ERROR ANALYSIS FOR THE NUMERICAL EVALUATION OF THE DIAGONAL FORMS OF THE SCALAR SPHERICAL ADDITION THEOREM*

S. KOC†, JIMING SONG‡, AND W. C. CHEW‡

Abstract. The numerical solution of wave scattering from large objects or from a large cluster of scatterers requires excessive computational resources and it becomes necessary to use approximate—but fast—methods such as the fast multipole method; however, since these methods are only approximate, it is important to have an estimate for the error introduced in such calculations. An analysis of the error for the fast multipole method is presented and estimates for truncation and numerical integration errors are obtained. The error caused by polynomial interpolation in a multilevel fast multipole algorithm is also analyzed. The total error introduced in a multilevel implementation is also investigated numerically.

Key words. fast multipole method, multilevel fast multipole algorithm, error analysis, truncation error, integration error, interpolation error

AMS subject classifications. 65G99, 78A40

PII. S0036142997328111

1. Introduction. The major computational task in scattering problems is the solution of dense systems of linear equations. When iterative methods are employed to solve these systems, the computation of the product of the coefficient matrix with a trial vector is the basic part of the algorithm and would require $\mathcal{O}(N^2)$ operations with traditional methods. The matrix vector product can be considered as the evaluation of all pairwise interactions between scatterers or parts of a scatterer. The fast multipole method (FMM) was proposed by Greengard and Rokhlin [1] for particle simulations and later extended for acoustic and electromagnetic scattering calculations [2, 3, 4, 5, 6, 7]. By using a multilevel fast multipole algorithm (MLFMA), all pairwise interactions can be evaluated with $\mathcal{O}(N)$ complexity for a volume distribution of scatterers.

The FMM/MLFMA algorithm is based on the diagonal forms of the spherical addition theorem. These diagonal forms are given as integrals of truncated infinite sums. In the numerical evaluation of these expressions, two kinds of errors are introduced. First, the infinite expansions are truncated at some finite value. Second, the integrals are evaluated by a numerical quadrature. Therefore, it is imperative to derive estimates for these errors in terms of the basic parameters of the algorithm. In a multilevel implementation of the fast multipole algorithm, one resorts to interpolating the far field expansions in order to achieve $\mathcal{O}(N)$ complexity resulting in a third kind of error.

The purpose of this paper is to present error estimates for the errors introduced by FMM and MLFMA in the solution of scalar wave equations. The algorithm is described in detail in [7]. An analysis of the errors in the FMM algorithm for the monopole and dipole terms is given in [8] and [9]. This work extends the error analysis to higher-order multipole terms, and includes an analysis for integration and interpolation errors.

2. Scalar addition theorem. The scalar wave functions are the characteristic solutions of the scalar wave equation

$$\nabla^2 u + k^2 u = 0, \tag{2.1}$$

which form a complete set. In spherical coordinates, such a set of characteristic solutions which satisfy the radiation condition (outgoing wave functions) are given by

$$\Psi_{lm}(k, \mathbf{r}) = h_l^{(1)}(kr) Y_{lm}(\theta, \phi), \tag{2.2}$$

where (r, θ, ϕ) are the spherical coordinates of the vector $\mathbf{r}$, $h_l^{(1)}(\cdot)$ denotes the spherical Hankel function of the first kind, and

$$Y_{lm}(\theta, \phi) = (-1)^m \left[\frac{(l-m)!}{(l+m)!} \frac{2l+1}{4\pi} \right]^{1/2} P_l^m(\cos\theta) \, e^{im\phi} \tag{2.3}$$

are the spherical harmonics. A time dependence of $e^{-i\omega t}$ is assumed and the associated Legendre functions $P_l^m(\cos\theta)$ are as defined in [10]. We will also use the shorthand notation $Y_{lm}(\hat{\mathbf{s}})$ to denote the spherical harmonics where $\hat{\mathbf{s}}$ is a unit vector whose direction is defined by the spherical angles θ and ϕ.

In many scattering problems it becomes necessary to express a scalar wave function in a different coordinate system. An outgoing wave function with the origin at a point O can be expanded into a sum of regular wave functions with the origin at a point O' since these fields do not have a singularity at O'. This expansion, known as the addition theorem for the scalar wave functions [11, 12], is given as

$$\Psi_{lm}(k, \mathbf{r}) = \sum_{l'=0}^{\infty} \sum_{m'=-l'}^{l'} \mathcal{R}g\Psi_{l'm'}(k, \mathbf{r}') \, \alpha_{l'm',lm}(k, \mathbf{r}''), \tag{2.4}$$

where $\mathbf{r}''$ is a vector from O to O' and $\mathcal{R}g\Psi_{lm}(k, \mathbf{r}) = j_l(kr) Y_{lm}(\theta, \phi)$ with $j_l(\cdot)$ denoting the spherical Bessel function of order l. This relation is valid provided that $|\mathbf{r}'| < |\mathbf{r}''|$, and the coefficients $\alpha_{l'm',lm}(k, \mathbf{r}'')$ are given as

$$\alpha_{l'm',lm}(k, \mathbf{r}'') = \sum_{n=|l-l'|}^{l+l'} i^{(l'-l+n)} \, Y_{n,m-m'}(\theta'', \phi'') \, h_n^{(1)}(kr'')$$

$$(-1)^{m'} \left[\frac{4\pi(2l+1)(2l'+1)}{(2n+1)} \right]^{1/2} (l\,l'\,0\,0|l\,l'\,n\,0), (l\,l'\,-m\,m'|l\,l'\,n\,-m+m'), \tag{2.5}$$

where $(j_1\,j_2\,m_1\,m_2|j_1\,j_2\,j_3\,m_3)$ are the Clebsch–Gordan coefficients as defined in [10].

The matrix representation of a translation operator has entries $\alpha_{l'm',lm}(k, \mathbf{r}'')$ and, in general, is a full matrix. It was shown by Rokhlin [13] that the spherical translation operators have diagonal representations when expressed in terms of their

plane wave spectra. The diagonal translation theory was later summarized by Epton and Dembart [14]. A different derivation based on plane wave expansions is given in [15]. The basic result is to express the coefficients $\alpha_{l'm',lm}(k, \mathbf{r}'')$ as

$$(2.6) \qquad \alpha_{l'm',lm}(k, \mathbf{r}'') = \lim_{N_\alpha \to \infty} \oint i^{l'-l} Y_{lm}(\hat{\mathbf{k}}) \, \tilde{\alpha}_{N_\alpha;D}(\hat{\mathbf{k}}) \, Y_{l'm'}^*(\hat{\mathbf{k}}) \, e^{i\mathbf{k}\cdot \mathbf{d}} \, d\hat{\mathbf{k}},$$

where $\mathbf{r}'' = \mathbf{D} + \mathbf{d}$ and

$$(2.7) \qquad \tilde{\alpha}_{N_\alpha;D}(\hat{\mathbf{k}}) = \sum_{n=0}^{N_\alpha} i^n \, (2n+1) \, h_n^{(1)}(kD) \, P_n\left(\hat{\mathbf{k}} \cdot \hat{\mathbf{D}}\right),$$

and $\hat{\mathbf{k}}$ and $\hat{\mathbf{D}}$ denote the unit vectors in the direction of respective vectors.

