The Fast Multipole Method for the Wave Equation: A Pedestrian Prescription

Ronald Coifman\(^1\), Vladimir Rokhlin\(^1\), and Stephen Wandzura\(^2\)

\(^1\)Fast Mathematical Algorithms and Hardware Corporation
1020 Sherman Avenue
Hamden, CT 06514

\(^2\)Hughes Research Labs
3011 Malibu Canyon Road
Malibu, CA 90265

1. Introduction

The purpose of this article is to give a practical and complete, but not rigorous, exposition of the Fast Multipole Method (FMM). The aim is to give the computational physicist or engineer a sufficiently clear understanding of the method that he or she will be able to implement it with a minimum of difficulty. For mathematical background and rigor, we refer the reader to Rokhlin’s papers [1, 2].

The FMM provides an efficient mechanism for the numerical convolution of the Green’s function for the Helmholtz equation with a source distribution. It can be used to radically accelerate the iterative solution of boundary-integral equations. In the simple single-stage form presented here, it reduces the computational complexity of the convolution from \(O(N^2)\) to \(O(N^{3/2})\), where \(N\) is the dimensionality of the problem’s discretization. By implementing a multistage FMM [1,2], the complexity can be further reduced to \(O(N\log N)\). However, even for problems that have an order of magnitude more variables than those currently tractable using dense-matrix techniques (\(N \approx 10^7\)), we estimate that the performance of the single-stage algorithm should be near optimal.

Our development is given in terms of the method of moments (MoM), rather than the Nyström method [5]. We do this because

- Electrical engineers are more familiar with the MoM, and may therefore be more comfortable with the development.

- The prescription we present is sufficiently simple that it can be easily retrofitted to existing MoM codes.

- When used in the MoM, detailed comparisons to verify that results are identical to dense-matrix techniques are immediately available.

- We avoid all questions of singularity subtraction, as it is required only for matrix elements representing nearby interactions, and the computation of these is unchanged when the FMM is employed.

- The presentation demonstrates the independence of the FMM from the choice of discretization method, boundary-surface model, basis functions, etc.

The reader is cautioned not to interpret our choice of presentation as representing a preference toward the MoM. On the contrary, we think that the Nyström method is the appropriate tool for efficient and accurate boundary-integral-equation solvers.

For the purposes of demonstration, we first consider the MoM for the scalar wave equation, with Dirichlet boundary conditions on the surface of a scatterer. This is done for notational convenience only, the (naïve) equivalent application to the electric-field integral equation (EFIE) being straightforward (One can simply apply the scalar prescription to each Cartesian component of the vector expansion functions, and to their divergences; a more efficient method is described in Section 5.)

If the structure of this article seems somewhat confusing at an initial reading, it is because some considerations are intentionally delayed. We hope that the reasons for this become clear upon subsequent readings. In Section 2, we define notation, introduce the discretization of the scattering problem, relate the FMM to a more familiar fast algorithm, and introduce the fundamental analytic apparatus of the FMM. A detailed prescription for FMM implementation, except for the choice of some important parameters of the algorithm, is given in Section 3. After the structure of the method is exhibited, these parameters (the number of terms used in the multipole expansion, and the directions at which far-field quantities are tabulated) are analyzed in Section 4. The algorithm for the scalar problem then having been being completely defined, we exhibit the minor modifications necessary for application to vector (electromagnetic) scattering in Section 5. Before concluding, a physical interpretation of the analysis behind the FMM is given in Section 6.

2. Basics

2.1 Notation

Vectors in three-dimensional space are represented by boldface type (\(\mathbf{x}\)). The magnitude of a vector \(\mathbf{x}\) is written as \(|\mathbf{x}|\), unit vectors are written as \(\hat{\mathbf{x}} = \mathbf{x} / |\mathbf{x}|\), and integrals over the unit sphere are written as \(\int d^2\hat{\mathbf{x}}\). The imaginary unit is denoted by \(i\).
2.2 Time-independent scattering and the Method of Moments

A scattering problem [6, 7] can be defined by the scalar wave equation

$$\nabla^2 + k^2 \psi = 0,$$

where $Z'$, $V$, and $I$ are all sparse. As described in detail in this article, this allows computation of the product of $Z$ with an arbitrary vector (corresponding physically to the determination of the fields radiated by a known source distribution), with $O(N^{5/2})$ operations. The complexity can be further reduced to $O(N^{1/3}, N^{5/4}, \ldots)$ by recursive decomposition of $Z'$ and $V$:

$$Z' = Z'' + V'TV' \dagger,$$

$$V' = V'S.$$

This is entirely analogous to the FFT if one factors $F$ into only two factors (independent of $N$), the result would be an $O(N^{3/2})$ algorithm. We do not exhibit the details of the multi-stage FMM in this article.

