Fast Multipole Methods for the Helmholtz Equation in Three Dimensions

A Volume in the Elsevier Series in Electromagnetism

Elsevier Science Internet Homepage – http://www.elsevier.com
Consult the Elsevier homepage for full catalogue information on all books, journals and electronic products and services.

Elsevier Series in Electromagnetism

(Series formerly known as Academic Press Series in Electromagnetism)

Edited by

ISAAK MAYERGOYZ, University of Maryland, College Park, USA

Electromagnetism is a classical area of physics and engineering that still plays a very important role in the development of new technology. Electromagnetism often serves as a link between electrical engineers, material scientists, and applied physicists. This series presents volumes on those aspects of applied and theoretical electromagnetism that are becoming increasingly important in modern and rapidly development technology. Its objective is to meet the needs of researchers, students, and practicing engineers.

Books Published in the Series

Giorgio Bertotti, *Hysteresis in Magnetism: For Physicists, Material Scientists, and Engineers*
Scipione Bobbio, *Electrodynamics of Materials: Forces, Stresses, and Energies in Solids and Fluids*
Alain Bossavit, *Computational Electromagnetism: Variational Formulations, Complementarity, Edge Elements*
M.V.K. Chari and S.J. Salon, *Numerical Methods in Electromagnetism*
Goran Engdahl, *Handbook of Giant Magnetostrictive Materials*
Vadim Kuperman, *Magnetic Resonance Imaging: Physical Principles and Applications*
John C. Mallinson, *Magneto-Resistive Heads: Fundamentals and Applications*
Isaak Mayergoyz, *Nonlinear Diffusion of Electromagnetic Fields*
Giovanni Miano and Antonio Maffucci, *Transmission Lines and Lumped Circuits*
Shan X. Wang and Alexander M. Taratorin, *Magnetic Information Storage Technology*
Isaak Mayergoyz, *Mathematical Models of Hysteresis and Their Applications*
Fausto Fiorillo, *Measurements and Characterization of Magentic Materials*

Related Books

John C. Mallinson, *The Foundations of Magnetic Recording, Second Edition*
Reinaldo Perez, *Handbook of Electromagnetic Compatibility*

Related Journals

A sample journal issue is available online by visiting the homepage of the journal (homepage details at the top of this page). Free specimen copy gladly sent on request. Elsevier Science Ltd, The Boulevard, Langford Lane, Kidlington, Oxford, OX5 1GB, UK

Fast Multipole Methods for the Helmholtz Equation in Three Dimensions

NAIL A. GUMEROV
RAMANI DURAISWAMI

University of Maryland
Institute for Advanced Computer Studies
College Park, Maryland
USA

2004

ELSEVIER

Amsterdam • Boston • Heidelberg • London • New York • Oxford
Paris • San Diego • San Francisco • Singapore • Sydney • Tokyo

ELSEVIER B.V.
Radarweg 29
P.O. Box 211, 1000 AE Amsterdam
The Netherlands

ELSEVIER Inc.
525 B Street, Suite 1900
San Diego, CA 92101-4495
USA

ELSEVIER Ltd
The Boulevard, Langford Lane
Kidlington, Oxford OX5 1GB
UK

ELSEVIER Ltd
84 Theobalds Road
London WC1X 8RR
UK

First edition 2004

Library of Congress Cataloging in Publication Data
A catalog record is available from the Library of Congress.

British Library Cataloguing in Publication Data
A catalogue record is available from the British Library.

ISBN: 0-08-044371-0
ISSN: 1567-2131

♾ The paper used in this publication meets the requirements of ANSI/NISO Z39.48-1992 (Permanence of Paper).
Printed in The Netherlands.

To our wives
(Larisa Gumerov and Shashikala Duraiswami),
Children and Parents

This page is intentionally left blank

Contents

Preface

Since Isaac Newton introduced a new descriptive method for the study of physics by using mathematical models for various physical phenomena, the solution of differential equations and interpretation of mathematical results have become one of the most important methods for scientific discovery in many branches of science and engineering. A century ago only mechanics and physics, and to a much smaller extent chemistry, enjoyed the use of the predictive and explanatory power of differential equations. At the end of the 20th century, mathematical models have become a commonplace in biology, economics, and many new interdisciplinary areas of science. The necessity for more accurate modeling and prediction, and the exponential growth and availability of computational capabilities has given rise to such disciplines as "computational physics", "computational chemistry", "computational biology", and more generally to "scientific computation." Contemporary engineering, physics, chemistry and biology actively use software for the solution of multidimensional problems. Material science, aerospace, chemical engineering, nuclear and environmental engineering, medical instrumentation—indeed, this list can be continued to include all sciences. Today, to a large extent modern technology depends on mathematical modeling and capabilities for the numerical solution of equations constituting these models.