In order to obtain (2.6), the order of an integration and an infinite summation must be exchanged—which is not permissible in the strict sense. As a result, the infinite sum in the expression of $\tilde{\alpha}_{\infty;D}(\hat{\mathbf{k}})$ does not converge. However, we need to truncate the infinite sum at some N_α for numerical purposes. Thus the series can be truncated first, after which exchanging the order of integration and a finite sum does not pose a problem. The number of terms that must be kept in the summation, N_α, depends on the value of kd, as well as the desired accuracy. As d is increased, more terms are required to keep the accuracy at the desired level. On the other hand, if too many terms are used so that N_α is large, the spherical Hankel functions oscillate wildly with exponentially large values, causing numerical inaccuracies. This occurs when the order exceeds the argument, hence the condition

$$(2.8) \qquad N_\alpha \lesssim kD$$

must be satisfied.

3. Error analysis. There are two kinds of errors introduced in the single-level FMM calculations. First, the infinite sum over n in the expression for $\tilde{\alpha}_{\infty;D}(\hat{\mathbf{k}})$ is truncated. Second, the integral appearing in (2.6) is computed numerically, resulting in discretization error. In a multilevel algorithm, an error due to interpolation is also introduced which propagates through levels.

3.1. Truncation error. We first write (2.6) as

$$\alpha_{l'm',lm}(k, \mathbf{r}'') = \lim_{N_\alpha \to \infty} \oint i^{l'-l} Y_{lm}(\hat{\mathbf{k}}) \, \tilde{\alpha}_{N_\alpha;D}(\hat{\mathbf{k}}) \, Y_{l'm'}^*(\hat{\mathbf{k}}) \, e^{i\mathbf{k}\cdot \mathbf{d}} \, d\hat{\mathbf{k}}$$

$$(3.1) \qquad = \sum_{n=0}^{\infty} i^{n-l+l'}(2n+1)h_n^{(1)}(kD) \oint Y_{lm}(\hat{\mathbf{k}}) \, P_n(\hat{\mathbf{k}} \cdot \hat{\mathbf{D}}) \, Y_{l'm'}^*(\hat{\mathbf{k}}) \, e^{i\mathbf{k}\cdot \mathbf{d}} \, d\hat{\mathbf{k}},$$

where the order of integration and summation has been exchanged. The integrals in this equation will be denoted by

$$(3.2) \qquad I_{lm,l'm'}^{(n)}(\hat{\mathbf{D}}, \mathbf{d}) = \oint Y_{lm}(\hat{\mathbf{k}}) \, P_n(\hat{\mathbf{k}} \cdot \hat{\mathbf{D}}) \, Y_{l'm'}^*(\hat{\mathbf{k}}) \, e^{i\mathbf{k}\cdot \mathbf{d}} \, d\hat{\mathbf{k}},$$

and we define the truncation error as

$$\epsilon_T = \alpha_{l'm',lm}(k, \mathbf{r}'') - \sum_{n=0}^{N_\alpha} i^{n-l+l'} \, (2n+1) \, h_n^{(1)}(kD) \, I_{lm,l'm'}^{(n)}(\hat{\mathbf{D}}, \mathbf{d}))$$

$$(3.3) \qquad = \sum_{n=N_\alpha+1}^{\infty} i^{n-l+l'} \, (2n+1) \, h_n^{(1)}(kD) \, I_{lm,l'm'}^{(n)}(\hat{\mathbf{D}}, \mathbf{d}).$$

The exponential term appearing in the integral $I_{lm,l'm'}^{(n)}(\hat{\mathbf{D}}, \mathbf{d})$ can be expanded into a spherical harmonic series as

$$(3.4) \qquad e^{i\mathbf{k}\cdot\mathbf{d}} = \sum_{\nu=0}^{\infty} i^\nu \, (2\nu+1) \, j_\nu(kd) \, P_\nu(\hat{\mathbf{k}} \cdot \hat{\mathbf{d}}),$$

and using the addition theorem

$$(3.5) \qquad P_n(\hat{\mathbf{k}} \cdot \hat{\mathbf{D}}) = \frac{4\pi}{2n+1} \sum_{\mu=-n}^{n} Y_{n\mu}^*(\hat{\mathbf{D}}) \, Y_{n\mu}(\hat{\mathbf{k}}) \,,$$

the integrals can be rewritten as

$$I_{lm,l'm'}^{(n)}(\hat{\mathbf{D}}, \mathbf{d}) = \frac{(4\pi)^2}{2n+1} \sum_{\nu=0}^{\infty} i^\nu \, j_\nu(kd) \sum_{\mu=-n}^{n} Y_{n\mu}^*(\hat{\mathbf{D}}) \sum_{\mu'=-\nu}^{\nu} Y_{\nu\mu'}(\hat{\mathbf{d}})$$

$$(3.6) \qquad\qquad \oint Y_{lm}(\hat{\mathbf{k}}) \, Y_{\nu\mu'}^*(\hat{\mathbf{k}}) \, Y_{n\mu}(\hat{\mathbf{k}}) \, Y_{l'm'}^*(\hat{\mathbf{k}}) \, d\hat{\mathbf{k}}.$$

The integral of the product of four spherical harmonics can be evaluated by using the formulas in [11], yielding

$$I_{lm,l'm'}^{(n)}(\hat{\mathbf{D}}, \mathbf{d}) = \frac{4\pi\sqrt{(2l+1)(2l'+1)}}{\sqrt{2n+1}} \sum_\nu, i^\nu \sqrt{2\nu+1} j_\nu(kd) \sum_\mu Y_{n\mu}^*(\hat{\mathbf{D}}) Y_{\nu,\mu+m-m'}(\hat{\mathbf{d}})$$

$$(-1)^{m'+\mu} \sum_\eta (2\eta+1) \begin{pmatrix} l & l' & \eta \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} l & l' & \eta \\ m & -m' & -m+m' \end{pmatrix}$$

$$(3.7) \qquad\qquad \begin{pmatrix} n & \nu & \eta \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} n & \nu & \eta \\ \mu & -\mu-m+m' & m-m' \end{pmatrix}.$$

Due to the Wigner 3-j symbols, the summations in this formula are all finite, and their ranges are

$$(3.8) \qquad \begin{array}{ccccc} \max\{0, |l-l'|-n, n-l-l'\} & \leq & \nu & \leq & n+l+l', \\ \max\{-n, -m+m'-\nu\} & \leq & \mu & \leq & \min\{-m+m'+\nu, n\}, \\ \max\{|l-l'|, |n-\nu|\} & \leq & \eta & \leq & \min\{l+l', n+\nu\}. \end{array}$$

