \sqrt{\frac{l^2 - m^2}{4l^2 - 1}}.
$$

The $\theta$ transform is defined by:

$$
\tilde{F}_m(\tilde{\theta}_k) = \sum_{n=|m|}^{N} \left( \sum_j F_m(\theta_j)\omega_j \bar{P}_n^m(\cos \theta_j) \right) \bar{P}_n^m(\cos \tilde{\theta}_k),
$$

the integer $N$ being the highest degree for the spherical harmonics

$$
= \sum_j F_m(\theta_j)\omega_j \epsilon_{n+1} \frac{\bar{P}_n^{m+1}(\cos \tilde{\theta}_k)\bar{P}_n^{m}(\cos \theta_j) - \bar{P}_n^{m}(\cos \tilde{\theta}_k)\bar{P}_n^{m+1}(\cos \theta_j)}{\cos \tilde{\theta}_k - \cos \theta_j},
$$

thanks to the Christoffel–Darboux formula.

$$
\tilde{F}_m(\tilde{\theta}_k) \epsilon_{n+1}^{m} = \frac{\bar{P}_n^{m+1}(\cos \tilde{\theta}_k)}{\cos \tilde{\theta}_k - \cos \theta_j} \sum_j F_m(\theta_j)\omega_j \bar{P}_n^{m}(\cos \theta_j) - \bar{P}_n^{m}(\cos \tilde{\theta}_k) \frac{\epsilon_{n+1}^{m}}{\cos \theta_j - \cos \theta_j}.
$$

Such a computation can be seen as two matrix-vector products with matrix

$$
1 \cos \tilde{\theta}_k - \cos \theta_j.
$$

A one-dimensional FMM can be applied successfully to this type of matrix yielding an optimal algorithm (see the article by Yarvin and Rokhlin for one-dimensional FMMs [25] and also [26] and [11]).
A backward FFT is performed
\[ \hat{F}(\tilde{\phi}_l, \tilde{\theta}_k) = \frac{2\pi}{I} \sum_m \hat{F}_m(\tilde{\theta}_k)e^{im\tilde{\phi}_l}. \]

3. Asymptotic complexity.

3.1. Error analysis. Our goal here is to prove the asymptotic complexity of the FMM. We will see that the proof relies on two error estimates stated in Propositions 1 and 2.

**Proposition 1** (number of terms in the transfer function). For all \( v, u \in \mathbb{R}^+ \) such that
\[ \frac{u}{v} \geq \frac{\sqrt{5}}{2}. \]
For all \( m \geq 0 \), we denote the Gegenbauer sequence by
\[ G_m \overset{def}{=} (2m + 1)f_m(v)h_m^{(1)}(u)P_m(w). \]
Then for all \( \epsilon < 1 \), there exists four constants \( C_1, C_2, C_2', \) and \( C_3 \) such that
\[ m \geq M \overset{def}{=} C_1 + C_2v + C_2'\log v + C_3\log \epsilon^{-1} \implies \left| \sum_{k=m}^{+\infty} G_k \right| \leq \epsilon. \]
Thus we can use \( M \) terms to compute the transfer function.

**Remark.** This a much better result than is usually given in earlier papers since this condition is independent of \( u \), provided that \( u/v \geq \sqrt{5}/2 \) and depends only on \( v \). Moreover the different constants appearing in the previous theorem are independent of all the parameters of the problems. We thus proved a uniform convergence of the series for all the points in a cluster of a given size.

The second proposition can be proved once the previous one has been established. It corresponds to the error due to the fact that we can only consider a finite number of directions \( s_k \) in our computation. It gives us a criteria to calculate the number of points needed on \( S^2 \).

**Proposition 2** (number of directions on \( S^2 \)). We suppose that the hypotheses of Proposition 1 are satisfied. We denote by \( \epsilon > 1 \) some bound on the error.

There exists constants \( C_1, C_2, C_2', \) and \( C_3 \) and constants \( D_1, D_2, D_2', \) and \( D_3 \) such that, if for some level we keep \( M \) terms in the transfer function (see Proposition 1) and take enough points on the sphere to integrate exactly the first \( K \) spherical harmonics on \( S^2 \), then
\[ M \geq C_1 + C_2v + C_2'\log v + C_3\log \epsilon^{-1} \quad \implies \quad \left| \sum_{k=m}^{+\infty} G_k \right| \leq \epsilon. \]
Practically a good choice for \( K \) is \( K \geq 2M \).

**Remark.** Using a Gaussian quadrature for \( \theta \) and a uniform quadrature for \( \phi \) we know that to integrate exactly the first \( K = 2M \) spherical harmonics we need \( 2M^2 \) points on \( S^2 \). From Proposition 2 we deduce that the number of points \( s_k \), \( \#\{s_k\} \approx 2M^2 \), is roughly proportional to the number of points in the cluster for
a uniform distribution of points on the surface (see below for the definition of this term).

The two previous theorems were already found empirically by others (see [22], [10], and [24], for example). From numerical tests it is found that with \( M = v + 3. \log(v + \pi) \) the error is \( 10^{-3} \) while with \( M = v + 5. \log(v + \pi) \) the error is \( 10^{-6} \). This is in total accordance with the theory.

As a simplification we will suppose from now on that \( \epsilon \) has been set to some value and we will consider it as a constant. The reader will note that in practice the terms proportional to \( v \) are the dominant ones.

3.2. Asymptotic complexity theorem.

Definition 3.1. We call “uniform distribution” of points on the surface of an object a distribution such that there exists two constants \( \gamma_2 > \gamma_1 > 0 \) such that the distance between one point and its closest neighbor is greater than \( \gamma_1 \lambda \) and less than \( \gamma_2 \lambda \) where \( \lambda \) is the wavelength. Typically \( \gamma_1 \) and \( \gamma_2 \) are greater than \( 1/20 \) and less than \( 1/5 \).

Another algorithm developed by Greengard addresses the problem of \( \gamma_1 \) tending to zero. This is the case if there is an accumulation of points somewhere on the surface (see [13]).

This constant will fix some of the other constants appearing later in the paper. In particular, in the theorem below, we have the following.

Theorem 3.2 (asymptotic complexity theorem). We consider a set of points uniformly distributed on the surface of an object. Those points are partitioned using an oct-tree until the leaves of the tree have a size of a fraction of a wavelength (typically \( \kappa \times \text{Diam} \) is of order 1). Then using a fast interpolation/subsampling as described above we can perform a matrix-vector product \( M \sigma \) with complexity

\[ O(n \log^2 n), \]

where \( n \) is the size of the matrix and

\[ M_{(i,j)} = \frac{e^{i \kappa |x_i - x_j|}}{4\pi |x_i - x_j|}. \]

There is a minimum size for the leaves of the tree. Under a certain critical size the roundoff errors due to the explosion of the functions \( h^{(1)} \) prevent the use of the FMM on a finite precision computer (even with 16 digits). Moreover the interaction between close neighbors at the lowest level in the tree has to be computed in a traditional way. With a uniform distribution of points this additional cost is \( O(n) \) (the constant depends on the critical size of the leaves and on \( \gamma_1 \) and \( \gamma_2 \)). However, this cost is in \( O(n^2) \) if there is an accumulation of points somewhere in the mesh.