The history of science knows many revolutions: the computational revolution at the end of the 20th century is closely related to the availability of cheap processing power through advances in electronics and materials science and improved algorithms and operating systems due to computer science and related disciplines. These have brought powerful desktop/laptop personal computers to researchers and engineers. These computers have sufficient speed and memory for the solution of such mathematical tasks as the three-dimensional boundary value problems for various partial differential equations. The availability of sophisticated front-end packages such as Matlab and Mathematica allows relatively naive users to access highly sophisticated algorithms, and makes simulation, and the analysis of simulation results a fundamental component of scientific discovery.

The computational capabilities of a modern computer, which children play with, is larger by several orders than the capabilities of huge mainframe computers and systems, exploited in the 1960s, 1970s, and even the 1980s. We should not forget that with such ancient computers humanity went to space, designed nuclear power stations, and experienced revolutions in science and technology of the 1960s. Mainframes and clusters of computers (supercomputers) of the end of the 20th century and the beginning of the 21st century have capabilities exceeding of those used just 10 years ago by orders of magnitude. Some limits for this growth are close enough today due to the limits of semiconductors and high-frequency electrical communications. However, new technologies based on new optical materials, optical switches, and optical analogs of semiconductor devices are under active research and development, which promise further growth of computational capabilities in the following decades. The exponential growth of computational power is captured in various "Moore's laws" named after the scientist Gordon Moore, a cofounder of Intel. In its original form [GM65], the law states that the number of components on a circuit doubles every 18 months. Today, this law is taken to mean that the capability of technology X doubles in Y months [K99].

Nevertheless, the evolution of computers (hardware) itself does not guarantee adequate growth of scientific knowledge or capabilities to solve applied problems unless appropriate algorithms (software) are also developed for the solution of the underlying mathematical problems. For example, for the solution of the most large-scale problems one needs to solve large systems of linear equations, which may consist of millions or billions of equations. Direct solution of a dense linear system for an $N \times N$ matrix requires $O(N^3)$ operations. Using this as a guideline we can say that the inversion of a million by million matrix would require about 10^{18} operations. The top computer in early 2004, the "Earth Simulator" in Yokohama, Japan, has a speed of about 36×10^{12} operations per second and would require about 8 h to solve this problem. If we were to consider a problem 10 times larger, this time would rise to about 1 year. It is impossible to conceive of using simulation as a means of discovery with direct algorithms even using such advanced computers.

Nevertheless, in many practical cases inversions of this type are routinely performed, since many matrices that arise in modeling have special structure. Using specially designed efficient methods for the solution of systems with such matrices, these systems can be solved in $O(N^2)$ or $O(N \log N)$ operations. This highlights the importance of research related to the development of fast and efficient methods for the solution of basic mathematical problems, particularly, multidimensional partial differential equations, since these solvers may be called many times during the solution of particular scientific or engineering design problems. In fact, improving the complexity of algorithms by an order of magnitude (decreasing the exponent by 1) can have a much more significant impact than even hardware advances. For a million variables, the improvement of the exponent can have the effect of skipping 16 generations of Moore's law!

It is interesting to observe how problems and methods of solution, which were formulated a century or two centuries ago, get a new life with advances in computational sciences and computational tools. One of the most famous examples here is related to the Fourier transform that appeared in the Fourier memoir and was submitted to public attention in 1807. This transform was first described in relation to a heat equation, but later it was found that the Fourier method is a powerful technique for the solution of the wave, Laplace, and other fundamental equations of mathematical physics. While used as a method to obtain analytical solutions for some geometries, it was not widely used as a computational method. A new life began for the Fourier transform only in 1965 after the publication of the paper by Cooley and Tukey [CT65], who described the Fast Fourier Transform (FFT) algorithm that enables multiplication of a vector by the $N \times N$ Fourier matrix for an expense of only $O(N \log N)$ operations as opposed to $O(N^2)$ operations. In practice, this meant that for the time spent for the Fourier transform of length, say, $N \sim 10^3$ with a straightforward $O(N^2)$ algorithm, one can perform the Fourier transform of a sequence of length $N \sim 10^5$, which is hundred times larger! Of course, this discovery caused methods based on the Fourier transform to be preferred over other methods, and revolutionized areas such as signal processing. This algorithm is described as one of the best ten algorithms of the 20th century [DS00].