3.1.1. Monopole-to-monopole translations. The special case of monopole-to-monopole translation, i.e., $l = l' = 0$, is of particular importance since, in many scalar scattering problems and the potential formulation of the vector scattering problem, one requires only monopole-to-monopole translations given by

$$(3.9) \qquad \alpha_{00,00}(k, \mathbf{r}'') = h_0^{(1)}(kr'') = -i\frac{e^{ikr''}}{kr''},$$

which is the Green's function of the problem. It is easy to show that

$$(3.10) \qquad I_{00,00}^{(n)}(\hat{\mathbf{D}}, \mathbf{d}) = \frac{1}{4\pi} \oint P_n(\hat{\mathbf{k}} \cdot \hat{\mathbf{D}}) \, e^{i\mathbf{k}\cdot\mathbf{d}} \, d\hat{\mathbf{k}} = i^n P_n(\hat{\mathbf{D}} \cdot \hat{\mathbf{d}}) j_n(kd),$$

and the truncation error becomes

$$(3.11) \qquad \epsilon_T = \sum_{n=N_\alpha+1}^{\infty} (-1)^n \, (2n+1) \, h_n^{(1)}(kD) \, P_n(\hat{\mathbf{D}} \cdot \hat{\mathbf{d}}) j_n(kd).$$

To obtain an error bound, we observe that the error is maximum when $\mathbf{D}$ and $\mathbf{d}$ are collinear, and assuming that $kD \to \infty$ we replace the spherical Hankel functions by their large-argument asymptotic forms to get

$$(3.12) \quad |\epsilon_T| \leq \sum_{n=N_\alpha+1}^{\infty} (2n+1) \left| h_n^{(1)}(kD) P_n(\hat{\mathbf{D}} \cdot \hat{\mathbf{d}}) j_n(kd) \right| \leq \frac{1}{kD} \sum_{n=N_\alpha+1}^{\infty} (2n+1) \left| j_n(kd) \right|,$$

which can also be interpreted as the truncation error in the radiation pattern of a source [6, 7]. If kd is sufficiently small (i.e., $kd \leq N_\alpha$), $j_n(kd)$ for $n > N_\alpha + 1$ are all positive, and we have

$$(3.13) \quad |\epsilon_T| \leq \frac{1}{kD} (2N_\alpha + 3) j_{N_\alpha+1}(kd) \leq \frac{1}{kD} \frac{(kd)^{(N_\alpha+1)}}{(2N_\alpha + 1)!!},$$

where $(2n+1)!!$ denotes the product of all odd integers from 1 up to $2n+1$. The last inequality, obtained by using the upper bound for Bessel functions given by [10], is

$$(3.14) \quad |J_\nu(z)| \leq \frac{|z/2|}{\Gamma(\nu + 1)} \qquad \left(\nu \geq -\frac{1}{2} \right),$$

which can be restated for the spherical Bessel functions as

$$(3.15) \quad |j_n(z)| \leq \frac{z^n}{(2n+1)!!} \sim \sqrt{\frac{e}{2}} \frac{(ez)^n}{(2n+1)^{n+1}}.$$

Equation (3.13) shows that the decrease in the truncation error, as N_α is increased, is faster than exponential. However, there are two important points to be considered for the choice of N_α. First, its value also determines the number of quadrature points to be used for numerical integration (as will be discussed in section 3.2). Thus its value must be as small as possible to keep the computational load at a minimum. Second, the expression for $\tilde{\alpha}_{N_\alpha;D}(\hat{\mathbf{k}})$ given in (2.6) is a divergent sum; hence N_α should be small enough to avoid numerical problems.

To obtain a more detailed analysis for finite values of kD, we can divide the sum in (3.11) into two parts in which the index of summation runs from $N_\alpha + 1$ to N_1 for the first sum, and it runs from $N_1 + 1$ up to infinity for the second sum where $N_1 \sim kD$. For the first sum, we have $kd < n < kD$, and since the magnitude of the spherical Hankel function is a slowly varying function of the order in this interval, we can replace it by $h_{N_\alpha}^{(1)}(kD)$ and, using (3.15) for the spherical Bessel functions, we get

$$(3.16) \quad |(2n+1) h_n^{(1)}(kD) j_n(kd)| \sim h_{N_\alpha}^{(1)}(kD) \frac{(kd)^n}{(2n+1)!!},$$

which is a rapidly decreasing function of the order, n. Therefore, this sum can be represented by its first term, i.e.,

$$(3.17) \quad S_1 \triangleq \sum_{n=N_\alpha+1}^{N_1} (2n+1) |h_n^{(1)}(kD) j_n(kd)| \sim h_{N_\alpha+1}^{(1)}(kD) \frac{(kd)^{N_\alpha+1}}{(2N_\alpha + 1)!!}.$$

For the second sum, we have $kd < kD < n$ and we can use the large-order asymptotic forms of both the spherical Hankel and Bessel functions to obtain

$$(3.18) \quad |(2n+1) h_n^{(1)}(kD) j_n(kd)| \sim \frac{1}{kD} \left(\frac{d}{D} \right)^n.$$

Thus, the second sum is approximated by a geometric series and we get

$$(3.19) \qquad S_2 \triangleq \sum_{n=N_\alpha+1}^{N_1} (2n+1) \, |h_n^{(1)}(kD) \, j_n(kd)| \sim \frac{1}{kD} \frac{(d/D)^{N_1+1}}{1-d/D}.$$

The total truncation error is bounded by the sum of S_1 and S_2 and, for a given value of N_α, it decreases as the ratio d/D decreases. However, it is important to note that the relative error, which is defined as the ratio of the truncation error to the actual value of translation coefficient, has a nonzero limiting value as $kD \to \infty$. On the other hand, for a given d/D ratio, the error can be decreased by increasing N_α.

Figure 3.1 shows a typical behavior of the relative error along with the error estimates given by (3.13) and $S_1 + S_2$. It can be seen that these formulas are quite satisfactory. This result also shows that, for small d/D, the error is only weakly dependent on D, and the converse is true for larger d/D.

3.1.2. Multipole-to-multipole translations. It is quite difficult to put a tight bound on the summations in (3.7) since the ranges of the summations depend on each other. On the other hand, the bound obtained for monopole-to-monopole translations is given by the asymptotic behavior of the error for small kd. Motivated by this result, we will investigate the asymptotic behavior of the error for the multipole-to-multipole translation.

For this purpose, we first observe that if $n > l + l'$, the dominant term in (3.7) is for $\nu = n - l - l'$, and for this value of ν the only possible value of η is $l + l'$, reducing the summation over η to a single term. We get

$$(3.20)$$
$$|I_{lm,l'm'}^{(n)}(\hat{\mathbf{D}}, \mathbf{d})|$$
$$\approx \frac{4\pi\sqrt{(2l+1)(2l'+1)}}{\sqrt{2n+1}} \, \sqrt{2\nu+1} \, j_\nu(kd) \, W(l,m,l',m') \, S(n,l+l',m-m',\hat{\mathbf{D}},\hat{\mathbf{d}}),$$

where

$$(3.21)$$
$$W(l,m,l',m') = (2l+2l'+1) \left| \begin{pmatrix} l & l' & l+l' \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} l & l' & l+l' \\ m & -m' & -m+m' \end{pmatrix} \right|,$$

$$(3.22)$$
$$S(n,L,M,\hat{\mathbf{D}},\hat{\mathbf{d}}) = \sum_\mu \left| Y_{n\mu}(\hat{\mathbf{D}}) Y_{\nu,\mu+M}(\hat{\mathbf{d}}) \begin{pmatrix} n & \nu & L \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} n & \nu & L \\ \mu & -\mu-M & M \end{pmatrix} \right|$$

with $L = l + l'$, $\nu = n - L$, $M = m - m'$.