Note that the constant in \( O(n \log^2 n) \) also depends on \( \gamma_1 \) and \( \gamma_2 \). The true expression is actually (we denote by \( R \) the diameter of the object)

\[ O(n) + O\left((\kappa R)^2 \log^2(\kappa R)\right), \]

\( O(n) \) being the only function of \( \gamma_1 \) and \( \gamma_2 \). We have

\[ n \sim (R/\lambda)^2 \sim (\kappa R)^2. \]

Proof. The hypothesis of uniform distribution of the points is equivalent to the fact that \( |P|^2 \) is proportional to the number of points contained in \( P \).
We now consider all the operations to be performed for all the clusters of a given level. For one cluster, the transfer + interpolation + subsampling require $|P|2\log|P|$ operations (see Propositions 1 and 2). Thus for the clusters at a given level the complexity is proportional to $n\log n$. This is possible only with the use of fast interpolation methods.

Therefore the total complexity is $n\log^2 n$.

4. Gegenbauer sequence theorem. We now prove Propositions 1 and 2. They can be deduced from the following key theorem.

**Theorem 4.1 (Gegenbauer sequence theorem).** For all $v, u \in \mathbb{R}^+$ such that $1 < v < u$ and $m \in \mathbb{N}$. Let $p$ be the smallest integer greater than $v - 1/2$, $q$ the smallest greater than $u + 7/2$.

We denote by $f_k(v)$ and $g_k(u, v)$ the two following sequences:

$$f_k(v) = \frac{v/(2k + 1)}{1 - (2k+1)(2k+3)},$$

$$g_k(u, v) = \frac{1 - \frac{(2k-1)(2k-3)}{u^2} - (2k-1)(2k-5)}{1 - \frac{(2k+1)(2k+3)}{u^2} - 2(2k+1)(2k+5)}.$$  

For some constant $C > 1$

$$|G_m| \leq \begin{cases} 
\frac{C}{v^{5/6}}(2m + 1) \prod_{p=1}^{m} f_k(v) & \text{for } p \leq m \leq u - \frac{1}{2}, \\
\frac{C}{(uv)^{5/6}}(2m + 1) \prod_{p=1}^{m} f_k(v) & \text{for } u - \frac{1}{2} < m < q, \\
\frac{C}{(uv)^{5/6}}\left(\frac{v}{u}\right)^{m-q}(2q + 1) \prod_{p=1}^{q} f_k(v) \prod_{q+1}^{m} g_k(u, v) & \text{for } q \leq m.
\end{cases}$$

Moreover, with the additional hypothesis that $v/u \leq 2/\sqrt{5} < 1$,

$$k \geq q \Rightarrow 0 < g_k(u, v) < 1.$$  

A similar theorem may be proved for the case $v < 1$. The mathematical tools which have to be used when $v < 1$ are very similar to the ones presented here. However, this case is of limited practical interest since in numerical experiments $v$ is always greater than 1.

The behavior of $G_m$ is thus the following. For $u \geq m \geq v$, we observe a fast convergence. Indeed, $f_k(v)$ decreases very rapidly to 0. For $k \sim v$, $f_k(v) \sim v/k$, while for $k \gg v$, $f_k(v) \sim \frac{v}{2k}$. This is thus a fast convergence region.

For $m \geq u$, $G_m$ converges to zero like a geometrical sequence of ratio $v/u$.

4.1. Proof of the error analysis. First we notice that if the points on the surface are partitioned using an oct-tree then $v/u \leq \sqrt{3}/2 < 2/\sqrt{5}$. Thus $g_k(u, v) < 1$.

**Proof of Proposition 1.** We now suppose that $v \geq 1$.

1. **Convergence before $u + \frac{7}{2} \geq m$.** The first point consists in the following bound which simplifies the expression. Suppose that we have $u + \frac{7}{2} \geq m \geq
Then
\[ f_m(v) \leq \frac{8}{7} \frac{v}{2m+1}. \]
\[ |G_m| \leq \frac{8C}{7} v^{1/6} \prod_{k=p+1}^{m-1} f_k(v). \]

We have the following lemma.

**Lemma 4.2.** We denote
\[ F(\alpha) \stackrel{\text{def}}{=} \left( 1 + (\alpha - 1)(2 + \alpha^{-1}) \right)^{-1}. \]

\( F(\alpha) \) is a monotonous decreasing function with \( F(1) = 1 \). We have
\[ |f_k(v)| \leq F(v/k). \]

The proof is left to the reader.

We now have a bound on \( |G_k| \) for all \( q > k \geq 2v + 1 \):
\[ |G_k| \leq \frac{8C}{7} v^{1/6} F(2)^{-2v}. \]

Before using (3), which applies to \( q > k \), we bound the series for \( k \geq q \).

**Lemma 4.3** (bound for \( k \geq q \)). There exists a constant \( C \) such that
\[ \left| \sum_{k=q}^{+\infty} G_k \right| \leq CG_{q-1}. \]

We see that the proof is complete if we bound \( G_k \) for all \( k < q \). Indeed, for all \( m < q \)
\[ \left| \sum_{k=m}^{+\infty} G_k \right| \leq (1 + C) \left| \sum_{k=m}^{q-1} G_k \right|. \]

The bound for all \( k < q \) is obtained below after the proof of Lemma 4.3, with the help of (3).

**Proof.** For \( k \geq q \), thanks to Theorem 4.1, the sequence \( G_k \) may be bounded by
\[ G_k \leq C G_{q-1} \left( \frac{v}{u} \right)^{k-q}. \]

Thus
\[ \left| \sum_{k=q}^{+\infty} G_k \right| \leq C_1 G_{q-1} \frac{1}{1 - \frac{v}{u}}. \]

Since \( \frac{v}{u} < \frac{2}{\sqrt{5}} \), there is a constant \( C_2 \) such that
\[ \left| \sum_{k=q}^{+\infty} G_k \right| \leq C_2 G_{q-1}. \]
There exists a constant $r_0$ such that with the following definition

$$r_1 \overset{\text{def}}{=} r_0 - \frac{\log v}{6 \log F(2)}$$

we have

$$8C \frac{v^{1/6} F(2)^{r_1}}{7} \leq 1. \tag{4}$$

If $m \geq 2v + r_1$, we can use inequalities (3) and (4):

$$\left| \sum_{k=m}^{+\infty} g_k \right| \leq (1 + C_2) \sum_{k=m}^{q-1} F(2)^{k-2v-r_1} \leq (1 + C_2) \frac{F(2)^{m-2v-r_1} - F(2)^{q-2v-r_1}}{1 - F(2)} \leq C_3 F(2)^{m-2v-r_1}. \tag{5}$$

Thus we have converged as soon as

$$m \geq M \overset{\text{def}}{=} 2v - \frac{\log v}{6 \log F(2)} + r_2 + \frac{\log \epsilon}{\log F(2)} \Rightarrow \sum_{k=m}^{+\infty} g_k \leq \epsilon$$

and this is a fast convergence since $F(2) = \frac{2}{7} \approx 0.29$. The integer $r_2 \geq r_0$ accounts for the constant $C_3$ in estimation (5).

As a matter of fact this estimation is almost optimal up to a constant. This is illustrated by Figure 2. We see that this zone is a fast convergence zone.