Another example from these top ten algorithms is related to the subject of this book. This is an algorithm due to Rokhlin and Greengard [GR87] called the "Fast Multipole Method" (FMM). While it was first formulated for the solution of the Laplace equation in two and three dimensions, it was extended later for other equations, and more generally to the multiplication of $N \times N$ matrices with special structure by vectors of length N. This algorithm achieves approximate multiplication for expense of $O(\alpha N)$ operations, where α depends on the prescribed accuracy of the result, ε, and usually $\alpha \sim \log N + \log \varepsilon^{-1}$. For computations with large N, the significance of this algorithm is comparable with that of the FFT. While the algorithm itself is different from the FFT, we note that as the FFT did, it brings "new life" to some classical methods developed in the 19th century, which have not been used widely as general computational methods.

These are the methods of multipoles or multipole expansions, which, as the FFT, can be classified as spectral methods. Expansions over multipoles or some elementary factorized solutions for equations of mathematical physics were known since Fourier. However, they were used less frequently, say, for the solution of boundary value problems for complex-shaped domains. Perhaps, this happened because other methods such as the Boundary Element, Finite Element, or Finite Difference methods appeared to be more attractive from the computational point of view. Availability of a fast algorithm for solution of classical problems brought research related to multipole and local expansions to a new level. From an algorithmic point of view, the issues of fast and accurate translations, or conversions of expansions over different bases from one to the other have become of primary importance. For example, the issue of development of fast, computationally stable, translation methods and their relation to the structured matrices, for which fast matrix–vector multiplication is available, were not in the scope of 19th or 20th century researchers living in the era before the FMM. A more focused attention to some basic principles of multipole expansion theory is now needed with the birth of the FMM.

The latter sentence formulates the motivation behind the present book. When several years ago we started to work on the problems of fast solution of the Helmholtz equation in three dimensions we found a substantial lack in our knowledge on multipole expansions and translation theory for this equation. Some facts were well known, some scattered over many books and papers, and several things we had to rediscover by ourselves, since we did not find, at that time, the solution to our problems. A further motivation was from our desire to get a solution to some practically important problems such as scattering from multiple bodies and scattering from complex boundaries. Here again, despite many good papers from other researchers in the field, we could not find a direct answer to some of our problems, or find appropriate solutions (e.g. we were eager to have FFT-type algorithms for the translation and filtering of spherical harmonics, which are practically faster than our first $O(p^3)$ method based on a rotation–coaxial translation decomposition). We also found that despite a number of publications, some details and issues related to the error bounds and the complexity of the FMM were not worked out. In the present book we attempt to pay significant attention to these important issues. While future developments may make some of the results presented in this book less important, at the time of its writing, these issues are essential to the development of practical solvers for the Helmholtz equation using these fast algorithms.

This page is intentionally left blank

Acknowledgments

We would like to thank the National Science Foundation for the support of much of the research on which the book is based through NSF awards 0086075 and 0219681. We would also like to acknowledge the support from the University of Maryland Institute for Advanced Computer Studies for the preparation of the book. In addition, the department of Computer Science and the Applied Mathematics and Scientific Computing Program at the University provided us the opportunity to develop some of this material while teaching a graduate course on the Fast Multipole Method. In particular we are very grateful to Profs Joseph F. JaJa (Director, UMIACS) and Larry S. Davis (Chair, Computer Science) for their support and encouragement. We would also like to thank Prof. Isaak Mayergoyz for his encouragement in writing this book.

Finally, writing this book required an investment of time, that necessarily reduced the time we could have spent otherwise. We would like to thank our families for their love, support, encouragement and forbearance.

This page is intentionally left blank

Outline of the Book

The book is organized as follows.