By expanding the Wigner 3-j terms, it can be shown that

$$(3.23) \qquad W(l,m,l',m') \leq \frac{C_L^l}{\sqrt{C_{2L}^{2l}}} \leq 1,$$

where $C_n^m = n!/(m!(n-m)!)$ denotes the binomial coefficients.

The function $S(n,L,M,\hat{\mathbf{D}},\hat{\mathbf{d}})$ attains its maximum value when $M = 0$ and $\hat{\mathbf{D}} = \hat{\mathbf{d}} = \hat{\mathbf{a}}_z$, which is

$$(3.24) \qquad S_{\max}(n,L) = \frac{1}{4\pi} \sqrt{\frac{2\nu+1}{2n+1}} \frac{(C_n^L)^2}{C_{2n}^{2L}},$$
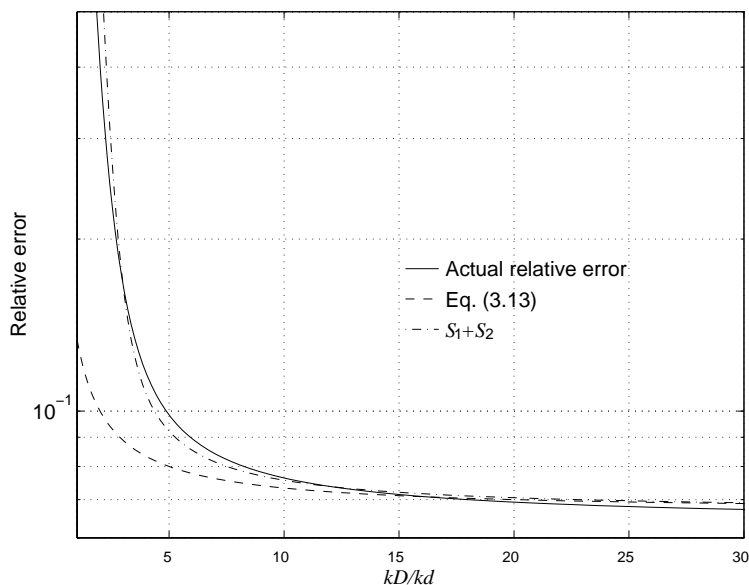
FIG. 3.1. *Comparison of the error estimates and the actual error for $kd = 1$, $N_\alpha = \lfloor kd + \log(kd + \pi)\rfloor = 2$ as a function of kD.*

from which follows

$$\frac{4\pi\sqrt{(2l+1)(2l'+1)}}{\sqrt{2n+1}}\sqrt{2\nu+1}\,S \le \sqrt{(2l+1)(2l'+1)}\,\frac{2\nu+1}{2n+1}\,\frac{\left(C_n^L\right)^2}{C_{2n}^{2L}}$$

$$(3.25) \qquad\qquad\qquad\qquad\qquad \le \frac{(2L+1)(2\nu+1)}{2n+1}\,\frac{\left(C_n^L\right)^2}{C_{2n}^{2L}} \le 1.$$

Combining all these results gives

$$(3.26) \qquad |I_{lm,l'm'}^{(n)}(\hat{\mathbf{D}},\mathbf{d})| \approx j_\nu(kd) \approx \frac{(kd)^\nu}{(2\nu+1)!!}, \qquad \nu = n - l - l'.$$

The asymptotic behavior of the truncation error is also determined by the first term of the sum in (3.3), and for $kD \to \infty$, we can write

$$(3.27) \qquad |\epsilon_T| \approx \frac{2N_\alpha + 3}{kD}\,\frac{(kd)^P}{(2P+1)!!},$$

where $P = N_\alpha + 1 - l - l'$. It must be noted that, although this expression is obtained by assuming $kd \to 0$, numerical experimentation shows that this is indeed a bound for the truncation error.

Finally, by using Stirling's formula to approximate the factorial, we get

$$(3.28) \qquad |\epsilon_T| \le \frac{2N_\alpha + 3}{kD}\,\frac{1}{\sqrt{2}}\,\left(\frac{ekd}{2P}\right)^P,$$

which shows that the decrease in the truncation error with N_α is faster than exponential.

An analysis for finite values of kD follows the lines described at the end of the previous section and again shows that the rate of decrease of the truncation error with N_α is faster than exponential.

3.2. Numerical integration error. In implementation, the integral representation for the translation coefficient appearing in (2.6) must be evaluated numerically. For this purpose, we note that the product formula

$$(3.29) \qquad \oint f(x,y,z)d\hat{\mathbf{s}} = \int_0^{2\pi} \int_0^{\pi} f(\theta,\phi)\sin\theta\, d\theta\, d\phi = \sum_{i=1}^{N_\theta} \sum_{j=1}^{2N_\theta} w_i^g w_j^c f(\theta_i,\phi_j)$$

is exact for polynomials $x^\alpha y^\beta z^\gamma$ if $\alpha + \beta + \gamma < 2N_\theta$ (see [16]), where θ_i are chosen as the Gauss–Legendre points, ϕ_j are $2N_\theta$ equally spaced points over the interval $[-\pi,\pi]$, w_i^g are the Gauss–Legendre weights, and $w_j^c = \pi/N_\theta$. Since the spherical harmonics $Y_{nm}(\theta,\phi)$ are also polynomials of order n on the surface of the sphere, the above quadrature rule is exact for harmonics of order $n < 2N_\theta$. A more efficient method for numerical quadrature on the sphere is given in [17].

The integrals $I_{lm,l'm'}^{(n)}(\hat{\mathbf{D}},\mathbf{d})$ can be expanded as in (3.6). If a numerical integration with $N_\theta \times 2N_\theta$ points is used, the product of four spherical harmonics appearing in (3.6) can be integrated exactly provided that $2N_\theta - 1 \geq l + \nu + n + l'$. Since the upper limit of the summation over ν is $n + l + l'$, choosing $N_\theta = n + l + l' + 1$ will result in exact integration for all the nonzero terms in (3.6). However, although the higher-order terms are analytically zero, their numerical evaluation will not be zero. Thus the error due to numerical integration of $I_{lm,l'm'}^{(n)}(\hat{\mathbf{D}},\mathbf{d})$ will be

$$\epsilon_i(n) = \frac{(4\pi)^2}{2n+1} \sum_{\nu=2N_\theta-n-l-l'}^{\infty} i^\nu \, |j_\nu(kd)| \sum_{\mu=-n}^{n} Y_{n\mu}^*(\hat{\mathbf{D}}) \sum_{\mu'=-\nu}^{\nu} Y_{\nu\mu'}(\hat{\mathbf{d}})$$

$$(3.30) \qquad \overline{\oint} Y_{lm}(\hat{\mathbf{k}})\, Y_{\nu\mu'}^*(\hat{\mathbf{k}})\, Y_{n\mu}(\hat{\mathbf{k}})\, Y_{l'm'}^*(\hat{\mathbf{k}})\, d\hat{\mathbf{k}},$$

where $\overline{\oint}$ is used to denote numerical integration which is bounded by

$$(3.31) \qquad \left| \overline{\oint} Y_{lm}(\hat{\mathbf{k}})\, Y_{\nu\mu'}^*(\hat{\mathbf{k}})\, Y_{n\mu}(\hat{\mathbf{k}})\, Y_{l'm'}^*(\hat{\mathbf{k}})\, d\hat{\mathbf{k}} \right| \leq \frac{\sqrt{(2l+1)(2\nu+1)(2n+1)(2l'+1)}}{4\pi}.$$