2. Convergence after $u + \frac{7}{2} \leq m$. We have proved so far, roughly speaking, that there exists a constant $C$ such that if $u + 7/2 \geq Cv$ then convergence will take place for $Cv \leq M \leq u + 7/2$. This may not be the case if $u + 7/2 \leq Cv$. First let us get rid of the term $\frac{C}{(uv)^{5/6}} (2q + 1)$. With the remark above

$$\frac{C}{(uv)^{5/6}} (2q + 1) \leq \frac{C_1 u^{1/6}}{v^{5/6}} \leq \frac{C_2}{v^{4/6}} \tag{6}$$

because $q \sim u$ and $u + 7/2 \leq Cv$. Since $v \geq 1$ we can bound (6) by a constant.
Now since we have a geometrical series of ratio \( v/u \), we will have converged when

\[
m \geq u + \frac{7}{2} + \log \epsilon - C \log \frac{v}{u}.
\]

A sufficient condition is

\[
m \geq 2v - \frac{\log v}{6 \log F(2)} + r_2 + \frac{C}{\log \sqrt{5}/2} + \log \epsilon \left( \frac{1}{\log \frac{2}{\sqrt{5}}} + \frac{1}{\log F(2)} \right)
\]

because

\[
2v - \frac{\log v}{6 \log F(2)} + r_2 + \log \epsilon \log F(2) \geq u + \frac{7}{2}.
\]

Again the optimality is illustrated by Figure 3.

This concludes the proof.

**Proof of Proposition 2: Number of Points \( s_k \).** Thanks to Proposition 1, which can be applied in the case of an oct-tree decomposition of our points (hypothesis on \( v/u \) in particular), we proved that the Gegenbauer series can be truncated after \( m \geq C(|P| + |Q|) \) (we simplify this estimation for clarity) for all transfers between cluster \( |P| \) and \( |Q| \).

We consider two points \( x_i \) and \( x_j \) in \( P \) and \( Q \) and their interaction.

Up to now, the proof for the multistep algorithm is strictly identical to the proof for the single-step algorithm. The only difference is that interpolations and smoothing are performed from one level to the other in the multistep algorithm. The proof of the fact that the error introduced by those operations is smaller than some threshold \( \epsilon \) is very similar to the proof given now to evaluate the number of directions \( s_k \). Thus we will omit this proof to simplify and shorten the article. However we announce that such a proof will be posted on the web (see the author’s home page at www.ann.jussieu.fr).

Now we suppose that the interpolations and smoothing are performed exactly.

The two following formulae,

\[
e^{i \kappa(s,v)} = \sum_{n=+\infty} (2n + 1) j_n(v) P_n(\cos(s,V)),
\]

\[
\sum_{m} Y_{m}^l(s) Y_{m}^l(s') = \frac{2l + 1}{4\pi} P_l(\cos \gamma)
\]
\( \gamma \) is the angle between \( s \) and \( s' \), \( Y^l_m \) is the spherical harmonic), and the following property,

\[ Y^l_m \] is an orthonormal basis of \( L^2(S^2) \),

prove that

\[
\int \! ds \, P_m(\cos(s,U)) P_n(\cos(s,V)) = \delta_{m,n} \frac{4\pi}{2l+1} P_m(\cos(U,V)) \quad \text{and thus} \quad \int \! ds \, P_m(\cos(s,U)) e^{\imath \kappa(s,v)} = 4\pi \imath m j_m(v) P_m(\cos(U,V)).
\]

We proved (2) and obtained the following expansion for \( h^{(1)}_0 \):

\[
h^{(1)}_0(\kappa|s_i - s_j|) = \sum_{m \leq +\infty} \sum_{n \leq +\infty} (2m + 1)(2n + 1) \frac{4\pi}{4\pi j_m(v)} j_n(v) h^{(1)}_m(u) \int \! ds \, P_m(\cos(s,U)) P_n(\cos(s,V)).
\]

Using Proposition 1 we have some \( M \geq 2v \) such that we can truncate the series with an error \( \eta \):

\[
h^{(1)}_0(\kappa|s_i - s_j|) = \sum_{n \leq +\infty} \frac{(2n + 1) j_n(v)}{4\pi} \sum_{m \leq M} (2m + 1) \imath^{-m} h^{(1)}_m(u) \int \! ds \, P_m(\cos(s,U)) P_n(\cos(s,V)) + \eta.
\]

We have the following bound on \( \imath^{-m} h^{(1)}_m(u) \):

\[
|\imath^{-m} h^{(1)}_m(u)| \leq \begin{cases} 1 & \text{if } u + \frac{1}{2} > M, \\ |h^{(1)}_M(u)| & \text{if } u + \frac{1}{2} \leq M. \end{cases}
\]

Suppose we take enough points on \( S^2 \) so that we can integrate exactly the first \( K \) spherical harmonics with \( K \overset{\text{def}}{=} M + H, H \geq M \). Then

\[
I^C_{m,n} = \int \! ds \, P_m(\cos(s,U)) P_n(\cos(s,V)), \quad I^D_{m,n} = \sum_k \omega_k P_m(\cos(s_k,U)) P_n(\cos(s_k,V)),
\]

\[
h^{(1)}_0(\kappa|s_i - s_j|) = \tilde{h}^{M,K}_0(\kappa|s_i - s_j|) + \eta - \sum_{n > H} \frac{(2n + 1) j_n(v)}{4\pi} \sum_{m \leq M} (2m + 1) \imath^{-m} h^{(1)}_m(u) (I^C_{m,n} - I^D_{m,n}),
\]

where \( \tilde{h}^{M,K}_0(\kappa|s_i - s_j|) \) is the contribution to node \( i \) from node \( j \) as computed with the FMM. Since

\[
|I^C_{m,n} - I^D_{m,n}| \leq 8\pi,
\]
\[
\left| h_0^{(1)}(\kappa|x_i - x_j|) - \tilde{h}_0^{M,K}(\kappa|x_i - x_j|) \right|
\leq \eta + 2(M + 1)(2M + 1) \max(1, |h_M^{(1)}(u)|) \sum_{n > H} (2n + 1)|j_n(v)|.
\]

Using Proposition 3 we obtain
\[
|j_n(v)| \leq \frac{C}{\nu^{5/6}} \prod_{p+1} \prod f_k(v)
\]
for \( n + \frac{1}{2} \geq \nu. \)

We now suppose that \( M \) is such that \( M \geq 2
\nu. \)

1. **Case** \( M + \frac{1}{2} \leq \nu. \) Let us suppose, moreover, that \( M + \frac{1}{2} \leq \nu. \) The case \( M + \frac{1}{2} > \nu \) will be considered below. If \( M + \frac{1}{2} \leq \nu, \)

\[
|h_M^{(1)}(u)j_n(v)| \leq \frac{C}{\nu^{5/6}} \prod_{p+1} f_k(v).
\]

Using the fact that \( H \geq M \) and that
\[
(2M - 1)(2M + 1) f_{M-1}(v) f_M(v) \leq C \nu^2,
\]
\[
(2n + 1)f_n(v) \leq C \nu,
\]
we get, for some constant \( C_1, \)
\[
2(M + 1)(2M + 1) \max(1, |h_M^{(1)}(u)|) \sum_{n > H} (2n + 1)|j_n(v)|
\leq \frac{C_1 \nu^3}{\nu^{5/6}} \sum_{n > H} \prod_{p+1 \leq k < n; k \neq M} f_k(v)
\leq C_2 \nu^{13/6} \sum_{n > H} \prod_{p+1 \leq k < n; k \neq M} f_k(v).
\]

With Lemma 4.2 and with the techniques used for the proof of Proposition 1, we prove Proposition 2.

