
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 Magnetic 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



to search for more Elsevier books, visit the Books Butler at
<http://www.elsevier.com/homepage/booksbutler/>

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

© 2004 Elsevier Ltd. All rights reserved.

This work is protected under copyright by Elsevier Ltd, and the following terms and conditions apply to its use:

Photocopying

Single photocopies of single chapters may be made for personal use as allowed by national copyright laws. Permission of the Publisher and payment of a fee is required for all other photocopying, including multiple or systematic copying, copying for advertising or promotional purposes, resale, and all forms of document delivery. Special rates are available for educational institutions that wish to make photocopies for non-profit educational classroom use.

Permissions may be sought directly from Elsevier's Rights Department in Oxford, UK: phone (+44) 1865 843830, fax (+44) 1865 853333, e-mail: permissions@elsevier.com. Requests may also be completed on-line via the Elsevier homepage (<http://www.elsevier.com/locate/permissions>).

In the USA, users may clear permissions and make payments through the Copyright Clearance Center, Inc., 222 Rosewood Drive, Danvers, MA 01923, USA; phone: (+1) (978) 7508400, fax: (+1) (978) 7504744, and in the UK through the Copyright Licensing Agency Rapid Clearance Service (CLARCS), 90 Tottenham Court Road, London W1P 0LP, UK; phone: (+44) 20 7631 5555; fax: (+44) 20 7631 5500. Other countries may have a local reprographic rights agency for payments.

Derivative Works

Tables of contents may be reproduced for internal circulation, but permission of the Publisher is required for external resale or distribution of such material. Permission of the Publisher is required for all other derivative works, including compilations and translations.

Electronic Storage or Usage

Permission of the Publisher is required to store or use electronically any material contained in this work, including any chapter or part of a chapter.

Except as outlined above, no part of this work may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without prior written permission of the Publisher. Address permissions requests to: Elsevier's Rights Department, at the fax and e-mail addresses noted above.

Notice

No responsibility is assumed by the Publisher for any injury and/or damage to persons or property as a matter of products liability, negligence or otherwise, or from any use or operation of any methods, products, instructions or ideas contained in the material herein. Because of rapid advances in the medical sciences, in particular, independent verification of diagnoses and drug dosages should be made.

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.

Working together to grow
libraries in developing countries

www.elsevier.com | www.bookaid.org | www.sabre.org

ELSEVIER

BOOK AID
International

Sabre Foundation

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

This page is intentionally left blank

Contents

Preface	xvii
Acknowledgments	xxiii
Outline of the Book	xxv
Chapter 1. Introduction	1
1.1 Helmholtz Equation	1
1.1.1 Acoustic waves	1
1.1.1.1 Barotropic fluids	1
1.1.1.2 Fourier and Laplace transforms	3
1.1.2 Scalar Helmholtz equations with complex k	5
1.1.2.1 Acoustic waves in complex media	5
1.1.2.2 Telegraph equation	7
1.1.2.3 Diffusion	8
1.1.2.4 Schrödinger equation	8
1.1.2.5 Klein–Gordan equation	9
1.1.3 Electromagnetic Waves	10
1.1.3.1 Maxwell’s equations	10
1.1.3.2 Scalar potentials	11
1.2 Boundary Conditions	15
1.2.1 Conditions at infinity	15
1.2.1.1 Spherically symmetrical solutions	15
1.2.1.2 Sommerfeld radiation condition	16
1.2.1.3 Complex wavenumber	18
1.2.1.4 Silver–Müller radiation condition	19
1.2.2 Transmission conditions	21
1.2.2.1 Acoustic waves	21
1.2.2.2 Electromagnetic waves	23