Replacing this upper bound in (3.30) yields

$$(3.32) \qquad |\epsilon_i(n)| \leq \sqrt{(2l+1)(2l'+1)} \sum_{\nu=2N_\theta-n-l-l'}^{\infty} (2\nu+1)\, |j_\nu(kd)|.$$

The total error due to numerical integration is

$$(3.33) \qquad |\epsilon_I| = \left| \sum_{n=0}^{N_\alpha} (2n+1) h_n^{(1)}(kD) \epsilon_i(n) \right|.$$

Replacing the bound for ϵ_i in this expression and using large-argument asymptotic forms of the Hankel functions yields

$$(3.34) \qquad |\epsilon_I| \leq \frac{\sqrt{(2l+1)(2l'+1)}}{kD} \sum_{n=0}^{N_\alpha} (2n+1) \sum_{\nu=2N_\theta-n-l-l'}^{\infty} (2\nu+1)|j_\nu(kd)|.$$

Assuming $N_\theta > N_\alpha > l + l'$, it can be shown that

(3.35)

$$\sum_{n=0}^{N_\alpha}(2n+1)\sum_{\nu=2N_\theta-n-L}^{\infty}(2\nu+1)|j_\nu(kd)|$$

$$=\sum_{n=2N_\theta-N_\alpha-l-l'}^{\infty}(2n+1)|j_n(kd)|\sum_{\substack{i=2N_\theta-L-n\\i\geq0}}^{N_\alpha}(2i+1)\leq 2.0\,(2N_\alpha-1)\,\frac{(kd)^P}{(2P-1)!!},$$

in which the leading factor 2.0 is obtained empirically. Finally, we have

(3.36) $$|\epsilon_I|\leq(2N_\alpha-1)\,\frac{\sqrt{(2l+1)(2l'+1)}}{kD}\frac{2(kd)^P}{(2P-1)!!}.$$

It can be seen that if N_θ is chosen to be greater than N_α, the numerical integration error will have a behavior similar to the truncation error. When $N_\theta = N_\alpha + 1$ the error bounds for truncation and numerical integration appear to be of the same order. However, the bound used in (3.31) is too high since the value of the integrand is replaced by its maximum value for all sample points. This suggests that the integration error will be much smaller than the truncation error. Therefore, the total numerical error can be said to be bounded by the truncation error only.

Several important points are to be noted from the foregoing derivations:
- The choice of $N_\theta = N_\alpha + 1$ is appropriate since all the contributing integrals are evaluated exactly.
- The error decreases very rapidly as the number of terms in the truncated sum, N_α, increases. This is shown in Figure 3.2.
- The error is an increasing function of kd and a decreasing function of kD, as can be seen from Figures 3.3 and 3.4. This makes it necessary to use more terms for larger values of kd. On the other hand, due to the Hankel functions appearing in the expressions, the convergence is achieved if $N_\alpha < kD$. It must be noted that the relative error has a nonzero limiting value as $kD \to \infty$ for the wave equation, which is quite different from the static case.
- The error increases as l or l' increases, as can be seen from Figure 3.5.

An empirical formula for choosing N_α is

(3.37) $$N_\alpha = \lfloor kd + C_a \log(kd + \pi)\rfloor + l + l',$$

where $\lfloor\cdot\rfloor$ denotes the integer part, and C_a is a parameter chosen to control the accuracy of the calculations. If C_a is chosen to be 1, then the relative error in the evaluation of (2.6) is about 10%, and with each increase in C_a by 1 the error decreases roughly by an order of magnitude.

Another formula for choosing N_α is

(3.38) $$N_\alpha = \lfloor kd + C_a'\,(kd)^{1/3}\rfloor + l + l',$$

which can be obtained by using asymptotic expansions of the Bessel functions [18].

3.3. Interpolation error. In a multilevel implementation of the FMM (MLFMA), the multipole expansions of groups at a certain level are aggregated (disaggregated) to obtain the multipole expansions for the groups at a higher (lower) level. Thus, it
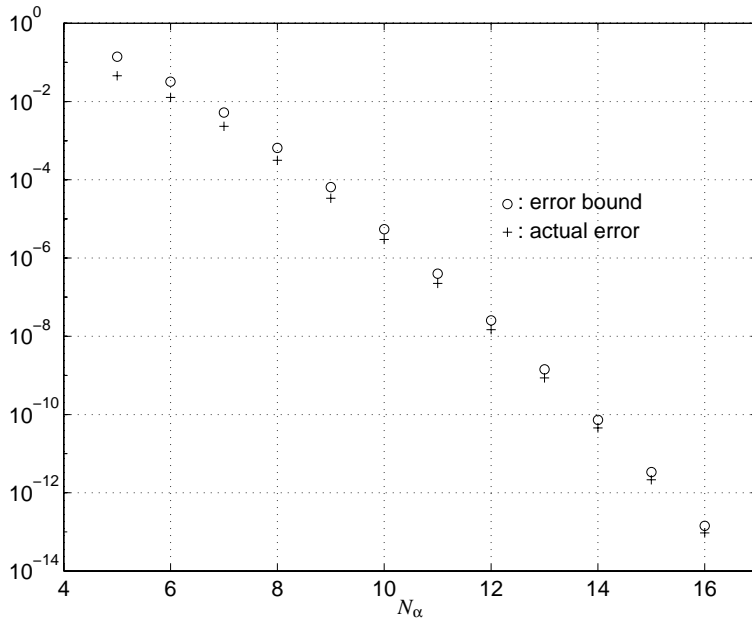
FIG. 3.2. *Actual errors (+) and the error bound given by* (3.27) *(○) as a function of N_α for* $k\mathbf{D} = 30\hat{\mathbf{a}}_z$, $k\mathbf{d} = \hat{\mathbf{a}}_z$, $l = 3$, $l' = 2$, $m = 0$, $m' = 0$, *and* $N_\theta = N_\alpha + 1$.
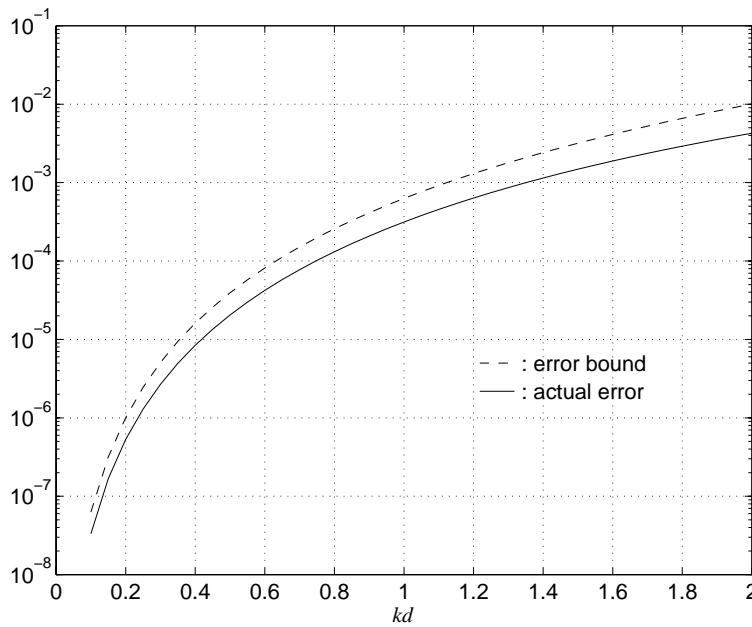