2. **Case** \( M \geq \nu. \) If \( M \geq \nu \) then (with Proposition 4)
\[
|h_M^{(1)}(u)j_n(v)| \leq \frac{C}{(uv)^{5/6}} \left( \frac{v}{u} \right)^{M-\nu} \frac{2q + 1}{2M + 1} \prod_{p+1} f_k(v) \prod_{M+1} f_k(v) \prod_{q+1} g_k(u,v).
\]

In the case of an oct-tree \( g_k(u,v) < 1. \) Moreover, with the same type of argument as for the proof of Proposition 1, we prove that
\[
\frac{2q + 1}{(uv)^{5/6}} \leq \frac{C_1}{\nu^{5/6}}.
\]
Thus we obtain, if $H \geq M + 1$,

$$
2(M + 1)(2M + 1) \max(1, |h_M^{(1)}(u)|) \sum_{n > H} (2n + 1)|j_n(v)|
\leq C_2 v^{2/6} \left( \frac{v}{u} \right)^{M-q} (M + 1) \sum_{n > H} \left( \prod_{k=p+1}^{q} f_k(v) \prod_{k=M+1}^{n-1} f_k(v) \right)
\leq C_3 v^{8/6} \left( \frac{v}{u} \right)^{M-q} \sum_{n > H} \left( \prod_{k=p+1}^{q} f_k(v) \prod_{k=M+2}^{n-1} f_k(v) \right).
$$

With the same type of arguments as for the proof of Proposition 1, the proof is complete.

5. Preliminary results: Spherical Bessel functions. In this section and the following we will prove the Gegenbauer convergence theorem.

We start with some basic observations regarding continued fractions (Lemmas 5.2 and 5.1).

Then we will observe that $j_k(v)/j_{k-1}(v)$ as well as $y_k(u)/y_{k-1}(u)$ are equal to simple continued fractions (Lemmas 5.3 and 5.4).

In the end (Propositions 3 and 4) we will derive from the previous results two bounds on $j_k(v)/j_{k-1}(v)$ and $y_k(u)/y_{k-1}(u)$.

These bounds are valid, respectively, for $k \gtrsim v$ and for $k \gtrsim u$. Thus in the last subsection we will look for bounds in the case, respectively, when $k \lesssim v$ and when $k \lesssim u$.

The Gegenbauer convergence theorem will be deduced from those results in the next section.

Continued fractions. We denote by

$$
b_0 + \frac{a_1}{b_1 + \frac{a_2}{b_2 + \frac{a_3}{b_3 + \cdots}}}$$

the following continued fraction:

$$
b_0 + \frac{a_1}{b_1 + \frac{a_2}{b_2 + \frac{a_3}{b_3 + \cdots}}}.
$$

We need two technical lemmas to proceed in our proof.

**Lemma 5.1.** Let $a_i$, $b_i$ be two sequences such that

$$
\begin{cases}
  a_i \leq 0, \\
  b_i \geq 0.
\end{cases}
$$

Then

$$
\begin{align*}
  & \text{for all } i \geq 0 \quad b_i b_{i+1} + 4a_{i+1} \geq 0 \\
  \Rightarrow & \text{ for all } i \geq 0 \text{ and } p \geq 0 \quad b_i + \frac{a_{i+1}}{b_{i+1}} + \cdots + \frac{a_{i+p}}{b_{i+p}} \geq \frac{1}{2} b_i.
\end{align*}
$$
Proof. It is true for \( k = i + p - 1 \) and \( p = 1 \) since
\[
b_{i+p-1} + \frac{a_{i+p}}{b_{i+p}} \geq \frac{1}{2} b_{i+p-1}.
\]

Then
\[
b_{i+p-2} + \frac{a_{i+p-1}}{b_{i+p-1}} + \frac{a_{i+p}}{b_{i+p}} \geq b_{i+p-2} + \frac{2a_{i+p-1}}{b_{i+p-1}} \geq \frac{1}{2} b_{i+p-2}.
\]

By recurrence from \( j = i + p - 1 \) to \( j = i \), the lemma can be proved.

Lemma 5.2. Let \( a_i, b_i \) and \( a'_i, b'_i \) be four sequences of \( \mathbb{R} \) such that
\[
\begin{align*}
& a_i \leq 0 \text{ and } a'_i \leq 0, \\
& b_i \geq 0 \text{ and } b'_i \geq 0, \\
& b_i b_{i+1} + 4a_{i+1} \geq 0 \text{ and } b'_i b'_{i+1} + 4a'_{i+1} \geq 0.
\end{align*}
\]

Then
\[
\begin{align*}
& a_i \leq a'_i \\
& b_i \leq b'_i
\end{align*}
\]

\[
\Rightarrow \quad b_0 + \frac{a_1}{b_1 + b_2 + b_3} + \frac{a_2}{b_1 + b_2 + b_3} + \frac{a_3}{b_1 + b_2 + b_3} + \cdots.
\]

Proof. Lemma 5.2 is a direct consequence of Lemma 5.1. We suppose that the hypotheses of the lemma are true for \( a_i, b_i \) and \( a'_i, b'_i \).

For all \( x \geq 0 \) and \( x' \geq 0 \) such that \( x \leq x' \),
\[
\begin{align*}
& a_i x' \leq a'_i x', \\
& a'_i x' \leq a'_i x, \\
& \Rightarrow a_i x' \leq a'_i x, \\
& \Rightarrow \frac{a_i}{x} \leq \frac{a'_i}{x'}.
\end{align*}
\]

Thus for all \( i \geq 0 \)
\[
b_i + \frac{a_i}{x} \leq b'_i + \frac{a'_i}{x'}.
\]

For all \( n \geq 0 \), we have that \( b_{n-1} + a_n/b_n \leq b'_{n-1} + a'_n/b'_n \).

Because (8) is true for \( a_i, b_i \) and \( a'_i, b'_i \), a recurrence from \( j = n - 1 \) to \( j = 0 \) proves the lemma.

5.1. Continued fraction: \( j_n/j_{n-1} \).

Lemma 5.3. For all \( n \in \mathbb{N} \) and \( z \in \mathbb{C} \) we have
\[
\frac{j_n(z)}{j_{n-1}(z)} = \frac{1}{\frac{2n+1}{z} - \frac{2n+3}{z} - \frac{2n+5}{z} - \cdots}.
\]

Proof. We prove it by recurrence.
\[
\frac{j_0(z)}{z} \quad \text{and} \quad \frac{j_1(z)}{j_0(z)} = \frac{\sin z}{z^2} - \frac{\cos z}{z} \Rightarrow \frac{j_1(z)}{j_0(z)} = \frac{1}{z} - \frac{1}{\tan z}.
\]
The function \( \tan z \) is equal to the following continued fraction:

\[
\tan z = \frac{z}{1 - \frac{z^2}{3 - \frac{z^2}{5 - \cdots}}} = \frac{1}{z^{-1} - \frac{1}{3z^{-1} - \frac{1}{5z^{-1} - \cdots}}}
\]

Thus the lemma is true for \( n = 1 \):

\[
\frac{j_1(z)}{j_0(z)} = \frac{1}{3z^{-1} - \frac{1}{5z^{-1} - \cdots}}.
\]

Using recurrence relation

\[
\frac{j_{n-1}(z)}{j_n(z)} + \frac{j_{n+1}(z)}{j_n(z)} = \frac{2n + 1}{z}
\]

we have

\[
\frac{j_{n-1}(z)}{j_n(z)} + \frac{j_{n+1}(z)}{j_n(z)} = \frac{2n + 1}{z}
\]

from which the lemma can be proved for all \( n \).