1.2.3	Conditions on the boundaries	23
1.2.3.1	Scalar Helmholtz equation	24
1.2.3.2	Maxwell equations	24
1.3	Integral Theorems	25
1.3.1	Scalar Helmholtz equation	26
1.3.1.1	Green's identity and formulae	26
1.3.1.2	Integral equation from Green's formula for ψ	30
1.3.1.3	Solution of the Helmholtz equation as distribution of sources and dipoles	31
1.3.2	Maxwell equations	34
1.4	What is Covered in This Book and What is Not	37
Chapter 2. Elementary Solutions		39
2.1	Spherical Coordinates	39
2.1.1	Separation of variables	40
2.1.1.1	Equation with respect to the angle φ	41
2.1.1.2	Equation with respect to the angle θ	41
2.1.1.3	Equation with respect to the distance r	43
2.1.2	Special functions and properties	44
2.1.2.1	Associated Legendre functions	44
2.1.2.2	Spherical Harmonics	49
2.1.2.3	Spherical Bessel and Hankel functions	55
2.1.3	Spherical basis functions	57
2.1.3.1	The case $\text{Im}\{k\} = 0$	58
2.1.3.2	The case $\text{Re}\{k\} = 0$	59
2.1.3.3	The case $\text{Im}\{k\} > 0, \text{Re}\{k\} > 0$	60
2.1.3.4	The case $\text{Im}\{k\} < 0, \text{Re}\{k\} > 0$	60
2.1.3.5	Basis functions	61
2.2	Differentiation of Elementary Solutions	65
2.2.1	Differentiation theorems	66
2.2.2	Multipole solutions	71
2.3	Sums of Elementary Solutions	73
2.3.1	Plane waves	73
2.3.2	Representation of solutions as series	74
2.3.3	Far field expansions	75
2.3.3.1	Asymptotic expansion	75
2.3.3.2	Relation to expansion over singular spherical basis functions	79
2.3.4	Local expansions	82
2.3.4.1	Asymptotic expansion	82

2.3.4.2	Relation to expansion over regular spherical basis functions	83
2.3.5	Uniqueness	86
2.4	Summary	86
Chapter 3. Translations and Rotations of Elementary Solutions		89
3.1	Expansions over Spherical Basis Functions	89
3.1.1	Translations	90
3.1.2	Rotations	92
3.2	Translations of Spherical Basis Functions	93
3.2.1	Structure of translation coefficients	97
3.2.1.1	Relation to spherical basis functions	97
3.2.1.2	Addition theorems for spherical basis functions	98
3.2.1.3	Relation to Clebsch–Gordan coefficients	98
3.2.1.4	Symmetries of translation coefficients	101
3.2.2	Recurrence relations for translation coefficients	105
3.2.2.1	Sectorial coefficients	108
3.2.2.2	Computation of translation coefficients	109
3.2.3	Coaxial translation coefficients	113
3.2.3.1	Recurrences	113
3.2.3.2	Symmetries	114
3.2.3.3	Computations	116
3.3	Rotations of Elementary Solutions	117
3.3.1	Angles of rotation	117
3.3.2	Rotation coefficients	121
3.3.3	Structure of rotation coefficients	126
3.3.3.1	Symmetries of rotation coefficients	128
3.3.3.2	Relation to Clebsch–Gordan coefficients	129
3.3.4	Recurrence relations for rotation coefficients	131
3.3.4.1	Computational procedure	135
3.4	Summary	137
Chapter 4. Multipole Methods		139
4.1	Room Acoustics: Fast Summation of Sources	139
4.1.1	Formulation	139
4.1.2	Solution	141
4.1.3	Computations and discussion	142
4.2	Scattering from a Single Sphere	143
4.2.1	Formulation	143
4.2.2	Solution	144

4.2.2.1	Determination of expansion coefficients	144
4.2.2.2	Surface function	146
4.2.3	Computations and discussion	148
4.3	Scattering from Two Spheres	150
4.3.1	Formulation	150
4.3.2	Solution	152
4.3.2.1	Determination of expansion coefficients	152
4.3.2.2	Surface function	156
4.3.3	Computations and discussion	157
4.4	Scattering from N Spheres	161
4.4.1	Formulation	161
4.4.2	Solution	161
4.4.3	Computations and discussion	165
4.5	On Multiple Scattering from N Arbitrary Objects	168
4.5.1	A method for computation of the T -matrix	169
4.6	Summary	170
 Chapter 5. Fast Multipole Methods		171
5.1	Preliminary Ideas	171
5.1.1	Factorization (Middleman method)	172
5.1.2	Space partitioning (modified Middleman method)	173
5.1.2.1	Space partitioning with respect to evaluation set	174
5.1.2.2	Space partitioning with respect to source set	177
5.1.3	Translations (SLFMM)	179
5.1.4	Hierarchical space partitioning (MLFMM)	183
5.1.5	Truncation number dependence	184
5.1.5.1	Geometrically decaying error	185
5.1.5.2	Dependence of the truncation number on the box size	186
5.1.6	Multipole summations	189
5.1.7	Function representations	190
5.1.7.1	Concept	190
5.1.7.2	FMM operations	192
5.1.7.3	SLFMM	194
5.2	Multilevel Fast Multipole Method	196
5.2.1	Setting up the hierarchical data structure	196
5.2.1.1	Generalized octrees (2^d trees)	196
5.2.1.2	Data hierarchies	199
5.2.1.3	Hierarchical spatial domains	200