FIG. 3.3. *Actual errors (solid) and the error bound given by* (3.27) *(dashed) as a function of* kd *for* $k\mathbf{D} = 30\hat{\mathbf{a}}_z$, $k\mathbf{d} = kd\hat{\mathbf{a}}_z$, $l = 3$, $l' = 2$, $m = 0$, $m' = 0$, $N_\alpha = 8$, *and* $N_\theta = N_\alpha + 1$.
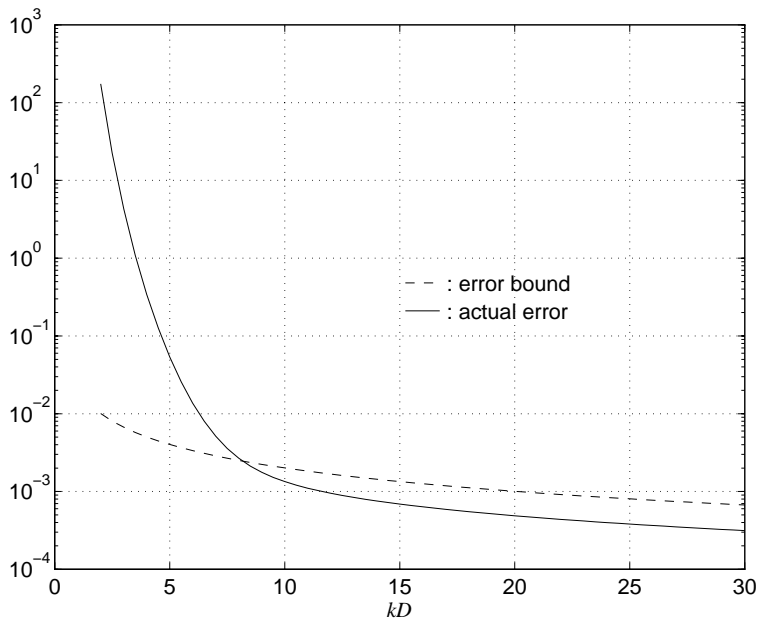
FIG. 3.4. *Actual errors (solid) and the error bound given by* (3.27) *(dashed) as a function of* kD *for* $k\mathbf{D} = kD\hat{\mathbf{a}}_z$, $k\mathbf{d} = \hat{\mathbf{a}}_z$, $l = 3$, $l' = 2$, $m = 0$, $m' = 0$, $N_\alpha = 8$, *and* $N_\theta = N_\alpha + 1$.
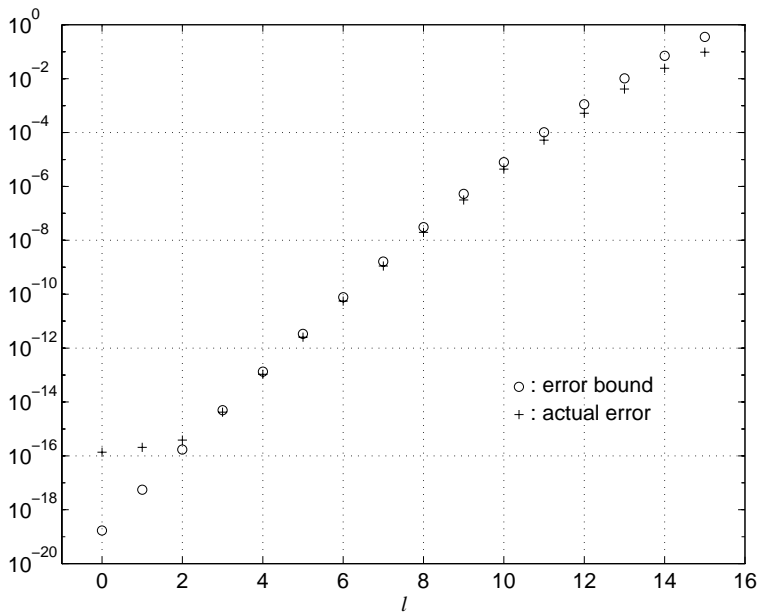

FIG. 3.5. *Actual errors* (+) *and the error bound given by* (3.27) (◦) *as a function of* l *for* $k\mathbf{D} = 30\hat{\mathbf{a}}_z$, $k\mathbf{d} = \hat{\mathbf{a}}_z$, $l' = 0$, $m = 0$, $m' = 0$, $N_\alpha = 15$, *and* $N_\theta = N_\alpha + 1$. *For* $l < 3$, *the actual error is below the machine precision.*

becomes necessary to evaluate the values of multipole expansions of a group at points other than they have already been calculated. If the multipole expansions are evaluated at the Gauss–Legendre points to facilitate the numerical integration, the grid structures of different levels will be quite different and nonuniform. Since the multipole expansion aggregated at a group is a spherical harmonic function of finite order, and if a sufficient number of points are used, it is possible to obtain the *exact* value of the multipole expansion at any point. However, such an approach does not reduce the computational complexity compared with a single-level algorithm. Therefore, it is necessary to use an appropriate interpolation algorithm. Since the grid points are nonuniform, the Lagrangian interpolation is a candidate for the purpose. However, other interpolators are possible as will be mentioned later in this section. Multilevel implementations of the three-dimensional MLFMA are described in [5, 6, 7].

The error in a polynomial interpolation is given by the Cauchy remainder formula [19, Chap. 3].

THEOREM 3.1. *Let $f(x) \in C[a, b]$ and suppose that $f^{(n)}(x)$ exists at each point of (a, b).*

If $a \leq x_1 < x_2 < \cdots < x_n \leq b$, then

$$R_n(f; x) \triangleq f(x) - p_n(f; x) = \frac{(x - x_1)(x - x_2) \cdots (x - x_n)}{(n)!} f^{(n)}(\xi),$$

where $p_n(f; x)$ is a polynomial approximation of order $(n - 1)$ to the function f. The point ξ depends upon $x, x_1, x_2, \ldots, x_n$ and f and is restricted as $\min(x, x_1) < \xi < \max(x, x_n)$.

In general the value of ξ is not known, and the following estimate becomes more useful.

COROLLARY 3.2. *If $f(x) \in C^n[a, b]$, then*

$$|R_n(f; x)| \leq \left\{ \max_{a \leq x \leq b} \left| f^{(n)}(x) \right| \right\} \frac{|x - x_1||x - x_2| \cdots |x - x_n|}{(n)!}.$$

To put this error bound into a more useful form, we need an estimate for the bound of the derivative of the function, and the following theorem for trigonometric polynomials (see Bernstein [20, Chap. 11]) proves useful.