**5.2. Continued fraction: \( y_n/y_{n-1} \).** A similar lemma can be proved for \( y_n \).

**Lemma 5.4.** For all \( n \in \mathbb{N} \) and \( z \in \mathbb{C} \) we have

\[
\frac{y_{n-1}(z)}{y_n(z)} = \frac{1}{2n-1} - \frac{1}{2n-3} - \frac{1}{2n-5} - \cdots.
\]

**Proof.** We prove it by recurrence.

\[
y_0(z) = -\frac{\cos z}{z} \quad \text{and} \quad y_1(z) = -\frac{\cos z}{z^2} - \frac{\sin z}{z} \quad \Rightarrow \quad \frac{y_1(z)}{y_0(z)} = \frac{1}{z} + \tan z.
\]

Thus the lemma is true for \( n = 1 \):

\[
\frac{y_1(z)}{y_0(z)} = z^{-1} + \frac{1}{z^{-1} - \frac{1}{3z^{-1} - \cdots}} = z^{-1} - \frac{1}{-z^{-1} - \frac{1}{-3z^{-1} - \cdots}}.
\]

Using recurrence relation

\[
y_{n-1}(z) + y_{n+1}(z) = \frac{2n + 1}{z} y_n(z)
\]

we prove the lemma for all \( n \).

**5.3. Continued fraction with constant coefficient.**

**Lemma 5.5.** For all \( t \in \mathbb{R}^+ \)

\[
2 \cosh t - \frac{1}{2 \cosh t - \frac{1}{2 \cosh t - \cdots}} = \exp t.
\]

**Proof.** If we denote by \( f_n \)

\[
f_n = b_0 + \frac{a_1}{b_1} + \ldots + \frac{a_n}{b_n}.
\]
Then, we can define two sequences \( A_n \) and \( B_n \) by

\[
\begin{align*}
A_{-1} &= 1 \text{ and } A_0 = b_0, \\
B_{-1} &= 0 \text{ and } B_0 = 1, \\
A_n &= b_n A_{n-1} + a_n A_{n-2}, \\
B_n &= b_n B_{n-1} + a_n B_{n-2},
\end{align*}
\]  

such that

\[
f_n = \frac{A_n}{B_n},
\]

with, for all \( n \),

\[
\begin{align*}
b_n &= 2 \cosh t, \\
a_n &= -1.
\end{align*}
\]  

We have the following relation \( B_n = A_{n-1} \). Thus with (12)

\[
f_n = 2 \cosh t - \frac{1}{f_{n-1}} \Rightarrow \lim_{n \to \infty} f_n = \exp t. \quad \Box
\]

We have an immediate corollary.

**Corollary 1.** For all \( x \in \mathbb{R} \)

\[
x \geq 1 \Rightarrow x^{-1} - \frac{1}{x - x^{-1}} - \frac{1}{x - x^{-1} - x - x^{-1} - \cdots} = 0.
\]

**Proof.** Consider \( x = \exp t, 2 \cosh t = x + x^{-1} \). Thus

\[
x + x^{-1} - \frac{1}{x - x^{-1}} - \frac{1}{x - x^{-1} - x - x^{-1} - \cdots} = x.
\]

**5.4. Bound on \( j_n/j_{n-1} \).**

**Proposition 3.** For all \( n \) and \( x \in \mathbb{R}^+ \)

\[
G(\alpha, n, x) = \left( 1 - \left( \frac{(2n+1)(2n+3)}{x^2} - \frac{\alpha(2n+1)}{2n+5} \right)^{1/2} \right)^{-1},
\]

\[
n + \frac{1}{2} \geq x \Rightarrow \\
\frac{x}{2n+1} G(1, n, x) \leq \frac{j_n(x)}{j_{n-1}(x)} \leq \frac{x}{2n+1} G(2, n, x),
\]

\[
\frac{x/(2n+1)}{1 - x^2/(2n+1)(2n+3)} \leq \frac{j_n(x)}{j_{n-1}(x)} \leq \frac{x/(2n+1)}{1 - x^2/(2n+1)(2n+3)}.
\]

The second inequality is not as good as the first one but the expression is more simple.

**Proof.** We consider the continued fraction of Lemma 5.3. It satisfies the hypothesis of Lemmas 5.2 and 5.1. Hence the bounds. \( \Box \)
5.5. Bound on $y_n/y_{n-1}$.

**Proposition 4.** For all $n$ and $x \in \mathbb{R}^+$

$$F(\alpha, n, x) = 1 - \left(\frac{(2n-1)(2n-3)}{x^2} - \frac{\alpha(2n-1)}{2n-5}\right)^{-1},$$

$$2n - 1 \geq x \Rightarrow \frac{2n-1}{x} F(2, n, x) \leq \frac{y_n(x)}{y_{n-1}(x)} \leq \frac{2n-1}{x} F(1, n, x).$$

**Proof.** We cannot apply Lemmas 5.1 and 5.2 this time because the continued fraction (Lemma 5.4)

$$\frac{1}{z} - \frac{1}{z} - \frac{1}{z} - \cdots$$

does not satisfy the hypothesis.

However, let $p \in \mathbb{N}$ be the smallest integer such that $p - \frac{1}{2} \geq x$.

We have the following recurrence relation which can be found in [1, Equation (10.1.22)]:

$$\frac{k}{x} y_k(x) - \frac{dy_k}{dx}(x) = y_{k+1}(x) \Rightarrow \frac{k}{x} - \frac{dy_k}{dx}(x) = \frac{y_{k+1}(x)}{y_k(x)}.$$

The inequality $k + \frac{1}{2} \geq x$ implies that $y_k(x) \leq 0$. Therefore, if $k + \frac{1}{2} \geq x$,

$$y_{k+1}(x)/y_k(x) \geq 1 \Rightarrow \frac{dy_k}{dx}(x) \geq 0.$$

The study of the zeros of $Y_\nu$ and $Y_\nu'$ shows that

$$\frac{dy_k}{dx}(x) = \sqrt{\frac{\pi}{2x}} \left(Y_{k+\frac{1}{2}}'(x) - \frac{1}{2x} Y_{k+\frac{1}{2}}(x)\right)$$

is always positive for $k + \frac{1}{2} \geq x$.

Therefore the sign of $y_p(x)/y_{p-1}(x)$ is constant and $y_p(x)/y_{p-1}(x) \geq 1$. For all $n \geq p + 3$, thanks to Lemma 5.4,

$$\frac{y_n(z)}{y_{n-1}(z)} = \frac{2n-1}{z} - \frac{1}{z} - \frac{1}{z} - \cdots - \frac{1}{y_p/y_{p-1}}.$$

Now we can apply Corollary 1:

$$\frac{y_n(z)}{y_{n-1}(z)} = \frac{2n-1}{z} - \frac{1}{z} - \frac{1}{z} - \cdots - \frac{1}{y_p/y_{p-1} + y_{p-1}/y_p} - \frac{1}{y_p/y_{p-1} + y_{p-1}/y_p}.$$

Because $y_p(x)/y_{p-1}(x) \geq 1$, we have

$$y_p/y_{p-1} + y_{p-1}/y_p \geq 2.$$

Therefore the hypothesis of Lemmas 5.2 and 5.1 are satisfied. Thus

$$\frac{y_n(x)}{y_{n-1}(x)} \leq \frac{2n-1}{x} F(1, n, x).$$

and

$$\frac{2n-1}{x} F(2, n, x) \leq \frac{y_n(x)}{y_{n-1}(x)}.$$
5.6. Bound on $j_n$ and $y_n$ for $n \leq x$.