In contrast to the FFT, the FMM decomposition is made possible by analytic rather than algebraic properties of the linear operator. Thus, while the FFT factorization is exact, the FMM decomposition is approximate. However, this does not constitute a practical limitation, as it is easy to control the FMM to achieve any desired level of precision (all the way to machine precision).

2.4 Identities

The FMM, as presented here, rests on two elementary identities. They, or formulas from which they may be easily derived, are found in many texts and handbooks on mathematical methods, such as Arfken [11] and Abramowitz and Stegun [12]. The first, an expansion of the kernel in the integral, Equation (3), for the impedance-matrix elements, is a form of Gegenbauer’s addition theorem,

$$e^{ik|x+d|} = \frac{1}{|x+d|} \sum_{j=0}^{\infty} (-1)^j (2j+1) j_i(\nu) j_k(\mu) j_l(kx) j_l(\mu)(\hat{d} \cdot \hat{X}).$$

where $j_\nu$ is a spherical Bessel function of the first kind, $h_\nu^1$ is a spherical Hankel function of the first kind, $l_i$ is a Legendre polynomial, and $d < X$. When using this expansion to compute the field at $x$ from a source at $x'$, $X$ will be chosen to be close to $x - x'$, so that $d$ will be small. This relationship of the various vectors is sketched in Figure 1. The special functions are as defined in [12]. The second is an expansion of the product $j_\nu(\hat{d} \cdot \hat{X})$ in propagating plane waves:

$$4\pi j_\nu(\hat{d} \cdot \hat{X}) = \int d^2 \hat{k} e^{i\hat{k} \cdot \hat{X}} j_\nu(\hat{k} \cdot \hat{X}).$$

Substituting Equation (12) into Equation (11), we get

$$e^{ik|x+d|} = \frac{ik}{4\pi} \int d^2 \hat{k} e^{i\hat{k} \cdot \hat{X}} \sum_{l=0}^{\infty} (2l+1) j_l(\nu)(kx) j_l(\mu)(\hat{k} \cdot \hat{X}).$$

Figure 1. The basic geometry, illustrating the relationship between the locations $x, x'$ and the displacements $x, x'$.
where we have performed the illegitimate but expedient interchange
of summation and integration. The key point is that we intend to
precompute the function

$$T_j(\kappa, \cos \theta) = \sum_{j=0}^{L} (-1)^j \psi_0(\kappa) \psi_j(\cos \theta), \tag{14}$$

for various values of $\kappa$. This is not a function in the $L \rightarrow \infty$, but
that need not concern us, as we obviously intend to truncate the
sum in numerical practice. The number of kept terms, $L + 1$, will
depend on the maximum allowed value of $kd$, as well as the desired
accuracy. The choice of $L$ is discussed in Section 4. It suffices, for
the present, to note that, in order to obtain accuracy from Equa-
tion (11), it must be slightly greater than $kd$, where $D$ is the maxi-
num value of $d$ for which the expansion will be used. Ignoring this
question for now (except for noting that the required number of
terms becomes small as $D \rightarrow 0$), we have

$$\frac{e^{i|X - d|}}{|X + d|} = \frac{ik}{4\pi} \int d^2 k e^{ik \cdot d} T_j(\kappa, \hat{k} \cdot \hat{X}). \tag{15}$$

Using this, the impedance-matrix element, Equation (3), is

$$Z_{m'm'} = \frac{k}{(4\pi)^2} \int d^2 x e^{ik(\kappa \cdot \kappa - X')} \int d^2 x' e^{ik(\kappa \cdot x' - X)} T_j(\kappa, \hat{k} \cdot \hat{X}) \tag{16}$$