Chapter 1: This is an introductory chapter whose main purpose is to present the scalar Helmholtz equation as a universal equation appearing in different areas of physics. Even though many problems are formulated in terms of systems of equations or are described by other well-known basic models, they can be reduced to the solution of the scalar Helmholtz equation using the scalar potentials and the Fourier or Laplace transforms. Here we also consider major types of boundary and transmission conditions and integral representation of solutions. Computation of the surface and volume integrals can be performed by discretization and reduction of the problem to summation of a large number of monopoles and dipoles. The rest of the book is dedicated to the solution of problems that arise from the scalar Helmholtz equation, whose solution can also be generalized to the summation of a large number of arbitrary multipoles.

Chapter 2: This chapter is dedicated to the fundamentals of the multipole and local expansions of the solutions of the Helmholtz equation. Most relations presented here are well known and one of the major goals of this chapter is to bring together in one place the necessary definitions and equations for easy reference. Another important goal is to establish the notation used in the book (because different authors use different functions under the same notation, e.g., spherical harmonics or "multipoles"). While the normalization factors to use may not seem important, our experience shows that one can spend substantial time to have a reliable analytical formula that can be used further. We introduce here the definition of the special functions used later in the book, and summarize useful relations for them.

Chapter 3: This is one of the key theoretical chapters. It introduces the concepts of reexpansion, translation, and rotation of solutions of the Helmholtz equation. Some equations and relations can be found in other sources while others are derived here for the first time. This chapter includes the basic concepts, structure of the reexpansion coefficients and special types and properties of these functions of vector argument. Since our major concern is the development and implementation of fast computational methods, we derive here some efficient methods for computations of the translation and rotation coefficients. While the explicit expressions for them via, say, Clebsch–Gordan coefficients, can be found elsewhere, these formulae are not practical for use in fast multipole methods. By designing and applying recursive methods, which allow one to compute all necessary coefficients spending not more than just a few operations for each of them, we achieve fast $O(p^4)$ and $O(p^3)$ translation methods, where p is the truncation number or bandwidth of functions used to approximate the solutions of the Helmholtz equation.

Chapter 4: The results of Chapter 3 can already be used for the solution of a number of problems of practical interest such as appearing in room acoustics and in scattering from multiple bodies. We identify the techniques used in this chapter as the "multipole reexpansion technique" or "multipole methods". In many cases this technique itself can substantially speed up solution of the problem compared to other methods (e.g. direct summation of sources or solution with boundary element methods). The purpose here is to show some problems of interest and provide the reader with some formulae that can be used for the solution of such complex problems as multiple scattering problem from arbitrarily shaped objects. This chapter comes before the chapters dedicated to fast multipole methods, and the methods presented can be speeded up further using the methods in subsequent chapters.

Chapter 5: In this chapter we introduce Fast Multipole Methods (FMM) in a general framework, which can be used for the solution of different multidimensional equations and problems, and where the solution of the Helmholtz equation in three dimensions is just a particular case. We start with some basic ideas related to factorization of solutions. We describe how rapid summation of functions can be performed. Next, we proceed to modifications of this basic idea, such as the "Single Level FMM", and the "Multilevel FMM", which is the FMM in its original form. While there exist a substantial number of papers in this area that may be familiar to the reader, we found that the presentation in these often obscure some important issues, which are important for the implementation of the method and for its understanding. This method is universal in a sense that it can be formalized and applied to problems arising not only in mathematical physics. One of the issues one faces is the data structures to be used and efficient implementation of algorithms operating with a large amount of data used in the FMM. This is one of the "hidden" secrets of the FMM that usually each developer must learn. We provide here several techniques based on spatial ordering and bit interleaving that enable fast "children" and "neighbor" search procedures in data organized in such structures as octrees. These techniques are known in areas which are not related to mathematical physics, and we tried to provide a detailed insight for the reader who may not be familiar with them.

Chapter 6: While one can consider the FMM for the Helmholtz equation as a particular case of a generalized FMM procedure, it has some very important peculiarities. In the form originally introduced by Rokhlin and Greengard for the solution of the Laplace equation, the FMM is practical only for the so-called "low-frequency" problems, where the size of the computational domain, D_0, and the wave number, k, are such that $kD_0 < A$, where A is some constant. While this class of problems is important, it prevents application of the FMM for "high-frequency" problems, which are equally important. The method to efficiently to solve these problems is to vary the truncation number with the level of hierarchical space subdivision. To illustrate this we introduce a model of the FMM for the Helmholtz equation, and derive several important theoretical complexity results. One of the basic parameters of this model is a parameter we call the "translation exponent" that characterizes the complexity of translations for some given truncation number. We also introduce some concepts such as the "critical translation exponent", which separates the complexity of the method for higher frequencies from one type to the other. The critical value of the exponent depends on the dimensionality and "effective" dimensionality of the problem, which is determined by the non-uniformity of the spatial distributions of the sources and receivers. We also provide some optimization results and suggest a fully adaptive FMM procedure based on tree-structures, opposed to the pyramid data structures used in the regular FMM. This method was found to be useful for the solution of some "low-frequency" problems, while additional research is needed for other problems.