**Proposition 5.** For all $z \in \mathbb{C}$ and $n \in \mathbb{N}$

$$n + \frac{1}{2} \leq |z| \Rightarrow \begin{cases} |j_n(z)| \leq \frac{\sqrt{\pi}}{|z|}, \\ |y_n(z)| \leq \frac{\sqrt{\pi}}{|z|}. \end{cases}$$

In particular they are bounded by 1.

**Proof.** From Formulae (10.1.26) and (10.1.27) in [1], we have

$$\begin{align*}
  j_n(z) &= \sqrt{\frac{\pi}{z}}M_{n+\frac{1}{2}}(z) \cos \theta_{n+\frac{1}{2}}(z), \\
  y_n(z) &= \sqrt{\frac{\pi}{z}}M_{n+\frac{1}{2}}(z) \sin \theta_{n+\frac{1}{2}}(z),
\end{align*}$$

where

$$\left(\frac{1}{2} \frac{\pi}{z}\right) M_{n+\frac{1}{2}}^2(z) = \frac{1}{z^2} \sum_{k=0}^{n} \frac{(2n-k)!(2n-2k)!}{k!(n-k)!^2} \frac{1}{(2z)^{2n-2k}}.$$

We now prove that

$$\left|\left(\frac{1}{2} \frac{\pi}{z}\right) M_{n+\frac{1}{2}}^2(z)\right| \leq \frac{n}{|z|^2}$$

from which Proposition 5 results.

Suppose the following lemma has been proved.

**Lemma 5.6.** For all $z \in \mathbb{C}$ and $k, n \in \mathbb{N}$ such that $0 \leq k \leq n$,

$$n + \frac{1}{2} \leq |z| \Rightarrow \frac{(2n-k)!(2n-2k)!}{k!(n-k)!^2} \frac{1}{(2z)^{2n-2k}} \leq 1. \tag{13}$$

With this lemma

$$\left|\left(\frac{1}{2} \frac{\pi}{z}\right) M_{n+\frac{1}{2}}^2(z)\right| \leq \frac{1}{|z|^2} \sum_{k=0}^{n} 1$$

which proves the first part of Proposition 5. Moreover, $|z|^2 \geq |z| - \frac{1}{2}$ for all $z \in \mathbb{C}$. Since $|z| \geq n + \frac{1}{2} \geq \frac{1}{2}$, $|z| - \frac{1}{2}$ is positive, the following identity holds:

$$\frac{\sqrt{|z| - \frac{1}{2}}}{|z|} \leq 1.$$

Therefore $j_n$ and $y_n$ are bounded by 1. \(\square\)

**Proof of Lemma 5.6.** First we bound $(2n-2k)!/[(n-k)!]^2$.

$$(2n-2k)!/[(n-k)!]^2 = \frac{2^{n-k}(n-k)!(2n-2k-1) \cdots 5.3.1}{[(n-k)!]^2}$$

$$= 4^{n-k} \frac{(2n-2k-1) \cdots 5.3.1}{(2n-2k) \cdots 6.4.2}.$$
Thus

\[
(2n - 2k)! / [(n - k)!]^2 \leq 4^{n-k}.
\]

We can bound the expression in (13):

\[
\frac{(2n - k)!((2n - 2k)!}{k![(n - k)!]^2} \frac{1}{(2|z|)^{2n-2k}} \leq \frac{(2n - k)!}{k!|z|^{2n-2k}}.
\]

We can rewrite the product as

\[
\frac{(2n - k)!}{k!|z|^{2n-2k}} = \prod_{1}^{2n-2k} \frac{k + p}{|z|} \prod_{1}^{n-k} \frac{n + p}{|z|}
\]

\[
= \prod_{1}^{n-k} \frac{(n + p)(n + 1 - p)}{|z|^2}
\]

\[
= \prod_{1}^{n-k} \frac{(n + \frac{1}{2} + (p - \frac{1}{2}))(n + \frac{1}{2} - (p - \frac{1}{2}))}{|z|^2}
\]

\[
= \prod_{1}^{n-k} \frac{(n + \frac{1}{2})^2 - (p - \frac{1}{2})^2}{|z|^2} \leq \prod_{1}^{n-k} \left( \frac{n + \frac{1}{2}}{|z|} \right)^2.
\]

The lemma is proved.  

6. Convergence of the Gegenbauer series. The main trick of the proof is the following. Instead of trying to prove a uniform bound on \( j_m \) and \( y_m \), which is a very difficult task, we use the following “expansion”:

\[
j_m = j_p \prod_{k=p+1}^{m} \frac{j_k}{j_{k-1}},
\]

\[
y_m = y_q \prod_{k=q+1}^{m} \frac{y_k}{y_{k-1}}.
\]

Lemma 6.1 will provide us with a good bound on \( j_p(v) \) and \( y_q(u) \). Then, using Propositions 3 and 4, we obtain bounds on \( j_m(v) \) and \( y_m(u) \). Finally, with the next subsection we will prove that \( g_k(u, v) \) is smaller than 1 when \( v/u \leq 2/\sqrt{5} \).

The proof of Theorem 4.1 results from all these elements.  

As a summary of our results, we notice that, with the bounds of Propositions 3 and 4, we proved that \( j_m(v)y_m(u) \) has a very simple behavior. For \( v \lesssim m \lesssim u \), we can roughly state that \( j_m(v)y_m(u) \) decrease faster than a “geometrical” series of ratio \( v/(2m + 1) \). This a very fast convergence region. On the contrary for \( m \gtrsim u \), this sequence behaves like a geometrical series of ratio \( v/u \) which is a much slower convergence. This is summarized on Figure 4.

**Lemma 6.1 (estimation of \( j_p(v) \) and \( y_q(u) \)).** With the hypothesis of Theorem 4.1, there exists some constant \( C > 0 \) such that

\[
|j_p(v)| \leq C \frac{1}{v^{1-1/6}},
\]

\[
|y_q(u)| \leq C \frac{1}{u^{1-1/6}}.
\]
Proof. From [1], we have two lemmas.

**Lemma 6.2 (bound on \( J_\nu(\nu) \) and \( Y_\nu(\nu) \)).** The following asymptotic estimations are valid for \( \nu \to +\infty \):

\[
\begin{align*}
J_\nu(\nu) &\sim \frac{2^{1/3}}{3^{2/3} \Gamma(2/3)} \frac{1}{\nu^{1/3}}, \\
Y_\nu(\nu) &\sim -\frac{2^{1/3}}{3^{1/6} \Gamma(2/3)} \frac{1}{\nu^{1/3}}.
\end{align*}
\]

(14)

We also have

\[
\begin{align*}
J_\nu(\nu + z\nu^{1/3}) &= 2^{1/3} \nu^{-1/3} Ai(-2^{1/3} z) + O(\nu^{-1}), \\
Y_\nu(\nu + z\nu^{1/3}) &= -2^{1/3} \nu^{-1/3} Bi(-2^{1/3} z) + O(\nu^{-1}),
\end{align*}
\]

(15)

where \( Ai \) and \( Bi \) are the Airy functions.