In infinite-precision arithmetic, and in the limit of large $L$, this result
would be independent of the choice of $X$ (for $X > |x - x'| > X$). In
practice, one chooses $X$ to make $x - x'$ relatively small, so that
excellent accuracy can be obtained with a modest value of $L$. (That
this can be done by the grouping scheme described below is a con-
sequence of the local support of the basis functions.) Notice that
Equation (16) gives the impedance-matrix element (for well-
separated interactions) in terms of the Fourier transforms with wave
number $k$ of the basis functions, i.e. the basis functions’ field

The acceleration provided by the FMM comes from the fact that
these far fields can be grouped together before the integral over $k$
is performed.

3. Algorithmic prescription

3.1 Setup

1. Divide the $N$ basis functions into $M$ localized groups,
labeled by an index $m$, each supporting about $N/M$ basis
functions. (For now, $M$ is a free parameter. Later it will
be seen that the best choice will be $M = \sqrt{N}$.) Thus,
establish a correspondence between the basis-function index, $n$, and a pair of indices $(m, \alpha)$, where $\alpha$ labels the
particular basis function within the $m$th group. Denote the
center of the smallest sphere enclosing each group as
$X_m$. The grouping and index correspondence is shown,
for a simple case, in Figure 2.

2. For group pairs $(m, m')$ that contain “nearby” basis
functions [defined for now as those whose regions of support
are separated by a distance comparable to or
smaller than a wavelength, $2\pi / k$, so that Equation (16)
is valid], construct the sparse matrix $Z'$, with matrix elements

$$Z_{m'm'} = Z_{m(m', \alpha)(n'(m', \alpha))}. \tag{17}$$

by direct numerical computation of the matrix elements,
Equation (3). For all other pairs, $Z_{m'm'} = 0$.

This part of the matrix computation is identical to
what is conventionally done. All matrix elements, the
computation of which requires subtraction of singularities,
belong to $Z'$. If the large-$N$ limit is taken with a
fixed discretization interval and nearness criterion, this
step would require $O(N)$ computations. In Section 4,
we define nearby regions precisely, and it turns out that
their volume increases as $\sqrt{N}$, so that this step requires
$O(N^{3/2})$ computations.

3. For $K$ directions $\hat{k}$, compute the “excitation vectors”
(Fourier transforms of the basis functions)

$$V_{ma} (\hat{k}) = \int d^2 x e^{i(k \cdot x')} f_{m(\alpha)} (x'), \tag{18}$$

where $k$ is considered to be a parameter of the problem,
not a variable. Because $K$ needs to be chosen to give
accurate numerical quadrature for all harmonics to some
order $\propto L - kd, K \propto L^2, (kd)^2$, and because (from
geometrical considerations) $kd \propto \sqrt{N / M}$, this step
requires $O(N^{2/ M})$ computations.

4. For each pair $(m, m')$ for which $Z_{m'm'} = 0$ (regions
that are not nearby), compute the matrix elements

![Figure 2. The grouping for a simple surface. It is assumed, for purposes of illustration only, that each patch supports only one
basis function. The correspondence $n(m, \alpha)$ is abbreviated in Table 1.](image)

<table>
<thead>
<tr>
<th>$m$</th>
<th>$\alpha$</th>
<th>$n(m, \alpha)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>1</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>7</td>
</tr>
</tbody>
</table>
\[ T_{\text{mm}}(\hat{k}) = \frac{k}{(4\pi)^2} \sum_{l=0}^{L} (2l+1) j_l(kX_{\text{mm}}) \hat{p}_l(\hat{k} \cdot \hat{X}_{\text{mm}}) \]  

(19)

for the same \( K \) directions \( \hat{k} \) as the previous step, where \( l \ll kR \). If done in a naive manner, this computation requires \( O(KLM^2 - M^{12}N^{30}) \) operations. However, it can be accomplished more rapidly in a number of ways, the most elegant being the fast Legendre expansion [13].