THEOREM 3.3. *If $f(x) = \sum_{k=-N}^{N} c_k e^{ikx}$ is a trigonometric polynomial of degree N, and M is the maximum of $|f(x)|$, then*

$$|f'(x)| \leq NM,$$

which immediately gives the following corollary.

COROLLARY 3.4. *For a trigonometric polynomial of the form $f(x) = \sum_{k=-N}^{N} c_k e^{ikx}$ we have*

$$|f^{(n)}(x)| \leq N^n M.$$

If we choose $n = 2p$, (i.e., even number of points) and restrict the interpolation point to the center, $x_p \leq x \leq x_{p+1}$, it can be easily verified that

$$(3.39) \qquad |x - x_1||x - x_2| \cdots |x - x_{2p}| \leq \left[\frac{(2p)!}{p!} \right]^2 \left[\frac{h}{4} \right]^{2p}$$

where $h \overset{\triangle}{=} \max_i\{x_i - x_{i-1}\}, i = 2, \ldots, 2p$, and the error estimate becomes

$$(3.40) \qquad |R_n(f;x)| \leq M\frac{(2p)!}{(p!)^2} \left(\frac{Nh}{4}\right)^{2p}.$$

For interpolation over the surface of a unit sphere, the polynomial approximation is applied in θ and ϕ variables separately, with $N_\theta = N_\alpha + 1$ Gauss–Legendre points in θ variable and $N_\phi = 2N_\theta$ equally spaced points in ϕ variable.

The functions to be interpolated are the far field expansions of groups of radii $d/2$, which means that the spherical harmonic expansion coefficients decay rapidly for orders greater than $N_\alpha/2$. Thus, these functions are oversampled by a factor of 2. This is because the sampling rate is dictated by the harmonic content of the translation coefficients. From (2.6), in which the vector $\mathbf{d}$ is the sum of the translations within two groups, we can see that the spherical harmonic coefficients of the translation coefficient are significant up to order N_α.

With these observations, the estimate for the relative error can be written as

$$(3.41) \qquad \delta = \frac{|R_n(f;\theta,\phi)|}{M} \leq \frac{(2p)!}{(p!)^2} \left(\frac{N_\alpha h_\theta}{8}\right)^{2p} + \frac{(2q)!}{(q!)^2} \left(\frac{N_\alpha h_\phi}{8}\right)^{2q}.$$

It can be shown that

$$h_\theta, h_\phi < \frac{\pi}{N_\alpha},$$

which yields

$$(3.42) \qquad \delta \leq 2\,\frac{(2p)!}{(p!)^2} \left[\frac{\pi}{8}\right]^{2p},$$

where we have taken $q = p$. Finally, using the Stirling's formula to approximate the factorials gives

$$(3.43) \qquad \delta \approx \sqrt{\frac{2}{\pi n}} \left(\frac{\pi}{4}\right)^n.$$

This result reveals that the interpolation error decreases exponentially as the number of interpolation points is increased.

There are other interpolation schemes for the interpolation of fields over spherical surfaces [21, 22]. These algorithms use the approximate prolate spheroid series [23] and the sampling window (Chebyshev sampling series) [24] with an oversampling factor χ defined as the ratio between the actual sampling rate and the Nyquist rate. Both algorithms have small error bounds (smaller than that given in (3.43)) and exhibit exponential decrease in error with the number of retained samples and the oversampling factor. However, these algorithms use uniformly spaced samples and are not suitable for integration. Furthermore, with an oversampling factor of 2, the error for polynomial interpolation becomes almost as small as the error for these algorithms.

The most important feature of the foregoing analysis is that the interpolation error does not depend on the harmonic content (bandwidth) of the function being approximated, which can be intuitively understood by noting the fact that as the harmonic content of the function is increased, the sampling rate is also increased. Hence the use of the same number of interpolation points at each level of the MLFMA is justified, even though the functions being interpolated get richer in harmonic content as we go up through the levels.
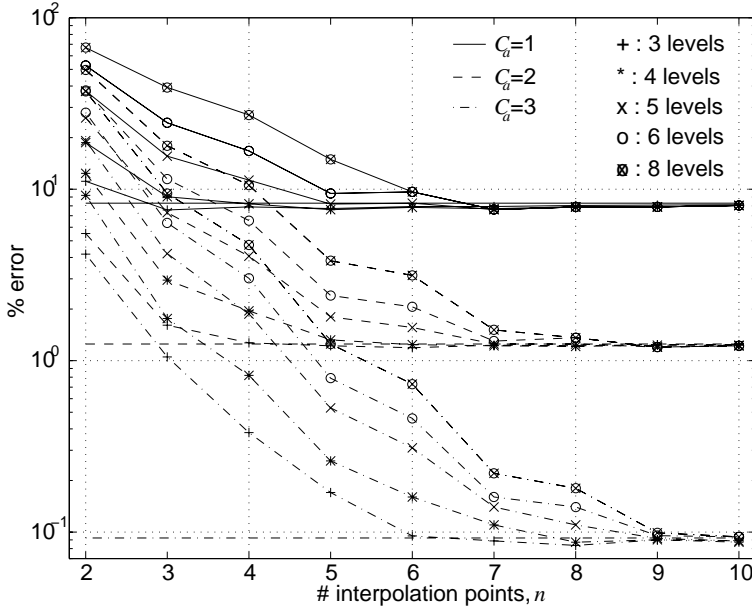
FIG. 3.6. *Relative error as a function of the number of points used for interpolation for different values of C_a and number of levels.*

3.4. Total error. It is quite difficult to obtain an error bound for the total error in the MLFMA algorithm since there are different sources of errors which propagate through levels. To investigate the total error introduced by the MLFMA, the translation matrix $\overline{\boldsymbol{\alpha}}_{LL'}(k, \mathbf{r})$ is calculated. The calculation of this matrix requires only a single matrix-vector multiplication. The relative error is defined as

$$(3.44) \qquad \epsilon_{\mathrm{rel}} = \frac{||\overline{\boldsymbol{\alpha}}_{LL'}(k, \mathbf{r}) - \hat{\overline{\boldsymbol{\alpha}}}_{LL'}(k, \mathbf{r})||}{||\overline{\boldsymbol{\alpha}}_{LL'}(k, \mathbf{r})||} \times 100,$$

where $||\cdot||$ denotes a matrix norm and $\hat{\overline{\boldsymbol{\alpha}}}_{LL'}(k, \mathbf{r})$ denotes the value calculated by the MLFMA. In the calculations, $k\mathbf{r} = 30(\hat{\mathbf{a}}_x + \hat{\mathbf{a}}_y + \hat{\mathbf{a}}_z)$ is used and the α matrices are truncated at $n_t = 3$, resulting in a 16×16 matrix.

Figure 3.6 shows the relative error as a function of the number of points, n, used in the interpolation for different number of levels and accuracy parameter C_a. It can be seen that for any given C_a and number of levels, the error decreases almost exponentially up to a certain value of n after which no improvement can be observed. The limiting value of the error as n increases is determined solely by the accuracy parameter C_a.