The second set of equations give sharper asymptotic estimates. Recall that

\[
\begin{align*}
j_n(z) &= \sqrt{\frac{2}{\pi z}} J_{n+\frac{1}{2}}(z), \\
y_n(z) &= \sqrt{\frac{2}{\pi z}} Y_{n+\frac{1}{2}}(z).
\end{align*}
\]

For all \( v \in \mathbb{R}^+ \), \( v \leq n + \frac{1}{2} \), \( J_{n+\frac{1}{2}}(v) \) is an increasing function (see zeros in 9.5.2 of [1]) of \( v \). Thus, there exists \( C > 1 \)

\[
J_{n+\frac{1}{2}}(v) \leq J_{n+\frac{1}{2}}(n + \frac{1}{2}) \leq C \left( n + \frac{1}{2} \right)^{-1/3}
\]

with (14).

For \( j_p(v) \) we have, for some \( C \)

\[
|j_p(v)| \leq \frac{C}{v^{1-1/6}}.
\]

For \( |y_q(u)| \) we cannot use the same proof since this is a decreasing function of \( u \) and thus cannot be bounded by \( |y_q(q + \frac{1}{2})| \). We need a more powerful estimate for \( y_n(z) \) given by (15).

Let us consider \( z = -5 \) for example. We denote \( \nu = q + \frac{1}{2} \). Then

\[
\nu + z\nu^{1/3} \leq q - 9/2 \quad (\nu \geq 1).
\]
We now use the monotonicity of $Y_\nu$ and for some $C > 1$

$$|y_q(u)| \leq \sqrt{\frac{\pi}{2|u|}}|Y_q + \frac{1}{2}(q - 9/2)|$$ since $q - 7/2 \geq u > q - 9/2$

$$\leq \sqrt{\frac{\pi}{2|u|}|Y_{q + \frac{1}{2}}(\nu + z\nu^{1/3})|}$$

$$\leq C' \sqrt{\frac{\pi}{2|u|}|u|^{-1/3}} \leq \frac{C}{|u|^{1-1/6}}. \quad \square$$

6.1. Bound on $g_k(u, v)$. In a code based on an oct-tree decomposition, $v$ is equal to the diameter of a cluster and $u$ is the smallest distance between the center of two “sufficiently” far apart clusters: such clusters cannot have in common a point or an edge. Thus if we choose to pave the space with cubes, we readily have

$$v \leq \sqrt{3}D, \quad u \geq 2D,$$

where $D$ is the length of an edge. Hence a realistic bound on $u/v$ is indeed

$$\frac{u}{v} \geq \frac{2}{\sqrt{3}}.$$

We prove that $g_k(u, v) \leq 1$ if $u/v \geq \sqrt{5}/2$ which is a sufficient condition in our case, since $\frac{2}{\sqrt{3}} \geq \frac{\sqrt{5}}{2}$.

Bounding $g_k(u, v)$ by 1 is equivalent to

$$\frac{(2k + 1)(2k + 3)}{v^2} - \frac{(2k - 1)(2k - 3)}{u^2} \geq \frac{2(2k + 1)}{2k + 5} - \frac{2k - 1}{2k - 5} = 1 - \frac{24k - 20}{4k^2 - 25}.$$

We prove a result that is sufficient to complete the proof

$$\frac{(2k + 1)(2k + 3)}{v^2} - \frac{(2k - 1)(2k - 3)}{u^2} \geq 1$$

since

$$\frac{24k - 20}{4k^2 - 25} \geq 0.$$

This is proved by

$$\frac{(2k + 1)(2k + 3)}{v^2} - \frac{(2k - 1)(2k - 3)}{u^2} \geq (2k - 1)(2k - 3) \left( \frac{1}{v^2} - \frac{1}{u^2} \right)$$

$$\geq 4 \left( \frac{k - 3/2}{u} \right)^2 \left( \frac{u^2}{v^2} - 1 \right) \quad \text{(since } k - 3/2 \geq u)$$

$$\geq 4 \left( \frac{u^2}{v^2} - 1 \right) \geq 1 \quad \text{(since } u/v \geq \sqrt{5}/2). \quad \square$$
7. Previous error estimations. We state Rahola’s lemma (see Rahola’s article [20]).

Lemma 7.1 (Rahola’s Lemma 4.1). For each $\kappa$ and $u$ there exists a constant $C$ such that the truncation error of the dipole potential is bounded by

$$\epsilon_T \leq \frac{C}{u-v} \left(\frac{v}{u}\right)^{L+1},$$

where $L$ is the number of terms in the truncated expansion, $u > v$, and $L > u$.

Similar lemmas can be found in [21] and [18].

Our new results improves on Rahola’s for two reasons:

- Rahola’s error estimation is valid when the number of poles in the multipole expansion, $L$, is greater than the distance between two clusters. However, our estimation is valid as soon as the convergence takes place, that is, as soon as $L$ is greater than the diameter of a cluster. This is an important improvement since we proved that the error and the convergence does not depend on the distance between the two clusters but rather on their size. Let us introduce some notations:

  - $v = \kappa \times$ Diameter of a cluster,
  - $u = \kappa \times$ Distance between two clusters,
  - $v < u$.

Grosso modo, we can truncate the expansion after

$$L > \begin{cases} 
  u + C \frac{\log \epsilon}{\log(v/u)} & \text{in Rahola’s theorem,} \\
  C.v + C' \log \epsilon & \text{in Theorem 1.}
\end{cases}$$

- A major improvement is that in our theorems all the constants are true constants and depend on no parameter of the problem (position of the points, distance between two clusters, or size of the cluster). In Rahola’s article this limitation was emphasized. Lemma 4.1 stated that there exists a “constant” $C$ such that the error in the FMM is bounded by a function of $C$, $u$ and $v$. The “constant” $C$ is independent of the number of poles taken in the multipole expansion but it depends on the distance $u$ between two clusters. Thus it cannot be used when we need constants valid for all $u$ and $v$, that is, for an implementation which allows any configuration (shape, size) of the object. On the contrary our estimations involve true constants and thus can be used safely.

Rahola uses the following expansion for $j_l$:

$$j_l(v) = \frac{v^l}{1.3.5 \cdots (2l+1)} \left(1 - \frac{\frac{1}{2} v^2}{1!(2l+3)} + \frac{\frac{1}{2} v^2}{2!(2l+3)(2l+5)} - \cdots \right),$$

and the following approximation

(16) $$j_l(v) \sim \frac{v^l}{1.3.5 \cdots (2l+1)} e^{-v^2/4l},$$

but this approximation is only valid if $l \gg 1$ and if the infinite sum can be approximated by its first terms, i.e., if $v^2/4l \lesssim 1$. The same type of
Fig. 5. Comparison of our error estimation with the other papers, $C = 1$.