### 3.2 Fast Matrix-Vector Multiplication

Rapid computation of the vector elements

\[ B_{m\alpha} = \sum_{n'\alpha'} Z_{mn'n'\alpha} f_{n'\alpha'} \]  

(20)

is accomplished by the following steps:

1. Compute the \( KM \) quantities

\[ s_m(\hat{k}) = \sum_\alpha s^*_m(\hat{k}) f_{m\alpha} \]  

(21)

which represent the far fields of each group \( m \). This step requires \( O(KN - N^2 / M) \) operations.

2. Compute the \( KM \) quantities

\[ s_m(\hat{k}) = \sum_{m'} s_{m'n}(\hat{k}) s_{m'\alpha}(\hat{k}) \]  

(22)

These represent the Fourier components of the field in the neighborhood of group \( m \), generated by the sources in the groups that are not nearby. This step requires \( O(KM^2 - KN) \) operations.

3. Finally, compute

\[ B_{m\alpha} = \sum_{n'\alpha'} Z_{mn'n'\alpha} f_{n'\alpha'} \]  

(23)

The first term is the standard MoM computation of near interactions, and the second term gives the far interactions, in terms of the far fields generated by each group. This step requires \( O(KN - N^2 / M) \) operations.

Straightforward substitution of Equations (18), (19), (21), and (22) into Equation (23), and of Equations (14)-(16) into Equation (20), shows that the two expressions for the vector \( B \), Equations (23) and (20), give equal results. Thus, computation of the vector \( B \) requires \( aMN + bN^2 / M \) operations, where \( a \) and \( b \) are machine and implementation dependent. The total operation count is minimized by choosing \( M = \sqrt{bN / a} \), the result is an \( O(N^{3/2}) \) algorithm.

### 4. Required number of multipoles and directions

In this section:

- We show how to choose the summation limit in the transfer function \( T_{\text{mm}}(\hat{k}) \), Equation (19), to achieve the desired accuracy (in the process, giving a precise definition of nearby regions).
- We discuss how to choose the \( K \) directions \( \hat{k} \), for the tabulation of angular functions.

One must choose \( L \) large enough that the multipole expansion of the Green's function, Equation (11), converges to the desired accuracy. As a function of \( l \), the Bessel functions \( j_l(z) \) and \( h_l^{(1)}(z) \) are of roughly constant magnitude for \( l < z \). For \( l > z \), \( j_l(z) \) decays rapidly and \( h_l^{(1)}(z) \) grows rapidly. While one must choose \( L > kd = k|x - x'| - X_{\text{mm}}| \) (so that the partial-wave expansion has converged), \( L \) cannot be taken to be much larger than \( kX_{\text{mm}} \), because the transfer function, Equation (14), will oscillate wildly, causing inaccuracies in the numerical angular integrations of Equations (15) and (23). This condition is a consequence of the interchange of summation and integration in Equation (13). An excellent semi-empirical fit to the number of multipoles required for single precision (32-bit reals) is

\[ L_d(kd) = kd + 5 \ln(kd + 1) \]  

(24)

where \( D \geq 1 / k \) is the maximum \( d \) which will be required (the "diameter" of the basis-function groups). For double precision (64-bit reals), our estimate is

\[ L_d(kd) = kd + 10 \ln(kd + 1) \]  

(25)

If the \( L \) dictated by the appropriate formula exceeds \( kX_{\text{mm}} \), then the groups are too close to use the FMM, and their interaction must be represented in the sparse matrix \( Z' \).

The \( K \) directions \( \hat{k} \), at which the angular functions are tabulated, must be sufficient to give a quadrature rule that is exact for all spherical harmonics of order \( l < 2L \). A simple method [2] for accomplishing this is to pick polar angles \( \theta \) such that they are zeros of \( P_l(\cos \theta) \), and azimuthal angles \( \phi \) to be \( 2L \) equally spaced points. Thus, for this choice of \( \hat{k} = (\sin \theta \cos \phi, \sin \theta \sin \phi, \cos \theta) \), \( K = 2L^2 \). If more-reefficient quadrature rules for the sphere (of the type described by McLaren [14]) are used, then \( K = (4/3)L^2 \). Since \( kd \approx \sqrt{N / M} \), this justifies the assertion made in Section 3.1 that \( N \approx \sqrt{M / N} \).