5.2.1.4	Spatial scaling and size of neighborhood	201
5.2.2	MLFMM procedure	203
5.2.2.1	Upward pass	203
5.2.2.2	Downward pass	205
5.2.2.3	Final summation	207
5.3	Data Structures and Efficient Implementation	207
5.3.1	Indexing	208
5.3.2	Spatial ordering	211
5.3.2.1	Scaling	211
5.3.2.2	Ordering in one dimension (binary ordering)	212
5.3.2.3	Ordering in d dimensions	214
5.3.3	Structuring data sets	219
5.3.3.1	Ordering of d -dimensional data	220
5.3.3.2	Determination of the threshold level	221
5.3.3.3	Search procedures and operations on point sets	222
5.4	Summary	222
 Chapter 6. Complexity and Optimizations of the MLFMM		225
6.1	Model for Level-Dependent Translation Parameters	225
6.2	Spatially Uniform Data	227
6.2.1	Upward pass	228
6.2.1.1	Step 1	228
6.2.1.2	Step 2	228
6.2.1.3	Step 3	228
6.2.2	Downward pass	231
6.2.2.1	Step 1	231
6.2.2.2	Step 2	232
6.2.3	Final summation	233
6.2.4	Total complexity of the MLFMM	233
6.3	Error of MLFMM	236
6.4	Optimization	238
6.4.1	Lower frequencies or larger number of sources and receivers	239
6.4.2	Higher frequencies or smaller number of sources and receivers	241
6.4.2.1	Volume element methods	244
6.4.2.2	Some numerical tests	245
6.5	Non-uniform Data	248

6.5.1	Use of data hierarchies	248
6.5.2	Surface distributions of sources and receivers: simple objects	249
6.5.2.1	Complexity of MLFMM	250
6.5.2.2	Error of MLFMM	253
6.5.2.3	Optimization for lower frequencies or larger number of sources and receivers	253
6.5.2.4	Optimization for higher frequencies or smaller number of sources and receivers	255
6.5.2.5	Boundary element methods	257
6.5.3	Surface distributions of sources and receivers: complex objects	258
6.5.4	Other distributions	263
6.6	Adaptive MLFMM	264
6.6.1	Setting up the hierarchical data structure	265
6.6.1.1	General idea	265
6.6.1.2	Determination of the target box levels/numbers	266
6.6.1.3	Construction of the D -tree	267
6.6.1.4	Construction of the \tilde{D} -tree	268
6.6.1.5	Construction of the C -forest	268
6.6.2	Procedure	270
6.6.2.1	Upward pass	270
6.6.2.2	Downward pass	272
6.6.2.3	Final summation	273
6.6.3	Complexity and optimization of the adaptive MLFMM	273
6.6.3.1	Data distributions	274
6.6.3.2	High frequencies	281
6.7	Summary	283
Chapter 7. Fast Translations: Basic Theory and $O(p^3)$ Methods		285
7.1	Representations of Translation and Rotation Operators	286
7.1.1	Functions and operators	286
7.1.1.1	Linear vector spaces	286
7.1.1.2	Linear operators	287
7.1.1.3	Groups of transforms	288
7.1.1.4	Representations of groups	289
7.1.2	Representations of translation operators using signature functions	295
7.1.2.1	$(R R)$ translation	296

7.1.2.2	(S S) translation	298
7.1.2.3	S R translation	301
7.1.2.4	Coaxial translations	304
7.1.2.5	Rotations	305
7.2	Rotational–coaxial translation decomposition	306
7.2.1	Rotations	308
7.2.2	Coaxial translation	310
7.2.3	Decomposition of translation	311
7.3	Sparse matrix decomposition of translation and rotation operators	313
7.3.1	Matrix representations of differential operators	314
7.3.1.1	Operator D_z	316
7.3.1.2	Operator D_{x+iy}	317
7.3.1.3	Operator D_{x-iy}	319
7.3.1.4	Operator D_t	320
7.3.1.5	Matrix form of the Helmholtz equation	321
7.3.2	Spectra of differential and translation operators	322
7.3.2.1	Continuous spectra of differential operators	322
7.3.2.2	Continuous spectra of translation operators	323
7.3.3	Integral representations of differential operators	325
7.3.4	Sparse matrix decomposition of translation operators	326
7.3.4.1	Matrix exponential	326
7.3.4.2	Legendre series	329
7.3.5	Sparse matrix decomposition of rotation operators	330
7.3.5.1	Infinitesimal rotations	334
7.3.5.2	Decomposition of the rotation operator for Euler angle β	336
7.4	Summary	338
Chapter 8. Asymptotically Faster Translation Methods		339
8.1	Fast Algorithms Based on Matrix Decompositions	340
8.1.1	Fast rotation transform	340
8.1.1.1	Toeplitz and Hankel matrices	340
8.1.1.2	Decomposition of rotation into product of Toeplitz and diagonal matrices	343
8.1.2	Fast coaxial translation	348
8.1.2.1	Decomposition of translation matrix	349
8.1.2.2	Legendre transform	352
8.1.2.3	Extension and truncation operators	353
8.1.2.4	Fast coaxial translation algorithm	354