The approximation is used for $y_l$:

$$y_l(u) = -\frac{1.3.5\cdots(2l-1)}{u^{l+1}} \left(1 + \frac{1}{2!}u^2 + \frac{(1/2)^2}{2!} + \cdots\right)$$

$$= -\frac{1.3.5\cdots(2l-1)}{u^{l+1}} \sum_r \frac{(1/2)^r}{r!(2l-1)(2l-3)\cdots(2l-2r+1)}$$

$$\sim -\frac{1.3.5\cdots(2l-1)}{u^{l+1}} e^{u^2/4l}.$$  

However, the infinite sum includes terms that are decreasing at the denominator

$$2l-1, 2l-3, \ldots, 2l-2r+1, \ldots, 1, -1, \ldots, -\infty.$$  

Thus the approximation by $e^{u^2/4l}$ is not justified unless $l \to +\infty$. The domain of validity of

$$y_l(u) \sim -\frac{1.3.5\cdots(2l-1)}{u^{l+1}}$$

cannot be assessed for finite $l$.

Figure 5 illustrates the fact that the constant $C$ in Rahola’s lemma depends on $u$.

Here we consider two vectors $x$ and $y$ with the relation $|y| = 0.7|x|$ and we compute the error

$$\epsilon_{\text{true}}(x,y) = \left| h_0^{(1)}(|x-y|) - \sum_{l=0}^{L} (2m+1) h_l^{(1)}(|x|) j_l(0.7|x|) P_l(\cos \gamma) \right|$$

for $L = |x| + 1$. We plot, as a function of $|x|$, $\epsilon_{\text{true}}(x,y)$ as well as $\epsilon_{\text{Rahola}}(x,y)$ (Rahola’s Lemma 4.1 with $C = 1$) and $\epsilon_{\text{ours}}(x,y)$ (Theorem 1). This figure confirms that there exists $C$ such that

$$\epsilon_{\text{true}}(x,y) \leq C\epsilon_{\text{ours}}(x,y).$$

However,

$$\frac{\epsilon_{\text{true}}(x,y)}{\epsilon_{\text{Rahola}}(x,y)}$$
cannot be bounded by a constant independent of $|x|$.

Similar remarks can be made regarding Koc, Song, and Chew in [18] as is illustrated by Figure 6. They use the following bound on $j_l$ and $h_l^{(1)}(u)$:

\begin{equation}
\begin{aligned}
|j_l(v)| &\leq \sqrt{\frac{e}{2}} \left(\frac{|v|}{e} \right)^l, \\
|h_l^{(1)}(u)| &\leq \frac{1}{|u|} \sqrt{\frac{2}{e}} \left(\frac{2l + 1}{e|u|} \right)^l.
\end{aligned}
\end{equation}

While the first bound is correct and can be found in [1, Equation (9.1.62)] the second one is not valid for all $l$ and certainly not for $l \sim |u|$. As a matter of fact the opposite seems to be true as can be seen on Figure 7.

Estimate (18) is used to bound the remaining terms in the series from $l \sim |u|$ to $+\infty$. However, this estimate is valid (up to a constant) only for $l \gg |u|$. This error is illustrated by Figure 6 where we chose $|u| + 5 = l$ and plot $|h_{l+5}^{(1)}(l)|$ and compare with the Koc, Song, and Chew estimate.

8. **Numerical results.** An article will be written about the implementation and performance of the FMM. In this article we only give a limited presentation of a few numerical results.
The benchmark chosen, called “CETAF,” was provided by Dassault (see mesh on Figure 8). It is a real life application and presents some numerical difficulties, such as the presence of a slit through the mesh and sharp edges. Reflection from one side of the slit to the other can be observed. Triangles on either side are strongly interacting. All these phenomena result in a bad condition number and a slower convergence.

The iterative scheme used is GMRES and the preconditioner is SPAI. We used the combined field integral equation (CFIE) which has nicer convergence property than EFIE or MFIE. This leads to a nonsymmetric matrix. This is why we chose to use GMRES as our iterative scheme. The reader interested in SPAI will read [9], [5], [4], [16], and [2]. SPAI was tested on a sphere. It was observed that a remarkable improvement (factor of 4-5) was obtained for the EFIE formulation. However, the improvement with CFIE was marginal (20%-30% improvement). As far as the author knows no preconditioner is known to be more efficient on a general configuration for electromagnetics problems.

The code was run on a HP PA 8000 200 MHz computer, on a single processor. The peak rate is 800 MFlops (two addition/multiplication units).

A comparison with a brute force resolution (= exact solution) showed that the relative error for the FMM (Euclidean norm) is less than 0.01. The “exact” solution was found by assembling the full matrix and inverting it with the Gauss method or using an iterative method with the full matrix. However, the largest test case for which the full matrix can be assembled corresponds to 5000 degrees of freedom. We found that for problems of size less than 5000 degrees of freedom the error is less than
<table>
<thead>
<tr>
<th>CETAF</th>
<th>10λ</th>
<th>11.7λ</th>
<th>25λ</th>
</tr>
</thead>
<tbody>
<tr>
<td>FMM CPU Time</td>
<td>1.44h</td>
<td>1.84h</td>
<td>12.81h</td>
</tr>
<tr>
<td>FMM CPU / Iter</td>
<td>1.55 min</td>
<td>2.35 min</td>
<td>13.03 min</td>
</tr>
<tr>
<td>FMM Memory</td>
<td>22.3 Megs</td>
<td>37.0 Megs</td>
<td>173 Megs</td>
</tr>
<tr>
<td># levels</td>
<td>5</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>Degrees of freedom</td>
<td>21564</td>
<td>45960</td>
<td>183840</td>
</tr>
<tr>
<td>Total CPU</td>
<td>1.71h</td>
<td>2.42h</td>
<td>21.7h</td>
</tr>
<tr>
<td>SPAI (construction)</td>
<td>3.15 min</td>
<td>14.16 min</td>
<td>37.29 min</td>
</tr>
<tr>
<td>Resolution</td>
<td>1.57h</td>
<td>1.99h</td>
<td>20.63h (7.16h)</td>
</tr>
<tr>
<td>Total memory</td>
<td>80 Megs</td>
<td>210 Megs</td>
<td>550 Megs</td>
</tr>
<tr>
<td># GMRES iterations</td>
<td>56</td>
<td>47</td>
<td>59</td>
</tr>
</tbody>
</table>

Fig. 9. CETAF test case.

0.01, but it was not possible to check the accuracy for the larger test cases shown here.

The figure in parentheses ("20.63h(7.16h)," Figure 9) indicates that the close interactions were not stored in core memory but were recomputed at each iteration. Thus for the 25λ case, 7.16h were spent recomputing the close interactions.

To visualize the asymptotic growth of the complexity and memory requirement we plot the CPU and memory for the 3 frequencies and compare them with the predicted $n \log^2(n)$ complexity for the CPU and $n \log(n)$ complexity for the memory (Figure 10). The intensity of the current on both sides of the CETAF is shown in Figures 11 and 12. The real part of the complex current (physical current) is shown in Figure 13.

9. Conclusion. In this article we have presented a theoretical tool to study the spherical Bessel functions and bound sequences and series involving these Bessel functions. From that we obtained an optimal and sharp criteria to truncate the
Gegenbauer series and thus correctly define the number of terms in transfer functions. This gives us a control on the error due to the FMM. Moreover we deduced from those results the asymptotic complexity of the FMM. All the results, truncation of the Gegenbauer series, convergence of the FMM, complexity of the algorithm were verified and measured on a complete Maxwell code based on the FMM. We were able to develop this Maxwell code with an optimal, i.e., minimal, complexity and reach
a million of unknowns on a parallel computer. Those numerical results will be the object of a forthcoming paper.

REFERENCES