### 5. Application to electromagnetic fields

In the solution of the electric-field integral equation, the impedance-matrix elements take the form [15]

\[ Z_{m'n'} = -i \sum_{l=1}^{3} \int_{\Gamma} d^2 x |d^2 x|^2 f_{m}(x) G_{m'n'}(x - x') f_{n'}(x') \]  

(26)

where

\[ G_{m'n'}(x - x') = \frac{1}{k^2} \frac{\hat{c}_r \cdot \hat{c}_l}{\hat{c}_m \cdot \hat{c}_n} \frac{e^{i(k|x - x'|)}}{4\pi |x - x'|} \]  

(27)

and the indices $j, j'$ label Cartesian components. As implied in Section 1, one can integrate by parts, and simply use the scalar prescription, given above, on the three components of $\mathbf{f}$ and the scalar $\mathbf{V} \cdot \mathbf{f}$. This is not, however, the most economical procedure. By differentiating with respect to $d$ under the integral in Equation (15), we get

$$G_j^j(X + d) = \frac{i k}{4 \pi} \int d^3 \hat{k} \left( \mathbf{\hat{d} \cdot \hat{k}} - \hat{k} \cdot \mathbf{\hat{d}} \right) \hat{k} \cdot \mathbf{d} \mathbf{J}_j(\mathbf{k}, \mathbf{\hat{k}} \cdot \mathbf{\hat{k}})$$

(28)

Now it can be easily seen that the scalar prescription presented in Section 3 can be modified to an electromagnetic one, by promoting the quantities $V_{max}$, $s_n(\mathbf{\hat{k}})$, and $g_{mn}(\mathbf{\hat{k}})$ to three-dimensional vectors, with

$$V_{max}(\mathbf{\hat{k}}) = \int d^2 \mathbf{\hat{x}} \ e^{i \mathbf{\hat{k}} \cdot \mathbf{\hat{x}}} \left[ r_{mn}(\mathbf{\hat{\alpha}})(\mathbf{x}) - \mathbf{\hat{k}} \cdot \mathbf{\hat{x}} \right] r_{mn}(\mathbf{\hat{\alpha}})(\mathbf{\hat{x}}) \right],$$

(29)

and using a dot product in the $\int d^2 \mathbf{\hat{k}}$ term of $B_{mn}$, Equation (23). This method can be implemented using about half the storage of the four-fold use of the scalar formula, because the vector $V_{max}$ has only two independent components: $\mathbf{\hat{k}} \cdot \mathbf{\hat{V}(\mathbf{\hat{k}})} = 0$.

6. Physical interpretation

The physics of the FMM rests on the following fact: given a field, $\psi(x)$, which satisfies the wave equation

$$(\nabla^2 + \kappa^2)\psi(x) = 0,$$

(30)

for all $x$ outside a given sphere, the field can be reconstructed everywhere outside that sphere from its far field [16, 17]. This means that if the field is radiated by a source density, $\rho(x)$, supported only within a sphere of radius $R$ centered at the origin,

$$\rho(x) = \int_{x \leq R} d^3 \mathbf{x} \ e^{i \mathbf{x} \cdot \mathbf{\hat{x}}}/(4\pi|x-x'|^2) \rho(x').$$

(31)

then the contribution of the "off-shell" ($q^2 = k^2$) components in the Fourier expansion of the Green's function [11],

$$e^{i \mathbf{q} \cdot \mathbf{x}}/4\pi|x-x'| = \int \frac{d^3 \mathbf{q}}{(2\pi)^3/2\pi} \left( \frac{2\pi}{q^2 - k^2 - i\varepsilon} \right),$$

(32)

(where $\varepsilon$ is an infinitesimal positive number, prescribing the correct treatment of the singularity at $q^2 = k^2$) are determined for $x \geq R$ (after integration over $d^3 \mathbf{x}$) by the radiation condition and the "on-shell" components. The on-shell components, coming from the residue of the pole at $q^2 = k^2$, give the imaginary part of the Green's function, and the off-shell components give the real part. It is important that the off-shell part is not determined by the on-shell part for $x < R$. This is related to the divergence of the series in Equation (11) for $d > X$. This interpretation explains why the far interactions can be computed [Equation (23)] from the radiation pattern $s_n(\mathbf{\hat{k}})$ of the nth group. It also clarifies why one only need keep two components in $\mathbf{V}$, $\mathbf{g}$, and $s$ for the electromagnetic case: the electromagnetic far field is transverse, and has only two independent components.