	8.1.2.5	Precomputation of diagonal matrices	355
8.1.3		Fast general translation	356
	8.1.3.1	Decomposition of the translation matrix	356
	8.1.3.2	Fast spherical transform	358
	8.1.3.3	Precomputation of diagonal matrices	360
8.2		Low- and High-Frequency Asymptotics	361
	8.2.1	Low frequencies	362
	8.2.1.1	Exponential sparse matrix decomposition of the $R R$ matrix	362
	8.2.1.2	Toeplitz/Hankel matrix representations	362
	8.2.1.3	Renormalization	366
	8.2.2	High frequencies	371
	8.2.2.1	Surface delta-function	371
	8.2.2.2	Principal term of the $S R$ translation	375
	8.2.2.3	Non-uniform and uniform asymptotic expansions	378
	8.2.2.4	Expansion of coaxial $S R$ matrix	380
	8.2.2.5	$R R$ translation	386
8.3		Diagonal Forms of Translation Operators	387
	8.3.1	Representations using the far-field signature function	389
	8.3.1.1	Spherical cubatures	389
	8.3.1.2	Signature functions for multipoles	391
	8.3.2	Translation procedures	395
	8.3.2.1	Algorithm using band-unlimited functions	395
	8.3.2.2	Numerical tests and discussion	396
	8.3.2.3	Deficiencies of the signature function method	400
	8.3.2.4	Algorithms using band-limited functions	403
	8.3.3	Fast spherical filtering	408
	8.3.3.1	Integral representation of spherical filter	408
	8.3.3.2	Separation of variables	410
	8.3.3.3	Legendre filter	413
8.4		Summary	416
Chapter 9. Error Bounds			417
9.1		Truncation Errors for Expansions of Monopoles	417
	9.1.1	Behavior of spherical Hankel functions	420
	9.1.2	Low frequency error bounds and series convergence	423

9.1.3	High frequency asymptotics	427
9.1.4	Transition region and combined approximation	431
9.2	Truncation Errors for Expansions of Multipoles	432
9.2.1	Low frequency error bounds and series convergence	436
9.2.2	High frequency asymptotics	438
9.3	Translation Errors	439
9.3.1	S S translations	439
9.3.1.1	Problem	439
9.3.1.2	Solution	440
9.3.2	Multipole-to-local S R translations	446
9.3.2.1	Problem	446
9.3.2.2	Solution	447
9.3.3	Local-to-local R R translations	452
9.3.3.1	Problem	452
9.3.3.2	Solution	453
9.3.4	Some remarks	456
9.3.5	FMM errors	458
9.3.5.1	Low and moderate frequencies	458
9.3.5.2	Higher frequencies	461
9.4	Summary	462

Chapter 10. Fast Solution of Multiple Scattering

	Problems	465
10.1	Iterative Methods	466
10.1.1	Reflection method	466
10.1.2	Generalized minimal residual and other iterative methods	470
10.1.2.1	Preconditioners	471
10.1.2.2	Flexible GMRES	471
10.2	Fast Multipole Method	472
10.2.1	Data structures	473
10.2.2	Decomposition of the field	475
10.2.3	Algorithm for matrix–vector multiplication	477
10.2.4	Complexity of the FMM	479
10.2.4.1	Complexity and translation methods for large problems	480
10.2.4.2	Smaller problems or low frequencies	481
10.2.5	Truncation numbers	482

10.2.6	Use of the FMM for preconditioning in the GMRES	485
10.3	Results of Computations	487
10.3.1	Typical pictures and settings	487
10.3.1.1	FMM for spatial imaging/field calculation	488
10.3.1.2	Surface imaging	489
10.3.2	A posteriori error evaluation	490
10.3.3	Convergence	493
10.3.4	Performance study	495
10.4	Summary	498
Color Plates		499
Bibliography		509
Index		515

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 multi-dimensional 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 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 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.