Chapter 7: This chapter is dedicated to the theory which underlies fast translation methods, and serves as a guide for further developments in this field. While providing substantial background theory, we focus here on two translation methods of complexity $O(p^3)$ which are based on rotation−coaxial decomposition of the translation operator and on sparse matrix decompositions of the operators. While the first method is known in the literature and can be applied to the decomposition of any translation for any space-invariant equation (which follows from the group theory), the second method is presented here, to the best of our knowledge, for the first time. This method can be derived from the commutativity properties of the sparse matrices representing differential operators and dense matrices representing translation operators. We implemented and tested both the methods and found them to be reliable and fast. While the first method seems to have smaller asymptotic constants and, so is faster, we believe that new research opportunities for fast translation methods are uncovered by the second method.

Chapter 8: In this chapter we consider both new and existing translation methods that bring the complexity of translations to $O(p^2 \log^{\alpha} p)$ with some α ranging from 0 to 2. They are based on the use of properties of structured matrices, such as Toeplitz or Cauchy matrices or on the diagonal forms of the translation and rotation operators. While some techniques developed over the last decade have been implemented and studied, this is still an active area for research. We have attempted to summarize and advance the knowledge in this area, though we are sure new fast techniques, filters, or transforms, will continue to be developed. We provide a link between the methods operating in the functional space of expansion coefficients and the methods operating in the space of samples of surface functions, where the transform from one space to the other can be done theoretically with $O(p^2 \log^{\alpha} p)$ complexity. We also present here some asymptotic results that can be used for the development of fast translation methods at low and high frequencies.

Chapter 9: One of the most important issues in any numerical method is connected with the sources of errors in the method, and bounds for these errors. This particularly relates to the FMM, where the error control is performed based on theory. There are several studies in the literature related to this issue for the Helmholtz equation, which are mostly concerned with proper selection of the truncation number for expansion of monopoles. Here we present some results from our study of the error bounds, which we extend to the case of arbitrary multipoles, and in addition establish the error bounds for the truncation of translation operators represented by infinite matrices. The theoretical formulae derived were tested numerically on some example problems for the expansion of single monopoles and while running the FMM for many sources. The latter results bring interesting findings, which should be theoretically explained by further studies. This includes, e.g. the error decay exponent at low frequencies, that shows that evaluations based on the "worst" case analysis substantially overestimate actual errors.

Chapter 10: In the final chapter we demonstrate the application of the FMM to the solution of the multiple scattering problem. We discuss this in details as well as some issues concerned with the iterative techniques combined with the FMM. Also we show how the FMM can be applied to imaging of the three-dimensional fields that are described by the Helmholtz equation. Finally, we present some results of numerical study of these problems including convergence of the iterative methods and overall method performance.

The book is written in an almost "self-contained" manner so that a reader with appropriate background in mathematics and computational methods, who, for the first time faces the problem of fast solution of the Helmholtz equation in three dimensions, can learn everything from scratch and can implement a working FMM algorithm. Chapter 8 is an exception, since there we refer to algorithms such as Fast Legendre Transform or Fast Spherical Filters, whose detailed presentation is not given, since it would require a special book chapter. As we mentioned, these algorithms are under active research, and so, if a beginning reader reaches this stage, we hope that he or she will be able to read and understand the appropriate papers from the literature that in any case may be substantially updated by that time.

An advanced reader can go directly to chapters or sections of interest and use the other chapters as reference for necessary formulae, definitions and explanations. We need to emphasize that while we have tried to use notations and definitions consistent with those used in the field, we found that different authors often define similar functions differently. As in any new work, at times we have had to introduce some of our own notations for functions and symbols, which are still not in common use. In any case we recommend that the reader be careful, especially if the formulae are intended to be used for numerical work, and follow the derivations and definitions presented carefully to avoid inconsistency with definitions in other literature.