7. Conclusion

Present methods for computing radar and other scattering cross sections are limited by computer-processing and memory requirements. The significance of the increase in problem size made possible by the FMM can be illustrated by considering the calculation of RCS for X-band radar. With current methods, the size of the largest body that can be accurately modeled is a few feet. With the same computing resources, the techniques that we have described will increase this by at least an order of magnitude. Such computational capabilities would significantly reduce the technological risk of expensive projects employing stealth technology. They may likewise revolutionize other applications of scattering computations, such as high-frequency circuit modeling, sonar, and geophysical applications.

Because the FMM accelerates computation of the matrix-vector product $Z \cdot I$, and thus only indirectly solution of $Z \cdot I = V$, we are frequently asked about the relative merits of direct and iterative solutions, and techniques to reduce the iterations required in a conjugate-gradient type of solution. These are important questions, and are under study by us as well as many others. We consider them to be mostly beyond the scope of this article, but note that the FMM is compatible with "complexification," and with preconditioning by a sparse matrix.

Although we have only demonstrated the use of the FMM for surface-scattering problems, its application to volume-integral equations (necessary for the analysis of penetrable inhomogeneous scatterers) is obvious. When comparison to other techniques for computing the fields of volume source distributions is made, it should be noted that in this case the matrix $I$ in Equation (8) is a strict convolution, and as such can be applied by FFT, resulting immediately in an $O(N^2 \log N)$ algorithm, without further decomposition.

8. Acknowledgment

This research was supported by the Advanced Research Projects Agency of the US Department of Defense, and was monitored by the Air Force Office of Scientific Research under Contracts No. F49620-91-C-0064 and F49620-91-C-0084. The United States Government is authorized to reproduce and distribute reprints for governmental purposes notwithstanding any copyright notation hereon.

9. References


### Introducing Feature Article Authors

**Ronald Coifman** is a professor of mathematics at Yale University. He received his PhD from the University of Geneva, in 1965. He received his License ès Sciences Mathematiques in 1962. Prior to coming to Yale, Prof. Coifman was a professor at Washington University, a visiting professor at Tel-Aviv University, and a visiting assistant professor at the University of Chicago. Prof. Coifman’s recent publications have been in the areas of scattering and inverse scattering, nonlinear harmonic analysis, and wavelet theory. He was chairman of the Yale Mathematics Department from 1986 to 1989.

**Vladimir Rokhlin** has been a professor of computer science and mathematics at Yale University since 1985. He received a PhD in Applied Mathematics from Rice University, in 1983, and a MS in mathematics from Vilnius University, Vilnius, Lithuania, in 1973. Prior to joining the faculty at Yale, Prof. Rokhlin worked as a Senior Research Specialist at Exxon, a partner in Livshitz and Associates, Houston, TX, a consultant at Computer Systems, Houston, TX; and as a mathematician at the Institute of Arctic Geology, Leningrad, Russia. Prof. Rokhlin’s research has been in the areas of numerical-scattering theory, elliptic partial-differential equations, numerical solution of integral equations, quadrature formulas for singular functions, and numerical complex analysis.

**Stephen Wandzura** is a Senior Scientist, Optical Physics Laboratory, Hughes Research Laboratories. He received his BS in music from UCLA, in 1971, and his PhD in physics from Princeton University, in 1977. His research has been in diverse areas of scattering and propagation. His thesis research could be correctly, if somewhat whimsically, characterized as the study of the “impedance matrix of the proton.” As a National Research Fellow with the NOAA, he studied scattering of light by atmospheric turbulence, and the occurrence of mountain lee waves. At HRL, he has worked on classical and quantum optics, especially the theoretical and numerical study of stimulated scattering. For the last three years, he has been studying improvements to moment-method techniques. He has published articles in *Physical Review, Physical Review Letters, Physics Letters, Optics Letters, Nuclear Physics, Wave Motion*, and other journals.