Figure 3.7 shows the relative error for a five-level MLFMA algorithm, as a function of C_a for different values of n (the number of points used in the interpolation). As C_a is increased, more terms are being included in the multipole expansions and the number of sampling points is increased accordingly. Also, the number of neighbors of a scatterer increases, thus more interactions are calculated directly. This increases the amount of work appreciably. It can be seen from the figure that the error for $C_a = 1, n = 5$ is smaller than the error for $C_a = 3, n = 2$ or 3. In the latter case the
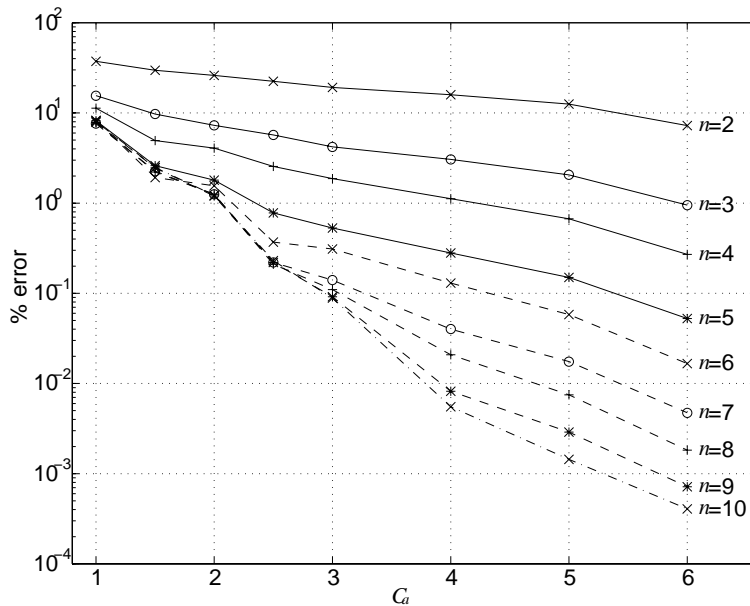
FIG. 3.7. *Relative error as a function of C_a for different values of the number of points used for interpolation for a five-level MLFMA algorithm.*

CPU time is much larger. Thus, C_a must be chosen as small as possible, but it must be large enough to give the desired accuracy. Numerical experimentation shows that $C_a = 1$ and $n = 4$ is an appropriate choice for a 10% relative error.

4. Conclusions. The truncation error and numerical integration error introduced by the dynamic FMM algorithm are analyzed and estimates are obtained. These estimates can be used in the implementation for choosing the values of several parameters with an a priori estimate of the overall error.

The two parameters that control the error in the MLFMA are C_a and p. The former controls the value of N_α that will be used at each level and the latter is the number of interpolation points used which is independent of the levels. This also means that the value of p is not affected by the problem size and is determined solely by the required accuracy, implying that the complexity of the MLFMA is indeed $\mathcal{O}(N)$ for volume problems and $\mathcal{O}(N \log N)$ for surface problems.

The main result of this study is to show that the decrease in the error, when either C_a or p is increased, is faster than exponential. Thus an increase in these parameters will decrease the error exponentially while the numerical work will be increased only linearly.

REFERENCES

[1] L. GREENGARD AND V. ROKHLIN, *A fast algorithm for particle simulations*, J. Comput. Phys., 73 (1987), pp. 325–348.
[2] V. ROKHLIN, *Rapid solution of integral equations of scattering theory in two dimensions*, J. Comput. Phys., 2 (1986), pp. 414–439.
[3] R. COIFMAN, V. ROKHLIN, AND S. WANDZURA, *The fast multipole method for the wave equation: a pedestrian prescription*, IEEE Trans. Antennas and Propagation, 35 (1993), pp. 7–12.

[4] C. C. LU AND W. C. CHEW, *A fast algorithm for solving the hybrid integral equation*, IEE Proc. Pt. H, 140 (1993), pp. 455–460.

[5] C. C. LU AND W. C. CHEW, *A multilevel algorithm for solving a boundary integral equation of wave scattering*, Microwave Opt. Tech. Lett., 7 (1994) pp. 466–470.

[6] J. M. SONG AND W. C. CHEW, *Multilevel fast-multipole algorithm for solving combined field integral equations of electromagnetic scattering*, Microwave. Opt. Tech. Lett., 10 (1995), pp. 14–19.

[7] S. KOC AND W. C. CHEW, *Calculation of acoustical scattering from a cluster of scatterers*, J. Acoust. Soc. Am., 103 (1998), pp. 721–734.

[8] J. RAHOLA, *Diagonal forms of the translation operators in the fast multipole algorithm for scattering problems*, Research Report A349, Helsinki University of Technology, Institute of Mathematics, Helsinki, Finland, 1995.

[9] J. M. SONG, C. C. LU, AND W. C. CHEW, *MLFMA for electromagnetic scattering by large complex objects*, IEEE Trans. Antennas and Propagation, 45 (1997), pp. 1488–1493.

[10] M. ABRAMOWITZ AND I. E. STEGUN, eds. *Handbook of Mathematical Functions*, Dover, New York, 1972.

[11] A. R. EDMONDS, *Angular Momentum in Quantum Mechanics*, Princeton University Press, Princeton, NJ, 1957.

[12] S. STEIN, *Addition theorems for spherical wave functions*, Quart. Appl. Math., 19 (1961), pp. 15–24.

[13] V. ROKHLIN, *Diagonal forms of translation operators for the Helmholtz equation in three dimensions*, Appl. Comput. Harmon. Anal., 1 (1993), pp. 82–93.

[14] M. A. EPTON AND B. DEMBART, *Multipole translation theory for the three-dimensional Laplace and Helmholtz equations*, SIAM J. Sci. Comput., 16 (1995), pp. 865-897.

[15] W. C. CHEW, S. KOC, J. M. SONG, C. C. LU, AND E. MICHIELSSEN, *A succinct way to diagonalize the translation matrix in three dimensions*, Microwave Opt. Tech. Lett., 15 (1997), pp. 144–147.

[16] A. H. STROUD AND D. SECREST, *Gaussian Quadrature Formulas*, Prentice–Hall, Englewood Cliffs, NJ, 1966.

[17] A. D. MCLAREN, *Optimal numerical integration on a sphere*, Math. Comput., 17 (1963), pp. 361–383.

[18] O. M. BUCCI AND G. FRANCESCHETTI, *On the spatial bandwidth of scattered fields*, IEEE Trans. Antennas and Propagation, AP-35 (1987), pp. 1445–1455.

[19] P. J. DAVIS, *Interpolation & Approximation*, Dover, New York, 1975.

[20] R. P. BOAS, *Entire Functions*, Academic Press, New York, 1954.

[21] O. M. BUCCI, C. GENNARELLI, AND C. SAVARESE, *Optimal interpolation of radiated fields over a sphere*, IEEE Trans. Antennas and Propagation, AP-39 (1991), pp. 1633–1643.

[22] O. M. BUCCI, C. GENNARELLI, R. RICCIO, AND C. SAVARESE, *Electromagnetic fields interpolation from nonuniform samples over spherical and cylindrical surfaces*, IEE Proc. Microwave Antennas Propagation, 141 (1994), pp. 77–84.

[23] J. J. KNAB, *Interpolation of band-limited functions using the approximate prolate series*, IEEE Trans. Inform. Theory, IT-25 (1979), pp. 717–720.

[24] J. J. KNAB, *The sampling window*, IEEE Trans. Inform. Theory, IT-29 (1983), pp. 157–159.